Model dugotrajne pohrane digitalno potpisanoga arhivskoga gradiva

Bralić, Vladimir

Sveučilište u Zagrebu

Filozofski fakultet

Vladimir Bralić

Model dugotrajne pohrane digitalno potpisanoga arhivskoga gradiva

DOKTORSKI RAD

Zagreb, 2023.

Sveučilište u Zagrebu

Filozofski fakultet

Vladimir Bralić

Model dugotrajne pohrane digitalno potpisanoga arhivskoga gradiva

DOKTORSKI RAD

Mentor:
dr. sc. Hrvoje Stančić, red. prof.

Zagreb, 2023.

University of Zagreb

Faculty of humanities and social sciences

Vladimir Bralić

A model for long term preservation of digitally signed archival records

DOCTORAL THESIS

Supervisor:
Ph. D. Hrvoje Stančić, Full Professor

Zagreb, 2023.

INFORMACIJE O MENTORU

Dr. sc. Hrvoje Stančić, red. prof. rođen je 2. listopada 1970. u Zagrebu gdje je završio osnovnu i srednju školu. Diplomirao je 1996. godine studijske grupe Informatologija (smjer Opća informatologija) i Engleski jezik i književnost na Filozofskom fakultetu u Zagrebu. Godine 1996. prihvaćen je kao znanstveni novak na projektu koji se vodi na Odsjeku za informacijske znanosti Filozofskog fakulteta u Zagrebu (broj znanstvenika 244003). Magistrirao je 2001. godine s temom Upravljanje znanjem i globalna informacijska infrastruktura. Doktorirao je 2006. s temom Teorijski model postojanog očuvanja autentičnosti elektroničkih informacijskih objekata. Od ožujka 2018. je u zvanju redovnoga profesora.

Urednik je knjige "Trust and Records in an Open Digital Environment", koautor "Arhivističkog rječnika. Englesko-hrvatskog, hrvatsko-engleskog", autor knjige "Digitalizacija", koautor knjige "Heritage Live. Upravljanje baštinom uz pomoć informacijskih alata", urednik pet zbornika međunarodne konferencije INFuture – The Future of Information Sciences. Objavio je, samostalno ili u koautorstvu, više od 80 znanstvenih i stručnih radova, te je vodio izradu 20 disertacija.

Kao istraživač sudjelovao je u radu četiriju nacionalnih znanstveno-istraživačkih projekata i jednom međunarodnom TEMPUS projektu. Vodio je, na razini fakulteta, međunarodni projekt HERITAGE Live koji se odvijao u okviru IPA operativnog programa prekogranične suradnje Slovenija-Hrvatska (2007.-2013.). Na razini Hrvatske koordinirao je aktivnosti europske koordinacijske inicijative Digital Preservation Europe – DPE. Na projektu InterPARES Trust (2013.-2019.) bio je voditelj europskog istraživačkog tima, a na projektu InterPARES Trust AI (2021.-2026.) član Istraživačkog odbora.

U Ministarstvu kulture RH bio je ili jest član više različitih radnih skupina. Član je Hrvatskog informacijskog i dokumentacijskog društva (HIDD), član Predsjedništva Hrvatskog arhivističkog društva (HAD) te član Centre for the International Study of Contemporary Records and Archives (CISCRA) na Sveučilištu Britanske Kolumbije, Vancouver, Kanada.

U Hrvatskom zavodu za normizaciju predsjednik je tehničkog odbora za razvoj međunarodne norme ISO/TC 307 Ulančani blokovi i tehnologija distribuirane glavne knjige (engl. Blockchain and Distributed Ledger Technologies). Dobitnik je dviju medalja za inovacije vezane uz *blockchain* na međunarodnim izložbama inovacija – brončane na izložbi ARCA 2021 te srebrne na izložbi ARCA 2022.

SAŽETAK

U radu je objašnjena teorijska osnova temeljnih pojmova područja arhivistike i diplomatike te se na osnovu njih ukazuje na nedostatnost suvremenih informacijskih sustava za dugoročno očuvanje (digitalnog) arhivskog gradiva, posebno digitalno potpisanih dokumenata. U radu je pokazano da suvremeni sustavi ne uzimaju u obzir temeljne arhivističke pojmove autentičnosti i arhivske veze na način na koji to potrebno. Rad pristupa ovom problemu s pretpostavkom da ga je moguće riješiti upotrebom ulančanih zapisa te je provedeno komparativno istraživanje postojećih sustava i tehnologija. U konačnici, na temelju provedenog istraživanja, razvijen je novi model informacijskog sustava za dugotrajno očuvanje digitalno potpisanoga arhivskoga gradiva koji je u skladu s temeljnim zahtjevima arhivistike i diplomatike.

Ključne riječi: autentičnost, arhivsko gradivo, arhivska veza, digitalni zapisi, digitalni potpis, digitalni certifikat, vremenski žig, dugotrajna pohrana, ulančani blokovi

PROŠIRENI SAŽETAK

Cilj ovog istraživanja je istražiti probleme povezane s dokazivosti autentičnosti dugotrajno pohranjenoga digitalnog arhivskog gradiva, a posebno digitalno potpisanih zapisa pohranjenih u arhivskim institucijama i probleme očuvanja arhivske veze u sustavima digitalnih arhiva te predložiti model rješenja prepoznatih problema. Motivacija za istraživanje proizlazi iz prethodnih istraživanja koja su ukazala na problem isteka digitalnih certifikata kojima je potpisana arhivska građa. Nakon isteka digitalnog certifikata o kojem je ovisan digitalni potpis gubi se mogućnost dokazivanja identiteta autora nekog zapisa te se na taj način narušava autentičnost zapisa. S obzirom na to da suvremeni potpisni digitalni certifikati imaju rok trajanja od dvije do tri godine a zakonska regulativa zahtjeva od arhivskih ustanova da zapise čuvaju značajno duže ova situacija stvara probleme pravne prirode jer je narušena mogućnost korištenja zapisa kao dokaza nekog događaja iz prošlosti, ali je i u sukobu s temeljnim zahtjevima arhivistike za očuvanjem autentičnosti arhivskog gradiva.

Postojeća rješenja ovog problema svode se na dodavanje novih potpisa ili upotrebu vremenskih žigova. Vremenski žigovi imaju značajno duže trajanje od digitalnog potpisa ali u osnovi pate od istog problema, ni oni ne traju zauvijek, a uz to sustavi temeljeni na njima pri aplikaciji žiga ne provjeravaju ispravnost potpisnog certifikata te se može tvrditi da vremenski žig granatira samo da podaci nisu promijenjeni od trenutka dodavanja žiga. Identitet autora je u ovakvom rješenju uglavnom zanemaren. Osim toga, specijalizirani arhivski sustavi moraju osigurati održavanje arhivske veze, što je još jedan element kojeg postojeća rješenja zanemaruju.

Zbog svega navedenog u sklopu InterPARES Trust (2013.-2019.) međunarodnog projekta započet je razvoj novog modela za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva. Razvoj ovog modela završen je u ovoj disertaciji.

Istraživanje u ovoj disertaciji započelo je na način da je odrađena analiza temeljnih zahtjeva arhivistike i diplomatike na očuvanje i dokazivost autentičnosti, uključujući arhivsku vezu. Cilj ovog istraživanja literature iz arhivistike i diplomatike je pokazati što arhivistika zahtjeva od sustava za dugotrajnu pohranu. Ukratko, autentičnost ovisi o mogućnosti dokazivanja identiteta autora i integriteta dokumenta od trenutka stvaranja, a arhivska veza podrazumijeva da je očuvan kontekst u kojem je zapis nastao, to jest mreža njegovih odnosa veza s drugim zapisima. Na osnovu ovih saznanja moguće je razviti novi model koji je prilagođen arhivskim svrhama i uzima u obzir suvremene probleme vezane uz digitalni potpis.

Osim zahtjeva arhivistike u obzir su uzeti i zahtjevi regulatornih tijela Republike Hrvatske. Zajedno, ovi zahtjevi definirali su funkcionalnosti novog modela.

Uz zahtjeve arhivistike razmotrene su i kriptografski algoritmi te industrijski standardi koji omogućavaju digitalni potpis. Ovo prije svega uključuje *hash* algoritme i asimetrične kriptografske algoritme te standarde koji reguliraju sadržaj potpisnih digitalnih certifikata i način dodavanja digitalnog potpisa raznim zapisima. Razumijevanje ovih (kriptografskih) algoritama i normi pruža zorni dokaz mogućnosti upotrebe digitalnog potpisa i povezanih sustava za dokazivanje autentičnosti prema zahtjevima arhivistike.

Istraživanje je obuhvatilo i pregled postojećih rješenja za vremenske žigove te analiziralo njihove prednosti i mane, objasnilo zbog čega u potpunosti ne ispunjavaju zahtjeve te razmotrilo u kojoj mjeri ovi sustavi mogu sudjelovati u novom modelu. Također, razmotreno je na koji način mogu biti primijenjene tehnologije na kojima se temelje u razvoju novog modela.

Novi model, naziva TrustChain, temeljen je na principu ulančanih zapisa, koji koriste i mnogi sustavi za vremenske žigove. Model je prilagođen potrebama arhivskih ustanova te uzima u obzir zahtjeve arhivistike i zakonske zahtjeve, prije svega one vezane uz zaštitu osobnih podataka. Model omogućava dugotrajno očuvanje autentičnosti digitalno potpisane arhivske građe bez potrebe za naknadnim dodavanjem potpisa ili vremenskih žigova. Osim toga, model je prilagođen i za dugotrajno očuvanje digitalnih potpisnih certifikata. Tijekom istraživanja, većim dijelom iz razloga povjerljivosti dijelova arhivske građe, pokazala se potreba i za ovom funkcionalnosti.

Ključne riječi: autentičnost, arhivsko gradivo, arhivska veza, digitalni zapisi, digitalni potpis, digitalni certifikat, vremenski žig, dugotrajna pohrana, ulančani blokovi

EXTENDED SUMMARY

The goal of this research is to explore problems related to the ability to prove the authenticity of digital archival records and especially digitally signed records in archival institutions as well as problems related to the preservation of the archival bond and to propose a solution to the identified problems. This research was motivated by previous research which has highlighted the problem of digital certificate expiry when such certificates are used to digitally sign archival records. After digital certificate expiry on which digital signatures are dependant, the ability to prove the identity of the records author is lost and the records authenticity is compromised. Contemporary digital signing certificates have a lifespan of two to three years and there is a legal requirement that archival institutions preserve their records for much longer. This discrepancy not only causes legal problems because the ability to use the record as proof of past events is compromised but also clashes with basic archival science requirements regarding archival records authenticity preservation.

Existing solutions to this problem include subsequent resigning of documents and the application of digital timestamps. Digital timestamps have a significantly longer lifespan then digital signatures but suffer from the same basic problem – just like digital signatures they do not last forever. In addition to this, timestamp systems do not check or otherwise account for the identity of the author when applying the timestamp. These systems guarantee that the record has remained unchanged since the timestamp application, but the identity of the record's author is largely ignored by timestamp systems. A further problem arises form the need to maintain archival bonds, an archival science requirement which is ignored by most systems.

These issues have prompted the development of a new model for long term preservation of digitally signed documents. This research began as part of the InterPARES Trust international project (2013-2019) and is concluded in this dissertation.

The research began with an analysis of archival and diplomatic requirements concerning the long term preservation and the ability to prove records authenticity, including archival bond preservation. The goal of this research into archival and diplomatic literature was to establish archival science requirements for long term record preservation. Summarily said, authenticity is dependent on the ability to prove the identity of a record's author and the integrity of the record since the moment of its creation and archival bond preservation requires that the full information relating to the context in which a document was created and its network of connections to other documents are preserved. Based on these conclusions it is possible to

develop a new model which accounts for archival requirements and addresses the problems related to contemporary digital signatures. Besides archival requirements, legal requirements in the Republic of Croatia were also considered. Together these requirements define the functionalities of the new model.

Alongside archival requirements, cryptographic algorithms and industrial standards which govern the use of digital signatures were investigated and assessed. This includes hash algorithms, asymmetric cryptography algorithms and standards which regulate digital signature content and the application of digital signatures to records. Understanding these (cryptographic) algorithms and regulations provides clear proof of the ability to use the digital signature and related systems to prove authenticity in accordance with archival science requirements.

The research also considered existing timestamp systems and analysed their advantages and shortcomings, explained why these systems do not fully fulfil archival science requirements and to which degree the systems themselves or the technology on which they are based can be used in the new model.

The new model, titled TrustChain, is based on the blockchain data structure which is also used by most existing timestamp solutions. The new model has been adapted to the needs of archival institutions and considers archival science and legal requirements, primarily those connected to personal data protection. The new model enables long-term preservation of digitally signed archival records without the need for subsequent resigning or application of timestamps. In addition, the model has been adapted to long term preservation of digital certificates. The need for this feature has arisen during the conducted research and model development, mostly motivated by the confidentiality of certain archival records.

Keywords: authenticity, archival records, archival bond, digital records, digital signature, digital certificate, timestamp, long-term preservation, blockchain

SADRŽAJ

Uvodna napomena

Prije početka, smatram nužnim napomenuti da je značajan dio prvog poglavlja parafraziranje sadržaja Dr.Sc.-01[1] obrasca doktoranda. Taj obrazac se koristi kao temelj prijave teme doktorske disertacije i sadrži informacije koje je nužno i ovdje (doslovno) prenijeti, poput argumentacije i hipoteza istraživanja, ali nije nigdje javno objavljen. Ovi dijelovi obrasca čine temelj na osnovu kojega je prihvaćeno istraživanje te daju uvjete koje ono mora ispuniti. Iako je sadržaj tog obrasca većim dijelom parafraziran i proširen smatram da je, osim na mjestima gdje je drukčije navedeno, s obzirom na to da citiranje spomenutog obrasca nije navedeno, korektno prvo poglavlje (poglavlje "1. Uvod") smatrati citatom sadržaja Dr.Sc.-01 obrasca kojega sam osobno autor. Sadržaj izvornog obrasca ovdje je proširen i promijenjen u mjeri koja je bila nužna nakon provedenog istraživanja. Neke pretpostavke vezane uz, tada budući tijek istraživanja pokazale su se drugačijima za vrijeme istraživanja i to se očituje u ovim razlikama. Na primjer, potreba za istraživanjem nekih postojećih sustava i potreba za razvojem određenih funkcionalnosti u novom modelu su se naknadno pokazale suvišne. Ostatak disertacije (poglavlja 2-7) navodi prenesene ili parafrazirane dijelove prema Chicago standardu za citiranje s potpunom referencom u fusnotama.

Osim ovoga potrebno je raspraviti i korištenu terminologiju. Rad je pisan na hrvatskom jeziku ali se zbog područja istraživanja spominje značajan broj termina na engleskom jeziku (sve strane riječi u radu su na engleskom jeziku). Svi termini kod kojih je bilo potrebno navedeni su na engleskom jeziku ali se koristi hrvatski prijevod. Ovo nije slučaj s terminom engl. *hash* i engl. *padding*. Oboje su engleske riječi za koje ne postoji u potpunosti jednoznačni prijevod na hrvatski jezik te se koriste u engleskom obliku. To je obrazloženo u kasnijim poglavljima u kojima se sporni termini pojavljuju. Osim stranih termina ima smisla odmah raspraviti i način korištenja dva termina na hrvatskom jeziku. Termin "zapis" koristi se kao univerzalna oznaka individualne jedinice arhivske građe. U nekim slučajevima ovaj termin je zamijenjen s terminom "dokument". Ovo su situacije u kojima je zapis nužno u obliku digitalne datoteke koja je čitljiva čovjeku (na primjer PDF datoteke). Osim ovoga termin "dokument" se na nekim mjestima koristi kao oznaka fizičkog pisanog dokumenta (u slučaju diplomatike) i kao oznaka posebne podatkovne strukture (u slučaju MongoDB sustava za pohranu podataka).

[1] Obrazac Sveučilišta u Zagrebu koji se koristi pri prijavi teme doktorske disertacije. Obrazac ne postoji u otvorenom pristupu već je dostupan putem OBAD sustava, vidi: http://www.unizg.hr/istrazivanje/doktorski-studiji/obrasci-dr-sc-dr-art/znanstvena-podrucja-dr-sc/

1. Uvod

Istraživanje novog modela za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva započeto u sklopu rada međunarodne istraživačke skupine na projektu InterPARES Trust[2]. Ova istraživačka skupina, u kojoj su sudjelovali doktorand i mentor, između ostalih aktivnosti provela je istraživanje načina upravljanja digitalno potpisanim zapisima koja je rezultirala s tri studije slučaja:

1) studije slučaja digitalno potpisanih zapisa Središnjeg registra osiguranika koje čuva FINA (e-Regos),[3]
2) studije slučaja digitalno potpisanih zapisa od strane lokalnih vlasti u Švedskoj regiji Skåne,[4]
3) studije slučaja digitalno potpisanih zapisa u sustavu e-Porez koji koristi Porezna uprava RH.[5]

Ove studije slučaja uzele su u obzir postojeće zapise te zahtjeve ustanova i utvrdile jasnu potrebu za sustavom koji će omogućiti dugotrajnu pohranu digitalno potpisanih zapisa bez gubitka dokazivosti autentičnosti zapisa. Prema ovim zaključcima pojavila se motivacija za razvojem rješenja koje bi uzelo u obzir arhivističke zahtjeve. Postojeći sustavi za pohranu nisu stvarani uzimajući u obzir zahtjeve arhivistike ili nisu uzimali u obzir digitalni potpis te je već tijekom InterPARES Trust projekta započelo istraživanje novog modela koji bi ovo omogućio.

Dokazivost autentičnosti digitalno potpisanoga arhivskoga gradiva središnji je problem ovog istraživanja. Autentičnost je jedan od osnovnih pojmova arhivistike i diplomatike te prethodna istraživanja iz ovih područja pružaju teorijsku osnovu za istraživanje. Diplomatika je skup koncepata i metoda izvorno razvijenih tijekom 17. stoljeća, Duranti je definira kao "proučavanje autentičnosti, valjanosti i autoriteta [...] dokumenata pregledavanjem njegovih

[2] InterPARES Trust Project, https://interparestrust.org/
[3] InterPARES Trust Project. (2018). *TRUSTER Preservation Model (EU31) – Case Study 1 – digitally signed retirement fund records.* Preuzeto 1. 7. 2022. s https://interparestrust.org/assets/public/dissemination/TRUSTERPreservationModel(EU31)-CaseStudy1v1_2.pdf
[4] InterPARES Trust Project. (2018). *TRUSTER Preservation Model (EU31) – Case Study 2 – digitally signed e-tax records.* Preuzeto 7. 1. 2022. s https://interparestrust.org/assets/public/dissemination/TRUSTERPreservationModel(EU31)-CaseStudy2v1_2.pdf
[5] InterPARES Trust Project. (2018). *TRUSTER Preservation Model (EU31) – Case Study 3 – digitally signed medical records […] and minutes of meetings.* Preuzeto 7. 1. 2022. s https://interparestrust.org/assets/public/dissemination/TRUSTERPreservationModel(EU31)-CaseStudy3v1_3.pdf

raznih elemenata".[6] Arhivistika, znanost koja je proizašla iz diplomatike, razvila je metode za očuvanje dokumentalističkih i funkcionalnih odnosa među zapisima te načine upravljanja zapisima.[7] Prethodna istraživanja primijenila su ove principe na digitalne (u nekim starijim istraživanjima "elektroničke") zapise. Prema Duranti zapis se može smatrati autentičnim kada "zaista jest ono što tvrdi da jest",[8] to jest kada se može dokazati da nije bio mijenjan nakon trenutka njegovog nastanka. Uz očuvanje pouzdanosti, integriteta i uporabljivosti, očuvanje autentičnosti jedan je od osnovnih zahtjeva i u slučaju digitalnih sustava za dugotrajnu pohranu arhivskog gradiva, to jest digitalnih arhiva.[9]

Projekt InterPARES Trust, četvrti InterPARES projekt u čijem radu su sudjelovali doktorand i mentor, istražuje specifičnosti digitalnih zapisa iz perspektive arhivistike te je i sam definirano osnovne zahtjeve. Projekt InterPARES 1 postavlja dva zahtjeva pri dokazivanju autentičnosti digitalnoga arhivskog gradiva – integritet i identitet.[10] Integritet podrazumijeva kompletnost i nepromjenjivost gradiva, sukladno općoj definiciji integriteta podataka.[11] InterPARES projekt definira identitet kao skup "atributa zapisa koji ga jedinstveno karakteriziraju i po kojima se on razlikuje od ostalih zapisa".[12] Ovo uključuje imena osoba ili institucija koje su ga stvorile.

U slučaju digitalno potpisanoga dokumenta ove podatke sadrži digitalni potpisni certifikat. Svi suvremeni digitalni certifikati, iz tehničkih i sigurnosnih razloga, što je detaljnije raspravljeno u kasnijim poglavljima, su vremenski ograničeni. Većina potpisnih certifikata ističe nakon dvije ili tri godine. Kada certifikat digitalno potpisanoga dokumenta istekne više ga nije na jednostavan način moguće povezati s institucijom ili fizičkom osobom koja je potpisala dokument. Istekom valjanosti potpisnog certifikata onemogućena je dokazivost autentičnosti zapisa. Ovisno o državi i vrsti, arhivsko gradivo se u arhivima, u skladu s propisanim rokovima čuvanja, mora čuvati dugi niz godina, a u nekim slučajevima i trajno. Bez

[6] Duranti, L. (1989). Diplomatics: New Uses for an Old Science, Part I. *Archivaria*, 28, 7-27. Preuzeto 16 .12. 2021. iz https://www.researchgate.net/publication/225035619_Diplomatics_New_Uses_for_an_Old_Science
[7] Duranti, L., & Macneil, H. (1996). The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project. *Archivaria, 42*, 46-67. Preuzeto 7. 1. 2022. s https://archivaria.ca/index.php/archivaria/article/view/12153/13158
[8] Duranti, L. (1995). Reliability and Authenticity: The Concepts and Their Implications. *Archivaria, 39*, 5-10. Preuzeto 7. 1 2022 s https://archivaria.ca/index.php/archivaria/article/view/12063/13035
[9] Stančić, H. (2006). Teorijski model postojanog očuvanja autentičnosti elektroničkih informacijskih objekata. Zagreb: Filozofski fakultet. Preuzeto 7. 1. 2022. s https://www.bib.irb.hr/244465
[10] InterPARES Project. (2001). *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project.* Preuzeto 7. 1. 2022. s http://www.interpares.org/book/
[11] Boritz, J. E. (2003). *IS Practitioners' Views on Core Concepts of Information Integrity.* Preuzeto s https://web.archive.org/web/20111005085820/http://www.fdewb.unimaas.nl/marc/ecais_new/files/boritz.doc
[12] InterPARES Project. (2001), *The Long-term Preservation*, n. dj.

obzira na točno trajanje nužnih vremenskih perioda čuvanja arhivskog zapisa (koji su u slučaju Republike Hrvatske raspravljeni u drugom poglavlju jasno je da oni višestruko premašuju period valjanosti digitalnog certifikata.

Jedno od rješenja ovog problema je naknadno, kontinuirano, periodičko potpisivanje arhivskog zapisa. Takvo rješenje često nije moguće jer zahtjeva da arhiv od izvornog potpisnika arhivskog zapisa zatraži ponovno potpisivanje. Zbog jednostavnosti arhivi potpisuju dokumente vlastitim certifikatom ili posežu za vremenskim žigom u pokušaju zaobilaženja ovog nedostatka. Volarević i Stančić ističu da "vremenski žigovi mogu bezuvjetno garantirati vjerodostojnost i integritet (ali ne i autentičnost) nekoga zapisa, no potrebno je voditi brigu o njima. Oni zahtijevaju periodičnu provjeru i pravovremeno obnavljanje".[13] Izvorni vremenski žigovi temelje se na tehnologiji povezanih zapisa,[14] a kasniji sustavi uveli su nove koncepte, poput prolaznih ključeva (engl. *transient key*). Ipak, oni ne dotiču problem isteka certifikata u kontekstu dokazivosti autentičnosti arhivskih zapisa. Neki od sustava za vremenske žigove bazirani su na tradicionalnim PKI (engl. *public key infrastructure*) sustavima te su i sami podložni isteku certifikata i sigurnosnim rizicima te zahtijevaju naknadno periodičko korištenje žiga. Izbor karakterističnih primjera postojećih sustava za vremenske žigove i njihovih nedostataka dan je u petom poglavlju.

Osim autentičnosti, arhivski sustav za dugotrajnu pohranu mora osigurati i očuvanje arhivske veze. Arhivska veza je "mreža odnosa zapisa koji su nastali kao rezultat iste aktivnosti"[15]. Arhivska veza nastaje postepeno, od trenutka nastanka zapisa do trenutka nastanka zadnjeg povezanog zapisa te se može, prema Duranti, opisati i kao "izraz razvoja aktivnosti u kojoj dokument sudjeluje".[16] Ova aktivnost može trajati godinama i sigurno zahtjeva naknadne promjene metapodataka vezanih uz zapis.

---

[13] Volarević, I., & Stančić, H. (2016). Norme za elektroničke vremenske žigove i mogućnosti njihove primjene u arhivskoj struci. (S. Babić, Ur.) Arhivi i domovinski rat, 425-435. Preuzeto 18. 12. 2021. s https://www.researchgate.net/profile/Hrvoje-Stancic/publication/341287614_Norme_za_elektronicke_vremenske_zigove_i_mogucnosti_njihove_primjene_u_arhivskoj_struci_Standards_for_electronic_time_stamps_and_the_possibilities_for_their_application_in_archival_

[14] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document. *Journal of Cryptology, 3*(2), 99-111. doi: https://doi.org/10.1007/BF00196791

[15] InterPARES Project. (2016). *InterPARES 2 Project Glossary.* Preuzeto 16. 12. 2021. s http://www.interpares.org/ip2/ip2_term_pdf.cfm?pdf=glossary

[16] Duranti, L. (1997). The archival bond. *Archives and Museum Informatics, 11*, 213-218. Preuzeto 16. 12. 2021. s https://www.researchgate.net/publication/226554280_The_Archival_Bond

Uz arhivske zahtjeve za dokazivošću autentičnosti i održavanja arhivske veze, novi model mora uzeti u obzir i novu zakonsku regulativu – Opću uredba o zaštiti podataka.[17] Ova uredba Europske Unije postavlja nove zahtjeve na arhivske sustave, napose na one koji pohranjuju osobne podatke.

Digitalni arhivski zapisi čije je dugotrajno očuvanje predmet ovog istraživanja su digitalno potpisane datoteke, na primjer: ugovori, potvrde, izjave i slični dokumenti pohranjeni u opće prihvaćenim formatima, poput PDF datoteka ili u drugom digitalnom obliku koji je reguliran standardima, prije svega eIDAS uredbom.[18] Ova vrsta digitalnog potpisa temelji se na konceptu poznatom kao PKI, to jest na infrastrukturi javnog ključa[19] čija je upotreba u potpisivanju digitalnih zapisa regulirana industrijskim normama poput PAdES, CADdES i XAdES.

Kod digitalno potpisanih zapisa, u slučaju dugotrajne pohrane u digitalnom arhivu, javlja se ranije spomenuti problem s vremenskim ograničenjima potpisnih certifikata koji se koriste u digitalnim potpisima. Zbog razvoja sigurnosnih standarda, razvoja tehnologije i sigurnosnih rizika potpisni certifikati nužno se izdaju s vremenskim ograničenjem nakon kojeg postaju nevaljani te zahtijevaju periodičku nabavu novog potpisnog certifikata. Stoga novi model sustava za dugotrajnu pohranu mora omogućiti dokazivost autentičnosti i nakon isteka digitalnog certifikata. Osim toga model uzima u obzir značaj arhivske veze, to jest omogućava organizaciju odnosa zapisa unutar linearne, slijedne, strukture ulančanih zapisa te naknadne izmjene u metapodacima veznima uz zapis (da bi omogućio dodavanje novih podataka o arhivskim vezama).

Novi model temelji se na tehnologiji ulančanih zapisa (engl. *blockchain*) te nosi naziv TrustChain. Postojeća istraživanja pokazala su da je tehnologija ulančanih zapisa pogodna za razvoj sustava za dugotrajnu pohranu osnovnih informacija o arhivskim zapisima, ali da je za potrebe digitalno potpisanih zapisa i očuvanje arhivske veze među njima potrebno razviti potpuno novi model. Između ostalih, do ovih zaključaka došlo se i na temelju prethodnih istraživanja InterPARES Trust projekta, doktoranda i mentora te drugih istraživača. Novi model sadržavat će mogućnost očuvanja arhivske veze i uzeti u obzir zahtjeve regulatornog okvira

[17] European Parliament and Council. Regulation (EU) 2016/679 – General Data Protection Regulation. Preuzeto 27. 4. 2016. s https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504
[18] European Parliament. (2014). eIDAS. Preuzeto 4. 12. 2021. s https://www.eid.as/home/
[19] Adams, C., & Lloyd, S. (2003). Understanding PKI: concepts, standards, and deployment considerations. Boston: Addison-Wesley.

EU. Iako suvremeni sustavi bazirani na ulančanim zapisima i distribuiranoj glavnoj knjizi (engl. *distributed ledger*) ponekad omogućavaju pohranu zapisa, a gotovo uvijek pohranu metapodataka, oni ignoriraju ove, za arhivistiku važne, podatke koji omogućavaju pregled potpunog konteksta u kojem je zapis nastao.[20] Zbog ovoga je nužno razviti novi model za dugotrajnu pohranu podataka.

1.1. Cilj istraživanja

Cilj istraživanja je izrada novog modela informacijskog sustava za dugotrajnu pohranu digitalnog potpisanoga arhivskog gradiva temeljenog na tehnologiji ulančanih blokova, koji će omogućiti dokazivanje autentičnosti digitalno potpisanih zapisa s većom razinom pouzdanosti od postojećih sustava te istovremeno omogućiti očuvanje arhivske veze u skladu s temeljnim arhivističkim zahtjevima neovisno o propisanom roku čuvanja i isteku potpisnih digitalnih certifikata.

1.2. Hipoteze

S obzirom na cilj istraživanja postavljene su sljedeće hipoteze:

[1] Upotrebom ulančanih zapisa produžuje se dokazivost autentičnosti digitalno potpisanoga arhivskoga gradiva.

[2] Sustavi za pohranu digitalno potpisanoga arhivskoga gradiva utemeljeni na ulančanim zapisima omogućuju očuvanje arhivske veze.

Hipoteze će biti potvrđene (ili opovrgnute) teorijskim istraživanjem arhivistike, diplomatike i informacijskih tehnologija te razvojem novog modela sustava za pohranu digitalno potpisanoga arhivskoga gradiva.

[20] Lemieux, V. L., & Sporny, M. (2017). Preserving the Archival Bond in Distributed Ledgers: A Data Model and Syntax. *WWW '17 Companion Proceedings of the 26th International Conference on World Wide Web Companion*, (str. 1437-1443). Perth, Australia. Preuzeto 16. 12. 2021. s http://papers.www2017.com.au.s3.amazonaws.com/companion/p1437.pdf

1.3. Metodologija i sadržaj

Istraživanje je organizirano u tri faze:

1) Komparativno istraživanje teorijske osnove i zahtjeva arhivistike i diplomatike za očuvanjem autentičnosti (digitalnih) zapisa, povezanih zakonskih uredbi i postojećih tehnoloških rješenja te tehnoloških osnova ovih rješenja i samog digitalnog potpisa.

2) Komparativno istraživanje postojećih sustava za očuvanje autentičnosti digitalnih zapisa.

3) Razvoj novog modela za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva.

U prvoj fazi provelo se komparativno istraživanje postojeće literature i rezultata prethodnih istraživanja iz područja koja predstavljaju teorijsku podlogu za daljnje istraživanje. Cilj ove faze je utvrditi koji su zahtjevi arhivistike i diplomatike prema dugotrajno pohranjenim digitalno potpisanim arhivskim zapisima te koji su zahtjevi zakonodavnih tijela vezani uz autentičnost, dugoročno očuvanje autentičnosti i zaštite osobnih podatka. Osim ovoga cilj je i raspraviti u kojoj mjeri i na koji način postojeća rješenja iz područja informacijske tehnologije mogu pomoći u ispunjavanju ovih zahtjeva.

Proučena je literatura iz područja diplomatike i arhivistike koja definira pojmove integriteta i autentičnosti, literatura povezana uz koncept arhivske veze, literatura koja se općenito dotiče digitalnih arhiva, literatura iz područja informacijske tehnologije i računarstva iz područja algoritama i struktura podataka, prije svega kriptografije te povezani zakonski propisi. U skladu s postavljenim ciljevima literatura korištena u ovoj fazi može se podijeliti u četiri skupine:

1) teorijska skupina koja postavlja zahtjeve arhivistike za dugotrajnim očuvanjem digitalno potpisanih dokumenata i arhivske veze;

2) hrvatski i međunarodni zakonski propisi koji se bave arhivskim gradivom i zaštitom podataka (npr. Zakon o arhivskom gradivu i arhivima, Opća uredba o zaštiti podataka itd.);

3) literatura vezana uz teoriju informacijskih tehnologija na kojima su sustavi za digitalni potpis i vremenske žigove izgrađeni;

4) literatura koja se izravno bavi implementacijom tehnologija iz treće skupine s ciljem ostvarenja ciljeva definiranih literaturom iz prve skupine (ovo su prije svega industrijske norme iz područja digitalnog potpisa – ISO, ANSI, ETSI i RFC dokumenti).

Rezultati prve faze istraživanja sadržani su poglavljima dva, tri i četiri. Literatura iz točke 2 i 4 (arhivistička teorija i zakonski propisi) istraženi su u drugom poglavlju. Zaključci ovog poglavlja predstavljaju temeljne zahtjeve za novi model dugoročne pohrane digitalno potpisanih arhivskih zapisa. Poglavlja tri i četiri sadrže tehnološki dio teorijske osnove istraživanja. Poglavlje tri daje pregled relevantnih koncepata iz područja kriptografije i norme kojima su oni regulirani. U poglavlju četiri istražena je posebna podatkovna struktura, ulančani blokovi, koja čini temelj novog modela, a na nju se oslanjaju i postojeći sustavi. Zajedno ova tri poglavlja čine potpunu teorijsku osnovu za izradu novog modela.

U drugoj fazi istraživanja proučeni su postojeći sustavi te se analizirani i raspravljeni njihovi nedostaci. Svrha ovog dijela istraživanja je pokazati da postoji potreba za novim modelom arhivskih sustava za dugotrajnu pohranu digitalno potpisanoga gradiva. Poglavlje se fokusira na sustave za vremenske žigove koji predstavljaju trenutno rješenje problema kojim se bavi ovo istraživanje (dugoročna pohrana digitalno potpisanoga gradiva). Analiza ovih sustava ukazala je do koje mjere postojeći sustavi zadovoljavaju zahtjeve arhivistike te jesu li ovi sustavi (i njihove tehnologije) iskoristivi kao temelj ili dodaci modelu koji je i rezultat ovog istraživanja. Ova skupina uključuje sustave poput Enigio time:beat, Proofspace i AbsoluteProof. Iako je izvorno planirano, u teorijskom dijelu istraživanja se pokazalo suvišnim posebno analizirati sustave za digitalni potpis. Riječ je o sustavima koji omogućuju, to jest automatiziraju postupak digitalnog potpisivanja. U ranim fazama pretpostavljeno je da će istraživanje ovakvih sustava biti nužno da bi se utvrdile specifičnosti koje se mogu očekivati pri računalnoj obradi digitalno potpisanih zapisa koji dolaze iz različitih izvora. Digitalni potpis je u prvom dijelu istraživanja (teorijska komparativna analiza) analiziran do razine kodiranog računalnog zapisa te je pokazano da je zapis standardiziran, pa čak i propisan zakonskim uredbama (eIDAS i povezanim uredbama), do te mjere da nije potrebno proučavati individualne sustave. Digitalni potpis koji je relevantan u postupku dokazivanja autentičnosti je u potpunosti standardiziran, ne postoje razlike koje proizlaze iz, eventualno, različitog postupka njegovog stvaranja.

U trećoj fazi pristupilo se praktičnom dijelu istraživanja, to jest razvoju modela za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva s ciljem očuvanja autentičnosti i integriteta dokumenta. Izvorno je pretpostavljen razvoj četiri različita modela koji bi bili iskoristivi u različitim slučajevima. Oni su:

1) Model za pohranu digitalnih potpisa pregledom potpisanog dokumenta i provjerom certifikata u vlastitom ulančanom zapisu uz mogućnost očuvanja arhivske veze;
2) Model za pohranu digitalnih potpisa pregledom potpisanog dokumenta i provjerom certifikata u javnom ulančanom zapisu uz mogućnost očuvanja arhivske veze;
3) Model za pohranu digitalnih potpisa pregledom potpisanog dokumenta i provjerom certifikata u javnom ulančanom zapisu korištenjem sustava za pametne ugovore uz mogućnost očuvanja arhivske veze;
4) Model za pohranu digitalnog certifikata i certifikacijskog lanca u vlastitom ulančanom zapisu.

Međutim, tijekom teorijskog dijela istraživanja i tijekom ranih faza razvoja modela pokazalo se da sustavi temeljeni na javnim ulančanim zapisima i sustavi temeljenim na pametnim ugovorima nisu adekvatna rješenja. Umjesto toga razvijen je jedinstveni model koji uključuje dva različita ali povezana modula. Ovi moduli omogućuju:

1) Pohranu digitalnih potpisa pregledom potpisanog dokumenta i provjerom certifikata u vlastitom ulančanom zapisu uz mogućnost očuvanja arhivske veze;
2) Pohranu digitalnog certifikata pregledom ispravnosti certifikata i certifikacijskog lanca u vlastitom ulančanom zapisu.

Novi model, koji omogućuje ove dvije navedene funkcionalnosti, predstavljen je u šestom poglavlju. Model jasno pokazuje jesu li potvrđene hipoteze te se može iskoristiti kao osnova pri izradi projekta informacijskog sustava za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva. Odnos novog modela i hipoteza, to jest razina do koje novi model potvrđuje hipoteze, raspravljen je u zadnjem poglavlju.

1.4. Znanstveni doprinos

Znanstveni doprinos provedenog istraživanja ogleda se u prilagodbi temeljnih arhivskih principa digitalnoj okolini kroz novi model, naziva TrustChain, za dugotrajno očuvanje digitalno potpisanoga arhivskoga gradiva. Izrađeni model osigurava očuvanje autentičnosti i integriteta digitalno potpisanoga arhivskoga gradiva unatoč isteku valjanosti potpisnih certifikata te pritom osigurava očuvanje arhivske veze. Novi model razrađen je do razine na kojoj može biti iskorišten kao osnova projekta informacijskog sustava.

2. Teorijska osnova i zahtjevi arhivistike na digitalni potpise

Cilj ovog istraživanja je utvrditi postoji li mogućnost upotrebe ulančanih zapisa za produživanje dokazivosti autentičnosti digitalno potpisanih arhivskih zapisa, na osnovu koje će biti razvijen novi model za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva. Prvi korak u istraživanju je utvrditi što, u kontekstu arhivistike, podrazumijeva pojam autentičnosti (arhivskih zapisa). Iako su temeljni pojmovi arhivistike koji su predstavljeni u ovom poglavlju primjenjivi na svo arhivsko gradivo novi model odnosi se isključivo na digitalno arhivsko gradivo. Ovo proizlazi iz činjenice da se istražuje digitalno potpisano arhivsko gradivo. Digitalni potpis, koncept koji će biti detaljno opisan u idućem poglavlju, primjenjiv je samo na digitalne zapise.

Naravno, moguće je digitalno potpisati zapis koji je nastao u nekom drugom obliku, na primjer papirnatom, te je naknadno digitaliziran (i digitalno potpisan). Ovakvi zapisi se za potrebe ovog istraživanja tretiraju jednako kao i oni koji su izvorno nastali u digitalnom obliku. Dokazivanje i dugotrajno očuvanje autentičnosti nedigitalnog arhivskog gradiva temeljito je istraženo tijekom prošlosti i ono nije predmet ovog istraživanja. Svejedno, dokazivanje autentičnosti digitalno potpisanoga gradiva oslanja se na arhivske koncepte proizašle iz dugotrajnog očuvanja tradicionalnog arhivskog gradiva pa su i oni tema ovog poglavlja.

Definiranje zahtjeva za model novog informacijskog sustava koji će omogućiti dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva uz očuvanje dokazivosti autentičnosti je prvi korak u njegovoj izradi. Ovo poglavlje će definirati zahtjeve tako da će pratiti sljedeće korake:

1) U prvom koraku potrebno je definirati temeljene pojmove arhivistike i diplomatike. Osim što će razumijevanje ovih temeljnih termina pružiti osnovu za definiranje zahtjeva za novi model ono je važno i za razumijevanje razloga nedostatnosti postojećih sustava. Autentičnost je termin koji nije ekskluzivan arhivistici i diplomatici – u kombinaciji s integritetom, termin se koristi u raznim granama ljudskog djelovanja. Za ovo istraživanje posebno je važno utvrditi postoji li značajna razlika u korištenju termina između arhivističkih definicija i onih koje se pojavljuju u računarstvu. Ove razlike (ako postoje) su relevantne jer je razlika između postojećih sustava za dugotrajnu pohranu digitalnih podatka (ne nužno arhivskih zapisa) i specijaliziranih arhivskih sustava vjerojatno uzrokovana upravo razlikama u razumijevanju ovih termina. Gubitak

autentičnosti digitalno potpisanoga zapisa do kojeg dolazi nakon isteka valjanosti povezanog digitalnog certifikata je glavna motivacija ovog istraživanja te je važno razumjeti zbog čega je problem dobrim dijelom ignoriran od strane postojećih sustava.

2) Nakon definicije temeljnih pojmova poglavlje će dati pregled postojećih istraživanja iz područja dokazivanja autentičnosti, prvenstveno digitalnih, arhivskih zapisa. U posljednja dva stoljeća značajan broj istraživanja proveden je u području očuvanja autentičnosti digitalnih i elektroničkih zapisa. Neka od ovih istraživanja bit će razmotrena te će na osnovu njih biti moguće definirati kako se temeljni koncepti iz arhivistike i diplomatike primjenjuju na digitalne zapise.

3) Arhivska veza je arhivistički koncept kojem se u ovom istraživanju pristupa s posebnom pažnjom. Iako spada u temeljene arhivske pojmove, arhivska veza dobila je posebnu pozornost jer je istraživanje mogućnosti dugotrajnog očuvanja podataka o arhivskoj vezi osnova druge hipoteze ovog istraživanja. Arhivska veza je ključan koncept u arhivistici koji se često ignorira u sustavima za pohranu podataka koji nisu namijenjeni upotrebi isključivo u (digitalnim) arhivima. S obzirom na to da je cilj ovog istraživanja razvoj modela koji može biti korišten kao osnova za izgradnju informacijskog sustava koji podupire digitalne arhive, arhivska veza je od posebnog interesa u ovom istraživanju.

4) Kratko će biti raspravljeni i aktualni propisi Republike Hrvatske i Europske unije vezani uz arhivistiku i novi model. Ovo se prije svega odnosi za hrvatski Zakon o arhivskom gradivu i arhivima[21] te eIDAS[22] i GDPR[23] uredbe Europske unije. eIDAS (engl. *Electronic Identification, Authentication and Trust Services*) je dokument Europskog Parlamenta kojim se na području Europske unije regulira elektronički potpis (uključujući digitalne certifikate) i servise koji pružaju dokaze autentičnosti i integriteta digitalnih zapisa (uključujući vremenske žigove). Kao takav, eIDAS je od izuzetne važnosti za ovo

[21] Hrvatski sabor. (2018). *Zakon o arhivskom gradivu i arhivima*. Preuzeto 19. 11. 2022. s https://www.zakon.hr/z/373/Zakon-o-arhivskom-gradivu-i-arhivima
[22] European Parliament. (2014). eIDAS, n. dj.
[23] European Parliament and Council. (2016). *Regulation (EU) 2016/679 – General Data Protection Regulation.* Preuzeto 4. 12. 2021. s https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504

istraživanje. GDPR (engl. *General Data Protection Regulation)* regulira upotrebu i pohranu osobnih podataka na području Europske unije te i on mora biti uzet u obzir pri razvoju bilo kojeg sustava koji pohranjuje (osobne) podatke.

5) Konačno treba razmotriti arhivističke i srodne standarde vezane uz pohranu metapodataka. Na temelju ovih saznanja moguće je definirati set metapodataka koje će koristiti novi model. Ovo istraživanje započeto je u radu Stančića i Bralića iz 2021.[24] te je tada razvijen set nužnih metapodataka za novi model ovdje ponovno razmotren i dopunjen. Metapodaci se u novom modelu većim dijelom pohranjuju u promjenjivu, pomoćnu, bazu podataka (umjesto u nepromjenjiv lanac zapisa) koja se temelji na postojećim tehnologijama. Zbog ovoga, osim izmjene sadržaja metapodataka, naknadno je moguće mijenjati i strukturu ovih zapisa. Iz ovog razloga nije pri razradi modela nužno uključiti sve moguće metapodatke (koji ovise o slučaju upotrebe). Konkretni slučaj upotrebe bit će poznat u trenutku početka razvoja informacijskog sustava temeljenog na prikazanom modelu te je tada moguće korigirati set metapodataka. Svejedno, važno je da model od početka uključuje osnovni set metapodataka da bi se pojasnilo na koji način će se pretraživati podaci i da se omogući pohrana podataka o arhivskoj vezi (koja je metapodatak).

Na osnovu gore definiranih pet područja teorijskog istraživanja diplomatike i arhivistike bit će moguće definirati što se očekuje do modela za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva, to jest koji su zahtjevi za informacijski sustav koji se temelji na takvom modelu.

[24] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain: Overcoming the Limitations of Data Immutability. *Computers, 10*(8). doi: https://doi.org/10.3390/computers10080091

2.1. Načela diplomatike i arhivistike

U ovom poglavlju kratko će se raspraviti svrha diplomatike i arhivistike te njihova temeljna načela. Prije svega potrebno je pokazati što diplomatika i arhivistika podrazumijevaju pod pojmom autentičnosti zapisa i na koji način su ta načela primjenjiva na digitalno potpisanu arhivsku građu. Da bi razumjeli autentičnost i potrebu arhivistike (i diplomatike) za posebnim definicijama pojma potrebno je sagledati i povijest ovih disciplina.

Diplomatika je starija disciplina, te ona tvori temelj suvremene arhivistike. Duranti je 1989. dala izvrstan pregled povijesti diplomatike.[25] Ovaj pregled je temelj za početak rasprave o autentičnosti arhivskog gradiva.

Prema Duranti, diplomatika je disciplina koja usko vezna uz povijest i paleografiju.[26] Diplomatika se pojavljuje u 17. stoljeću kao skup pravila po kojima se procjenjuje autentičnost dokumenata. U ovom periodu u Njemačkoj se diplomatika koristila u sudskim raspravama vezanim uz razna prava. U Francuskoj su sukobi vođeni principima diplomatike poprimili oblik ideološkog sukoba oko činjenica o životima svetaca između benediktinskog i jezuitskog reda.[27] Sam termin je skovan 1681. od strane benediktinskog redovnika Jeana Mabillona koji prvi put opisuje disciplinu u svom djelu "De re diplomatica".

Već u prvim trenutcima pojavljivanja diplomatike, bez obzira govori li se o sudskim sporovima oko vlasništva nad zemljom ili ideološkim sukobima kršćanskih redova, sasvim je jasna primarna svrha ove discipline – dokazivanje autentičnosti, to jest razlikovanje originalnih, pravovaljanih, dokumenta od onih koji to nisu. Povijesno, diplomatika ovo postiže upotrebom raznih tehnika, poput proučavanja potpisa, pečata (i drugih posebnih oznaka), medija, načina te navedenog datuma i vremena stvaranja dokumenta. Iz ovih tehnika očituje se sličnost diplomatike i paleografije, te njena povezanost s povijesti.

[25] Duranti, L. (1989). Diplomatics, n. dj.
[26] Disciplina koja se bavi proučavanjem drevnih pisanih sustava te interpretiranjem i datiranjem povijesnih rukopisa. Definicija prevedena s engl. *paleography* prema Merriam-Webster rječniku: https://www.merriam-webster.com/dictionary/paleography
[27] Duranti, L. (1989). Diplomatics, n. dj.

Duranti kao najbolju definiciju discipline bira definiciju Gioriga Cencettia:

"Proučavanje stanja i stvaranja dokumentacije, analiza stvaranja, unutarnje konstitucije i prijenosa dokumenata te njihovog odnosa činjenica koje su u njima predstavljene i (odnosa) prema njihovim stvaraocima." (Cencetti, 1985)[28]

Duranti kasnije u svojem radu pojednostavljuje i pojašnjava definiciju na način:

"...diplomatika je disciplina koja proučava stvaranje, oblike i prijenos arhivskih dokumenata i njihov odnos prema činjenicama koje su u njima sadržane te s njihovim autorom, sa svrhom prepoznavanja, procijene i iskazivanja njihove prirode." (Duranti, 1989)[29]

U obje definicije naglašena je uloga autora, to jest odnos dokumenta i njegovog stvaratelja. Od svojeg najranijeg perioda diplomatika (i kasnije arhivistika) ima za zadaću dokazivanje autentičnosti nekog dokumenta da bi on mogao poslužiti kao temelj u nekom, često pravnom, sporu. U ovom kontekstu identitet autora bilo kojeg zapisa je od kritične važnosti. Ugovori, proglasi, pravilnici i zakoni nemaju smisla ako se ne može dokazati tko je njihov autor. Osim autora, vrlo je važno moći dokazati da je dokument ostao nepromijenjen od trenutka njegovog nastajanja do trenutka u kojem se koristi kao dokaz nečega. I sama Duranti u svom radu naglašava ovaj aspekt diplomatike:

"Izvor diplomatike striktno je povezan s potrebom utvrđivanja autentičnosti dokumenta, s konačnim ciljem utvrđivanja prava ili istinitosti činjenica koje sadrži (dokument op.a.)." (Duranti, 1989)[30]

Duranti pojašnjava pojam autentičnost na način da ga dijeli na tri vrste:

1) Diplomatička autentičnost. Dokumenti koji su stvoreni na način koji odgovora uobičajenim postupcima za mjesto i period koji su navedeni u dokumentu, te koji je potpisan imenom osobe ili osoba koje su imale mandat stvoriti takav dokument.[31] Naglasak na potpisu dokumenta, to jest na dokazu identiteta autora je opet prisutan. U svijetu digitalnih dokumenata pod pojmom potpis riječ je o

[28] Cencetti, G. (1985). La preparazione dell'archivista. Antologia di Scritti Archivistici, 283-313. Preuzeto 17. 12. 2021. s http://2.42.228.123/dgagaeta/dga/uploads/documents/Saggi/543ba04bdab11.pdf
[29] Duranti, L. (1989). Diplomatics, n. dj.
[30] Ibid.
[31] Ibid.

digitalnim potpisima, stoga se može zaključiti je dugotrajno očuvanje izvorna svrha diplomatike.

2) Pravna autentičnost. Pod ovim se podrazumijeva autentičnost koja je proizašla iz činjenice da je stvaranju dokumenta svjedočio predstavnik javne službe koji garantira za autentičnost dokumenta.[32] Suvremeni primjer ovakvih dokumenata u Republici Hrvatskoj su dokumenti koji su ovjereni kod javnog bilježnika. Možda se može tvrditi da ovaj uvjet ispunjavaju i dokumenti potpisani kvalificiranim elektroničkim potpisom kakav je propisan u eIDAS uredbi.[33] Oni svakako imaju snagu pravnog dokaza te njihovu upotrebu regulira javna ustanova.

3) Povijesna autentičnost. Povijesno autentični dokumenti su oni koji dokumentiraju događaje koji su se zaista dogodili ili informacije koje su neosporivo istinite.[34]

Ove tri vrste autentičnosti nisu međusobno ovisne. Na primjer, moguće je da je dokument diplomatički autentičan, jer sadrži ispravan potpis i pravilno je formiran, ali nije povijesno autentičan jer sadrži neistine. Isti dokument nije pravno autentičan ako nije stvoren pred ovlaštenim svjedokom. Istraživanje modela za dugotrajno očuvanje digitalno potpisanoga gradiva prije svega uzima u obzir diplomatičku autentičnost, donekle pravnu, i potpuno zanemaruje povijesnu. Očuvanje diplomatičke autentičnosti upotrebom suvremenih kriptografskih tehnika je osnovna svrha novog modela.

Uloga diplomatike, i posebno arhivistike, je šira od ovoga (dokazivanja autentičnosti kroz dokazivanje nepromjenjivosti dokumenta i autorstvo), štoviše, Duranti dio diplomatike koja se ovime bavi svrstava u poddisciplinu koju naziva "posebna diplomatika" (a koristi sličan termin i u kontekstu arhivistike – "posebna arhivistika") ali se ova rasprava fokusira na ovo područje (autentičnost) jer je ono centralno za provedeno istraživanje. Danas se autentičnost digitalnog dokumenta dokazuje kroz digitalni potpis, i njegovo dugotrajno očuvanje odgovara i ovoj, izvornoj svrsi diplomatike.

[32] Duranti, L. (1989). Diplomatics, n. dj.
[33] Uredba je raspravljena u kasnijim poglavljima.
[34] Duranti, L. (1989). Diplomatics, n. dj.

"Diplomatika je disciplina koja proučava dokument ili, ako želimo, osnovnu arhivsku jedinicu (dokument ili registar) i analizira njene formalne aspekte da bi definirala njenu pravnu prirodu, s obzirom na njeno stvaranje i učinak." (Carucci, 1987)[35]

Diplomatika je preteča arhivistike, ali njihov odnos nije striktno odnos prethodnika i nasljednika. Prema Duranti, i u skladu s ranijim definicijama, diplomatika se koncentrira na proučavanje individualnih dokumenata te se ova disciplina razvijala prije utemeljenja arhivskih ustanova. S pojavom arhivskih ustanova i arhiva razvija se arhivistika koja u svojem djelovanju obuhvaća cjelokupni rad s arhivskim gradivom. Osim ispitivanja autentičnosti, koje je bilo centralna briga (posebne) diplomatike, arhivistika uključuje razne organizacijske i druge elemente koji dolaze do izražaja pri radu s velikim brojem dokumenta, to jest arhivskim gradivom. Riječima Duranti:

"Granična linija između ove dvije discipline se očituje se u serijama, fondovima, arhivima kao složenim strukturama dokumenta, koje u svojoj cijelosti tvore područje arhivistike. Umjesto toga, individualni dokument, osnovna arhivska jedinica, je područje diplomatike." (Duranti, 1989)[36]

Ova razlika je važna, jer ukazuje na to da je arhivistika preuzela metodologiju ispitivanja i dokazivanja autentičnosti iz diplomatike. Ako je tako, onda su prethodne definicije, koje se izravno odnose na diplomatiku, primjenjive i u slučaju arhivistike. Ove dvije discipline, ili znanosti, ne čine linearan slijed razvoja znanosti o upravljanju zapisima već, u svojoj osnovi, djeluju paralelno. Može se reći da je danas diplomatika dio arhivistike (koji je usredotočen na individualne dokumente).

U kasnijem radu Duranti dodatno razrađuje koncept autentičnosti te uvodi i koncept pouzdanosti arhivskih zapisa.[37] Duranti pouzdanost zapisa objašnjava na način:

"Zapis je pouzdan kada se prema njemu može odnositi kao prema činjeničnom stanju onoga što dokazuje." (Duranti, 1995)[38]

Odnos prema zapisu kao prema činjenici moguć je, to jest zapis je pouzdan, kada ispunjava zahtjeve forme i procedure. Forma je zadovoljena kada je zapis formiran na poseban

[35] Carucci, P. (1987). *Il documento contemporaneo. Diplomatica e criteri di edizione.* Rome, Italy: La Nuova Italia scientifica.
[36] Duranti, L. (1989). Diplomatics, n. dj.
[37] Duranti, L. (1995). Reliability and Authenticity, n. dj.
[38] Ibid.

način, u skladu s društveno-pravnim zahtjevima vremena i podneblja u kojem je stvoren. Duranti navodi da se ovo obično postiže time da zapis sadrži oznaku datuma kada je stvoren i potpis osobe koja preuzima odgovornost za sadržaj zapisa. Poštovanje procedure podrazumijeva da je zapis stvoren u skladu s pravilima koja se odnose na njegov sadržaj. Na primjer, da potpisnik zaista smije preuzeti odgovornost za sadržaj zapisa ili da se zapisom, tijekom njegovog stvaranja, rukuje na točno propisan način.[39] Iz pouzdanosti proizlazi osobina dokumenta da se kasnije koristi kao dokaz.

S druge strane, prema ovoj interpretaciji, zapis je autentičan kada se može dokazati da je ostao nepromijenjen od trenutka njegovog nastajanja do trenutka kada se pokušava koristiti kao dokaz neke radnje. Autentičnost se dokazuje ukazivanjem na posebne sigurnosne mjere i procedure u arhivu. Ako su ovakve procedure dostatne i može se dokazati da su poštovane, zapis je autentičan, to jest tijekom vremena je sačuvao razinu pouzdanosti koju je imao pri nastanku.[40]

U usporedbi s prethodnim radom, Duranti je u ovim definicijama premjestila činjenicu da je dokument (ispravno) potpisan iz pojma autentičnosti u pojam pouzdanosti zapisa, te se sada autentičnost odnosi samo na činjenicu da zapis nije promijenjen od trenutka nastajanja. Drugim riječima, autentičnost više nije povezana s dokazivošću identiteta autora, osim u smislu da dokazuje da je potpis u izvornom stanju. Po takvoj definiciji moguće je imati autentičan falsificirani zapis, to jest zapis može biti autentičan i nepouzdan.

Usprkos ovoj izmjeni u odnosu na definicije dane pri opisu diplomatike, Duranti u radu navodi da je pravovaljanost (izvornost, originalnost, engl. *genuineness*) zapisa, to jest njegova mogućnost da i nakon prolaza nekog vremena dokaže određenu radnju iz prošlosti, uvjetovana i pouzdanošću i autentičnošću zapisa.[41] U ovom kontekstu ranija definicija nije izmijenjena već je detaljnije razrađena podjelom zahtjeva jednog pojma (autentičnosti) na dva (autentičnost i pouzdanost). U naravi dokazivanje pravovaljanosti (engl. *genuineness*) postiže se na isti način i važnije, uloga digitalnog potpisa ostala je ista. Kao što će biti pokazano u idućem poglavlju, digitalni potpis sadrži sve elemente koje zahtjeva pouzdanost (datum i potpis osobe) i autentičnost (digitalni potpis sadrži kriptografske mehanizme koji garantiraju da je zapis ostao nepromijenjen).

[39] Duranti, L. (1995). Reliability and Authenticity, n. dj.
[40] Ibid.
[41] Ibid.

2.2. Arhivska veza

Arhivska veza koncept je koji datira iz vremena diplomatike. Prema Duranti, diplomatika, koja je uspostavljena kao alat koji pomaže u dokazivanju vlasništva (i crkvenih dogmi), najčešće nad nekretninama, prvi je primjer potrebe za uspostavljanjem odnosa između više zapisa, odnosa koji danas zovemo arhivska veza.[42] Kao što je ranije pokazano, aktivnosti koje spadaju pod diplomatiku uključuju proučavanje i dokazivanje autentičnosti dokumenata, koji najčešće nisu postojali sami već su bili dio cijelog niza dokumenta (na primjer, slijed prijenosa vlasništva). Pojam arhivska veza podrazumijeva odnos više dokumenta (ili druge vrste arhivskog gradiva) koji su nastali u postupku dokumentiranja nekog procesa. Na primjer, u Republici Hrvatskoj prodaja nekretnine će možda započeti potpisivanjem predugovora, nakon toga možda će uslijediti zahtjev za izradom energetskog certifikata koji je preduvjet za zahtjev za odobravanje kredita. Nakon potpisivanja ugovora o kreditu i njegove realizacije, potpisat će se i kupoprodajni ugovor kojim se prodaje predmetna nekretnina. Posljednji koraci ove aktivnosti uključivati će zahtjev za izmjenom vlasništva u katarskim knjigama i zahtjeve za izmjenom korisnika komunalnih usluga. Svi spomenuti dokumenti zajedno opisuju i dokumentiraju postupak promjene vlasništva nad nekretninom. Zapis kojim je dokumentiran zahtjev za izradu energetskog certifikata objekta sam po sebi nema posebno značenje, ali u kontekstu opisanog procesa on je ključni element bez kojeg naknadni koraci nisu mogući. Općenito može se reći da prema načelima diplomatike i, kasnije, arhivistike dokument postaje arhivski zapis tek kada je smješten u odnos koji se formira od svih dokumenta (ili drukčijih zapisa) koji sudjeluju u aktivnosti na koju se odnose. Postojanje ovog odnosa dio je dokazivosti autentičnosti zapisa.

Iako je koncept značajno stariji, prva definicija termina "arhivska veza" (engl. *archival bond*) pojavljuje se 1997. kada je Duranti u svom radu "Arhivska veza" ponudila dvije definicije termina. U prvoj, arhivska veza je definirana kao:

"...mreža odnosa svakog zapisa s drugim zapisima iz istog skupa." (Duranti, 1997)[43]

Gornja definicija odgovara objašnjenju koje je dano u uvodu ovog poglavlja. Kasnije u radu Duranti navodi detaljniju definiciju termina:

[42] Duranti, L. (1997). The archival bond, n. dj.
[43] Ibid.

"...izražaj razvoja aktivnosti u kojoj dokument sudjeluje, umjesto čina koji je utjelovljen dokumentom, jer (op.a. arhivska veza) u sebi sadrži smjer odnosa uzroka i posljedica." (Duranti, 1997)[44]

Osim definicija, Duranti dodatno objašnjava termin navodeći najvažnije značajke arhivske veze. Prema Duranti arhivska veza je:

1) izvorna (engl. *originary*), arhivska veza nastaje u trenutku nastanka zapisa;
2) nužna (engl. *necessary*), arhivska veza nužno postoji za svaki arhivski zapis, u suprotnom govorimo isključivo o zapisu;
3) (pred)određena (engl. *determined*), arhivska veza je kvalificirana funkcijom zapisa u skupu kojem pripada.[45]

Iz navedenih definicija jasno je da je arhivska veza jedan od najvažnijih metapodataka u kontekstu arhivske građe. Ona definira značenje (i značaj) zapisa. Arhivska veza se, iako definira skup zapisa kojem neki zapis pripada, razlikuje od klasifikacija na koje smo naviknuti izvan arhivistike. Uobičajeni načini klasifikacije razvrstavaju objekte prema vrsti njihovog sadržaja. Na primjer, takva kategorija može biti "zahtjevi za izradu energetskog certifikata" ili prema kontekstu u kojem su nastali, na primjer, "zahtjevi upućeni Gradskom uredu za prostorno uređenje Grada Zagreba 2021.". Ovakve klasifikacije mogu obuhvatiti zahtjev iz primjera s početka poglavlja, zbog konteksta u kojem je nastao ili njegovog sadržaja, ali ne odgovaraju arhivskoj vezi koja opisuje ranije navedeni postupak prodaje nekretnine. Bez očuvanja arhivske veze gubi se dio informacija vezanih uz navedeni zahtjev, konkretno, gubi se razlog njegova nastajanja. S obzirom da arhivistika često ima cilj pružiti (pravni) dokaz o nekom postupku, gubitak ovakvih podataka nije prihvatljiv i može se s pravom reći da je zapis koji je izgubio arhivsku vezu izgubio i autentičnost.

Poseban problem pri održavanju podataka o arhivskoj vezi proizlazi iz činjenice da svi zapisi, to jest sva građa ne mora biti pohranjena na istom mediju. Katie Rudolph je istražila problematiku održavanja arhivske veze u slučaju kada se dio građe koja sudjeluje u arhivskoj vezi sastoji od trodimenzionalnih objekata, artefakata.[46] Održavanje arhivske veze u

[44] Duranti, L. (1997). The archival bond, n. dj.
[45] Ibid.
[46] Rudolph, K. (2011). Separated at appraisal: Maintaining the archival bond between archives collections and museum objects. *Archival Issues, 33*(1), 25-40. Preuzeto 16. 12 .2021. s https://minds.wisconsin.edu/bitstream/handle/1793/72333/AI_Vol33_No1_KatieRudolph1.pdf?sequence=1

tradicionalnom arhivu postizalo se na način da se građi koja sudjeluje u vezi dodjeljivala posebna klasifikacijska šifra, koja je jednoznačno označavala arhivsku vezu, i tako da se građa koja je dio veze drži fizički blizu, ponekad u istom pretincu. Rudolph navodi da do problema dolazi jer se trodimenzionalni predmeti, koji možda sudjeluju u arhivskim vezama, često pohranjuju na mjestima udaljenim od papirnatih arhivskih zapisa. Oni se u nekim slučajevima prepuštaju na čuvanje drugim ustanovama, često muzejima. U slučaju digitalnih zapisa ne postoji koncept fizičke blizine zapisa, jedini način da se arhivska veza definira je spomenuto uvođenje šifre ili neke druge posebne oznake. Do problema dolazi kada je dio građe u digitalnom obliku, a dio u papirnatom, ili kako Rudolph navodi u obliku trodimenzionalnog artefakta. Rudolph predlaže upotrebu informacijskih sustava koji bi održavali podatke o arhivskoj vezi kao moguće rješenje, to jest sustava poput onih baziranih na TrustChain modelu. Ovo znači da će dio građe biti dio sustava, a dio neće (onaj koji nije digitalni). Štoviše u slučaju sustava temeljenog na TrustChain modelu problem će možda biti još više izražen. TrustChain je predviđen za očuvanje autentičnosti digitalno potpisanoga gradiva, situacija u kojoj je dio zapisa iz iste arhivske veze digitalno potpisan, a dio nije (ili uopće nije u digitalnom obliku) je vrlo vjerojatna. Ovo nas vraća na problem s kojim je Rudolph započela (gradivo koje je disperzirano kroz više ustanova ili, u ovom slučaju, informacijskih sustava).

Ovaj problem uočila je i Duranti u ranije spomenutom radu[47] ali ni ona ne daje konkretno rješenje problema već navodi da područje nije dovoljno istraženo te da mu treba posvetiti posebnu pažnju. Prema saznanjima autora TrustChain modela i ove disertacije, problematika održavanja arhivske veze gradiva koje je pohranjeno u različitim digitalnim arhivskim sustavima, pa čak i različitim arhivskim (ili drugim) ustanovama i dalje nije dovoljno istraženo. TrustChain model će uvažiti ovaj problem i ponuditi rješenje za njega koje, za sada, ostaje teorijsko rješenje nedovoljno istraženog problema.

Problem arhivske veze nedavno su razmotrili Lemieux i Sporny[48] u kontekstu suvremenih digitalnih sustava koji stvaraju veliki broj zapisa, poput Bitcoin kriptovalute. Ovi autori predložili su razvoj posebnih ontologija koje su dizajnirane za održavanje podataka o arhivskim vezama digitalnih zapisa. Osim toga autori su predložili i uvođenje arhivske veze u Bitcoin sustav upotrebom OP_RETURN polja u koje bi se mogli pohraniti podaci o Bitcoin transakcijama upotrebom posebnih ontologija.

[47] Duranti, L. (1997). The archival bond, n. dj.
[48] Lemieux, V. L., & Sporny, M. (2017). Preserving the Archival, n. dj.

Važniji za ovo istraživanje od prijedloga ontologije i metoda očuvanja arhivske veze u Bitcoin sustavu su uvidi u osobine arhivske veze koje Lemieux i Sporny navode. U svom radu oni naglašavaju dvije osobine koje nisu naglašene u ranijem radu Luciane Duranti. Ove osobine arhivske veze su:

1) Višestrukost – autori navode da, iako se o arhivskoj vezi obično govori u jednini, veliki broj zapisa sudjeluje u više od jedne arhivske veze. U prethodnom primjeru ugovor o kreditu realiziranom za kupovinu nekretnine naveden je kao dio arhivske veze koja opisuje postupak promjene vlasnika nekretnine. Isti ugovor mogao bi biti i dio arhivske veze koja opisuje kasniji spor između banke i korisnika kredita. Svaki arhivski zapis sudjeluje u barem (a ne samo) jednoj arhivskoj vezi.

2) Promjenjivost – arhivska veza nije statična. Lemieux i Sporny naglašavaju da se arhivske veze mijenjaju, to jest rastu kako se stvaraju novi zapisi. Ovo može biti izuzetno dug proces, na primjer u slučaju višegodišnjih sudskih sporova.

Važnost arhivske veze naglasila je i Dara Hofman u svom radu iz 2017.[49]. Hoffman istražuje mogućnosti upotrebe Ethereum pametnih ugovora u arhivske i svrhe dokaza na sudovima. Hoffman naglašava da, ako pametni ugovori imaju ambiciju biti korišteni kao pravno obvezujući, oni moraju podržavati održavanje podataka o arhivskoj vezi jer u suprotnom gube na autentičnosti. Hofman predlaže razvoj posebnog, pravnog, semantičkog sloja koji bi riješio pravne probleme vezane uz upotrebu pametnih ugovora (uključujući održavanje arhivske veze).

Semantički sloj koji autorica predlaže može biti primjenjiv i na TrustChain model. Ipak, za novi model je u ovoj fazi važnije naglasiti jasnu potrebu za uključivanjem ovih podataka. Jedna od hipoteza istraživanja vezana je za arhivsku vezu ali svejedno, već u ovoj ranoj fazi istraživanja, postaje očigledno da će uključivanje arhivske veze u sustav temeljen na nepromjenjivim podatkovnim strukturama zahtijevati znatno više od dodavanja polja koje omogućava klasifikaciju zapisa.

Bit će potrebno razviti posebne postupke koji omogućavaju naknadnu izmjenu podataka (o arhivskoj vezi) i trebat će omogućiti da zapis sudjeluje u više arhivskih veza. S druge strane, s obzirom na to da je pokazano da je arhivska veza dio dokaza autentičnosti, podatke vezane uz

[49] Hofman, D. L. (2017). Legally speaking: Smart contracts, archival bonds, and linked data in the blockchain. *26th International Conference on Computer Communication and Networks (ICCCN)* (str. 1-4). IEEE. Preuzeto 16. 12. 2021. s https://blockhack.osive.com/_downloads/d5a5c8e62a06c24e5791ab043a950796/61.pdf

nju je potrebno osigurati u najvećoj mogućoj mjeri. Svi navedeni autori (Duranti, Lemieux i Sporny, Hoffman) posebno naglašavaju da gubitak arhivske veze rezultira kompromitacijom autentičnosti arhivskog gradiva. Konkretno, Duranti navodi da se:

> *"...arhivski opis se, kao način objašnjavanja arhivske veze u dokumentarnom kontekstu, tradicionalno smatra primarnim načinom širenja i autentificiranja značenja zapisa i, u vremenima migracija (op.a. digitalnih migracija), vjerojatno najbolja metoda dugotrajnog održavanja autentičnosti."* (Duranti, 1997)[50]

Lemieux i Sporny naglašavaju:

> *"...ovaj međusobni odnos određuje jedinstveno značenje i identitet zapisa, i zbog toga što zapis mora imati jedinstven identitet prije nego što njegova autentičnost može biti potvrđena, on pruža temelj za uspostavljanje autentičnosti (op.a. zapisa)."* (Lemieux & Sporny, 2017)[51]

Hoffman piše da, u slučaju pametnih ugovora:

> *"Bez arhivske veze nemoguće je znati je li ugovor prihvaćen, jer je nemoguće rekonstruirati odnose zapisa na način koji dokazuje da se 'prihvaćanje' zaista dogodilo, za razliku od pokušaja prihvaćanja istekle ili povučene ponude."* (Hofman, 2017)[52]

Ovo arhivsku vezu postavlja na istu razinu važnosti, u kontekstu dokazivanja autentičnosti, kao i digitalni potpis (kada je riječ o digitalno potpisanim zapisima). Idealno, podaci o arhivskoj vezi su dio nepromjenjive podatkovne strukture.

[50] Duranti, L. (1997). The archival bond, n. dj.
[51] Lemieux, V. L., & Sporny, M. (2017). Preserving the Archival Bond, n. dj.
[52] Hofman, D. L. (2017). Legally speaking, n. dj.

2.3. Primjena načela diplomatike i arhivistike na digitalne zapis

Ovo poglavlje razmatra načine na koje se prethodno definirani koncepti, prije svega dokazivanje autentičnosti, koji su se ranije provodili pregledima žigova i potpisa papirnate arhivske građe, primjenjuju na digitalnu građu, prije svega digitalno potpisanu građu.

Duranti je sa suradnicima u sklopu InterPARES istraživanja ponudila daljnju razradu i suvremenije definicije temeljenih koncepata arhivistike[53] o kojima se govorilo u prethodnom poglavlju, te ih je primijenila na digitalne zapise. Rezultati ovog istraživanja dodatno su razrađeni i pojašnjeni u doktorskom radu Hrvoja Stančića u kojem je razvijen teorijski model sustava za očuvanje autentičnosti digitalnih zapisa.[54] Iako se razvijeni modeli značajno razlikuju, čak i na konceptualnoj razini, zahtjevi za (digitalne) arhivske sustave su isti te je ova disertacija logična početna točka pri istraživanju kako se ranije objašnjeni principi primjenjuju na digitalne zapise. Na sličan način primjenjiv je doktorski rad Hrvoja Brzice [55], koji je izravnije povezan s ovim modelom jer se oba oslanjaju na kriptografske tehnike. Ipak, Brzičin model ne uključuje izravnu upotrebu povezanih zapisa, iako navodi mogućnost njihove upotrebe, već se u svojem rješenju ograničava na upotrebu PKI sustava i digitalnih vremenskih žigova za dugotrajno očuvanje autentičnosti.

Prije rasprave o dugoročnom očuvanju autentičnosti potrebno je utvrditi na koji način se ona uopće postiže u kontekstu digitalnih zapisa. Godine 2002. pripremljen je "Izvještaj o autentičnosti elektroničkih zapisa" za UNESCO koji navodi:

> *"Dokazivanje autentičnosti je postupak dokazivanja koji određuje je li objekt koji se vodi kao dokaz u pravnom postupku autentičan. U slučaju dokaza, autentičnost znači da objekt zaista jest ono što strana koja ga uvodi u postupak tvrdi da jest. Najlakši način dokazivanja autentičnosti je postojanje svjedoka koji ju (autentičnost, op. a.) može potvrditi."* (Gränström, i dr., 2002)[56]

---

[53] Duranti, L., Gilliland-Swetland, A., Guercio, M., Hamidzadeh, B., Iacovino, L., Lee, B., . . . Ross, S. (2001). *Authenticity Task Force Final Report.* InterPARES. Preuzeto 10. 12. 2021. s http://www.interpares.org/book/interpares_book_d_part1.pdf

[54] Stančić, H. (2006). Teorijski model postojanog očuvanja autentičnosti, n. dj.

[55] Brzica, H. (2018). *Koncept uspostave elektroničkoga arhiva u javnoj upravi.* Zagreb: FFZG. Preuzeto 14. 12. 2021. s http://darhiv.ffzg.unizg.hr/id/eprint/10282/1/Doktorski_rad_-_Koncept_uspostave_elektroni%C4%8Dkoga_arhiva_u_javnoj_upravi.pdf

[56] Gränström, C., Hornfeldt, T., Peterson, G., Rinaldi Mariana, M. P., Schäfer, U., & Zwicker, J. (2002). *Authenticity of Electronic Records, a report by ICA to UNESCO*. INTERNATIONAL COUNCIL ON ARCHIVES. Preuzeto 11.12.2021. s https://www.ica.org/sites/default/files/ICA_Study-13-1-Authenticity-of-electronic-records-ICA-Report-to-UNESCO_EN.pdf

Autori izvještaja dalje tvrde da je ova metodologija razvijena za papirnate dokumente i ne uzima u obzir digitalne zapise. I tvrdnja je sigurno na mjestu. Računarstvo je kao dokaz autentičnosti elektroničkih zapisa ponudilo digitalni potpis za koji Stančić u svojoj doktorskoj disertaciji navodi:

> *"Općenito, elektroničkim se potpisom može utvrditi identitet osobe koja ga je pridodala informacijskom objektu, ali i osigurati integritet podataka."* (Stančić H., 2006)[57]

Ako digitalni potpis utvrđuje identitet osobe i integritet podataka, to jest nepromjenjivost podataka onda on u potpunosti ispunjava uvjete za dokazivanje autentičnosti prema Duranti koji su objašnjeni u prethodnom poglavlju. Nadalje, Stančić pojašnjava način na koji se ovo postiže:

> *"Elektronički potpis osigurava integritet elektroničkog objekta, tj. njegovu nepromijenjenost, služeći se tehnologijom šifriranja (zakrivanja) javnim ključem (engl. public-key encryption) i funkcijom raspršenja (op.a. kriptografskog sažetka) (engl. hash function)."* (Stančić H., 2006)[58]

Funkcioniranje obje komponente digitalnog potpisa (funkcije za šifriranje i kriptografskog sažetka) detaljno je objašnjeno u idućem poglavlju te se na osnovu njega može neosporno tvrditi da digitalni potpis zaista ispunjava zahtjeve za dokazivanjem autentičnosti, možda i s većom razinom pouzdanosti nego što je slučaj s tradicionalnim potpisom i žigom. Problem s upotrebom ove tehnologije je vremenska ograničenost digitalnog potpisa, to jest povezanog digitalnog certifikata. Kao što će biti pokazano kasnije u ovom poglavlju, digitalni certifikati ističu i digitalni potpis (za razliku od tradicionalnog) je nužno vremenski ograničen. Postoji više mogućih rješenja ovog problema, jedan od kojih je novi model koji je tema ovog istraživanja.

Stančić u svojem doktorskom istraživanju utvrđuje još dva elementa sustava za dugoročno očuvanje digitalnog arhivskog gradiva koji su važni za razvoj novog modela te su izravno u njega implementirani. Prema Stančiću:

> *"Provjera autentičnosti bi, kad je riječ o uključivanju zapisa u sustav za očuvanje na dulji vremenski rok, trebala biti omogućena najmanje u dvije*

[57] Stančić, H. (2006). Teorijski model postojanog očuvanja autentičnosti, n. dj.
[58] Ibid.

situacije. Jednom, prilikom prihvata kako bi se provjerio identitet i integritet dostavljenih zapisa, te drugi puta na zahtjev korisnika prilikom pristupa očuvanim zapisima." (Stančić H., 2006)[59]

Pod "provjera autentičnosti", ako je dokaz autentičnosti ostvaren digitalnim potpisom, podrazumijeva se postupak provjere ispravnosti digitalnog potpisa. Funkcioniranje digitalnog potpisa detaljno je opisano kasnije ali u ovoj fazi rasprave može se ukratko navesti da se ovaj postupak (provjere ispravnosti digitalnog potpisa) može podijeliti na dva podpostupka:

1) Provjeru identiteta autora. Ova provjera podrazumijeva čitanje podataka o vlasniku digitalnog certifikata i provjeru njegove ispravnosti. Ispravnost certifikata se utvrđuje prateći niz potpisa (i provjeravajući njihovu ispravnost) od potpisnika dokumenta do certifikacijskog autoriteta (ustanove od povjerenja koja je izdala digitalni certifikat).

2) Provjeru integriteta podataka. Ovu provjeru moguće je izvršiti bez oslanjanja na vanjske ustanove (certifikacijski autoritet). Ona se svodi na matematičke izračune koji su dio algoritama za dešifriranje i izračun *hash* vrijednosti (oba su objašnjenja u kasnijem poglavlju).

Do problema dolazi kod točke broj 1. Upotrebom navedenih kriptografskih tehnika uvijek se može dokazati integritet podataka, ali identitet autora je dokaziv samo ako je digitalni certifikat (čije je trajanje u pravilu dvije godine) još uvijek valjan. Ako ove postupke, kako Stančić navodi, treba obaviti prilikom prihvata dokumenta te kasnije na zahtjev korisnika, digitalni arhiv se suočava sa značajnim problemom. Provjera ispravnosti potpisa je možda i bila moguća prilikom ulaska zapisa u arhiv (ako nije, to nije krivica arhiva, dostavljen mu je neautentični zapisa te ga on može sačuvati samo kao takvog) ali u nekom kasnijem trenutku postoji značajna šansa da certifikat više nije valjan. Razlog ovome i ne mora biti istjecanje certifikata, certifikat je možda povučen zbog krađe podataka ili se algoritam koji koristi pokazao nepouzdanim. Ako je certifikat istekao dok je zapis u nadležnosti arhiva onda se može reći da arhiv nije ispunio jednu od svojih zadaća – očuvanje autentičnosti (digitalno potpisanoga) gradiva.

[59] Stančić, H. (2006). Teorijski model postojanog očuvanja autentičnosti, n. dj.

Novi model, TrustChain, mora omogućiti provjeru autentičnosti gradiva u oba slučaja. Zbog tog model mora sadržavati dva različita postupka za provjeru autentičnosti:

1) Inicijalnu provjeru autentičnosti koja pri prihvatu dokumenta provjerava autentičnost na uobičajen način (provjerom ispravnosti potpisa).

2) Naknadnu provjeru autentičnosti koje se može provesti u bilo kojem trenutku neovisno o stanju digitalnog certifikata. Postojanje ovakve provjere podrazumijeva da model uključuje posebne metode kojima će ovo osigurati bez oslanjanja na certifikacijski autoritet. TurstChain model ovo ostvaruje upotrebom nepromjenjivih ulančanih zapisa.

Brzica je u svojoj disertaciji[60] raspravu o autentičnosti većim dijelom utemeljio na spoznajama do kojih su došli Jean-François Blanchette u svom radu iz 2006. godine[61] i InterPARES Trust projekt.[62] Na temelju ovih istraživanja Brzica utvrđuje da su se pojavila četiri moguća rješenja za problem dugoročnog očuvanja digitalno potpisanoga gradiva:

"1. Očuvanje elektroničkih potpisa,

2. Uklanjanje elektroničkih potpisa,

3. Bilježenje traga o elektroničkim potpisima u metapodacima,

4. Bilježenje valjanosti o elektroničkim potpisima u blokchainu." (Brzica H., 2018)[63]

Točka 2, metoda koju predlaže Blanchette podrazumijeva uklanjanje digitalnog potpisa. Smatram da je ovo gruba povreda autentičnosti digitalnog zapisa i kao takva nije rješenje problema već odustajanje od njega.

Točka broj 1 je Brzičino preferirano rješenje problema. Brzica navodi da sam elektronički potpis napreduje i reguliran je zakonima i propisima[64] te kasnije u radu predlaže produživanje dugotrajnosti potpisa upotrebom vremenskih žigova. Ovo je sasvim sigurno moguće rješenje, te je i ono koje je trenutno predloženo od strane regulatornih tijela. Vremenski žigovi su tehnika kojom se osigurava integritet zapisa od neke točke u vremenu i regulirana je

[60] Brzica, H. (2018). Koncept uspostave elektroničkoga arhiva, n. dj.
[61] Blanchette, J.-F. (2016). The digital signature dilemma. *Annales des télécommunications, 61*(7), 908-923. Preuzeto 18. 12. 2021. s https://escholarship.org/content/qt1kt3f8hx/qt1kt3f8hx.pdf
[62] InterPARES Trust, https://interparestrust.org/
[63] Blanchette, J.-F. (2016). The digital signature dilemma, n. dj.
[64] European Parliament. (2014). *eIDAS*, n. dj.

Uredbom eIDAS.[65] Način funkcioniranja vremenskih žigova detaljnije je raspravljen u poglavlju koje se bavi prethodnim rješenjima (jer se ona prvenstveno temelje na vremenskim žigovima). Nažalost, pri upotrebi vremenskih žigova, kao što će kasnije biti pokazano, se (samo) produžuje period dokazivosti integriteta zapisa, i dalje je vjerojatno, a ovisno o konkretno korištenoj tehnologiji, u nekim slučajevima i sigurno, da će vremenski žig trebati zamijeniti. Situacija je slična problemu s digitalnim certifikatima, ali je vrijeme isteka značajno duže (ovisno o tehnologiji 10 ili više godina).

U zaključku pregleda normi i uredbi koje se odnose na vremenske žigove iz 2016. Volarević i Stančić[66] utvrđuju da:

> *"Oni (vremenski žigovi op. a.) zahtijevaju periodičnu provjeru i pravovremeno obnavljanje (odnosno generiranje novih koji obuhvaćaju sve prethodne)"* (Volarević & Stančić, 2016)[67]

Osim ovog problema, upotreba vremenskog žiga sama po sebi garantira samo integritet podataka. Identitet autora treba provjeriti nezavisno od upotrebe vremenskih žigova, a onda i negdje zabilježiti te na neki način garantirati za integritet tih podataka.

Na tragu ovoga su 3. i 4. točka. Točka 3 temelji se na Blanchettovom prijedlogu prema kojem se podaci iz digitalnog potpisa prenose u neki drugi sustav, to jest u metapodatke zapisa i tamo se njihov integritet garantira na drukčiji način. Do sličnih zaključaka došli su Cullen i koautori u knjizi "Autentičnost u digitalnom okruženju".[68] Ovi autori predlažu digitalno potpisivanje "tvrdnji o metapodacima" (engl. *metadata claims*) kao jedno od rješenja problema dokazivosti autentičnosti digitalnog gradiva.

Točka 4 potječe iz rada autora ove disertacije i suradnika u sklopu InterPARES Trust istraživanja[69] koji je temelj i ovog istraživanja. Moglo bi se reći da je točka 4, "upotreba blockchaina" oblik konkretne implementacije Blanchettovog prijedloga (pod brojem 3).

---

[65] European Parliament. (2014). *eIDAS*, n. dj.
[66] Volarević, I., & Stančić, H. (2016). Norme za elektroničke vremenske žigove, n. dj.
[67] Brzica, H. (2018). Koncept uspostave elektroničkoga arhiva, n. dj.
[68] Cullen, C., Hirtle, P., Levy, D., & Lynch, C. R. (2000). *Authenticity in a Digital Environment.* Council on Library and Information Resources. Preuzeto 11. 12. 2021. s https://www.clir.org/wp-content/uploads/sites/6/pub92.pdf
[69] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation of digital signature validity: TrustChain. *INFuture2017 conference proceedings*, (str. 89-113). Zagreb. doi: https://doi.org/10.17234/INFUTURE.2017.10

Upotrebom povezanih zapisa (engl. *blockchain*) garantira se (nakon provjere da su ispravni pri prihvatu) integritet podataka koji omogućava dokazivanje autentičnosti digitalnih zapisa.

Da bi završili ovu raspravu treba prokomentirati i odnos autentičnosti prema digitalnim zapisima koji nisu potpisani. Od 90-ih godina prošlog stoljeća pa sve do danas arhivisti su (iz očiglednih razloga) izuzetno fokusirani na prokomentirani odnos autentičnosti i digitalnih zapisa ali, u većini slučajeva tijekom ovog perioda, fokus nije bio na problematici isteka digitalnih certifikata već na drukčijem problemu. Krajem prošlog i početkom ovog stoljeća veliki broj digitalnih zapisa morao je proći kroz proces migracije iz jednog formata u drugi. Ovo se događalo zbog zastarijevanja softverskih i hardverskih platformi koje su bile potrebne za reprodukciju zapisa. Čak i bez migracija bilo je upitno ostaje li zapis u svojem izvornom obliku prilikom bilo koje transmisije ili čak upotrebe u elektroničkim sustavima. Za razumijevanje ove problematike treba biti svjestan značajne razlike između papirnatih i drugih zapisa koji koriste fizičke medije. Papirnati zapis, jednom kada je završen proces njegovog stvaranja, postaje većim dijelom statičan (ako zanemarimo oštećenja materijala zbog starosti ili nezgoda) i trajno povezan za medij na kojem stvoren. Proučavanjem značajki izvornog medija, načina zapisa, potpisa i žigova, koji su svi statične strukture, diplomatika (i paleografija) su izvorno dokazivale autentičnost. Ovo uopće nije primjenjivo na digitalne zapise. Zbog načina na koji funkcioniraju sva suvremena računala, digitalni podatak nikada nije trajno povezan s medijem na koji je pohranjen. Digitalni podaci su dinamički, sustavi upravljanja radnom memorijom računala cijelo vrijeme premještaju podatke u njoj. Ista stvar se događa i na datotečnim sustavima. Rijetki primjeri stvarno statičnog digitalnog zapisa su mediji na koje se može pisati samo jednom, poput optičkog CD medija (engl. *compact disc*) ali čak i u slučaju njihove upotrebe prije samog zapisa, i tijekom svake reprodukcije zapis je morao biti prisutan u radnoj memoriji, dakle napravljena je kopija, kopija koja tijekom rada nije statična, ona se seli kroz medij (radnu memoriju). Nadalje, upotreba optičkih medija poput CD-a i magnetskih medija poput magnetske trake, koji su se koristili na način da bi se na njih pohranio zapis (češće zapisi) te bi se oni čuvali kao fizički predmeti u nekom arhivu, ukazao je i na problem isteka svojstva čuvanja podatka koji je prisutan kod ovih medija, što je izazvalo potrebu za migracijom podataka na novi medij (kojem još nije istekao rok trajanja) čak i kada nije došlo do potrebe da se to radi zbog razloga zastarijevanja tehnologije korištene u stvaranju zapisa.

Van Diessen i van der Warf-Davelaar u svom izvještaju IBM-u ovaj problem objašnjavaju na način:

> *"Problem digitalnih objekata leži u činjenici da su oni samo konceptualni objekti. Digitalni objekt je konceptualni objekt kojeg treba interpretirati (renderirati) upotrebom posebne IT infrastrukture (hardvera i softvera)."* (van Diessen & van der Werf-Davelaar, 2002)[70]

Zbog ove odvojenosti samog zapisa (digitalnog objekta) i medija na kojem je pohranjen, koja je stvorila značajne probleme za izvorne koncepte i metode dokazivanja autentičnosti, razvijena je velika količina modela koji pokušavaju riješiti problem. Od definicija koje dijele digitalne oblike na više razina (da bi razlučile promjenjiv dio od nepromjenjivog koji bi trebao očuvati autentičnost) do razvoja raznih administrativnih i tehničkih postupka koji bi trebali osigurati minimalne izmjene u digitalnom zapisu. Jedan od ranih primjera ovakvog istraživanja dala je i Duranti 1996. godine.[71] Slična istraživanja nastavljena su i u recentnijem periodu, kao što je vidljivo iz iscrpnog pregleda stručne literature Corinne Rogers.[72] I sam referentni model OAIS (Otvoreni Arhivski Informacijski Sustav)[73] pridaje ovome značajnu pažnju, na primjer kroz posebne postupke migracije podataka.

Važno je naglasiti da ova problematika, koja je i dalje aktualna, nema izravnog utjecaja na ovo istraživanje i razvoj novog modela. Ovo istraživanje proučava upotrebu kriptografskih tehnika koje garantiraju da su podaci ostali nepromijenjeni. Ako se prihvati činjenica da je digitalni objekt apstrakcija, to jest da za razliku od papirnog zapisa digitalni zapis ne postoji u obliku koji je statično zapisan na neki medij, onda se ove tehnike mogu prihvatiti kao dokaz autentičnosti. Ovo prihvaćanje se i dogodilo, već je pokazano prihvaćanje sa znanstveno-arhivističke strane, uskoro će biti pokazano da je ono prihvaćeno i u pravnom smislu, a u sljedećem poglavlju će biti objašnjeno kako se dokazivanje postiže na tehničkoj razini.

Ako sve navedeno stoji, onda problemi povezani uz migraciju digitalnih podataka i eventualni gubitak autentičnosti tijekom njihovog prijenosa nisu relevantni za ovo istraživanje.

---

[70] van Diessen, R. J., & van der Werf-Davelaar, T. (2002.). *Authenticity in a Digital Environment, IBM / Koninklijke Bibliotheek Long-Term Preservation Study Report Series.* Amsterdam: IBM Netherlands.
[71] Duranti, L., & Macneil, H. (1996). The Protection of the Integrity, n. dj.
[72] Rogers, C. (2016). A literature review of authenticity of records in digital systems: from 'machine-readable'to records in the cloud. *Acervo, 29*(2), 16-44. Preuzeto 16.12.2021. s https://www.researchgate.net/publication/320593846_A_Literature_Review_of_Authenticity_of_Records_in_Digital_Systems_From_%27Machine-Readable%27_to_Records_in_the_Cloud
[73] International Organization for Standardization. (2018). *ISO 14721:2012 Space data and information transfer systems — Open archival information system (OAIS) — Reference model.* International Organization for Standardization. Preuzeto 14. 12. 2021. s https://www.iso.org/standard/57284.html

Digitalno potpisani zapis nije potrebno na poseban način migrirati (moglo bi se reći da to nije ni moguće), digitalni potpis čini sadržaj digitalnog zapisa statičnim. Da bi se odradila migracija digitalnog zapisa moralo bi se prvo, kao što Blanchette predlaže,[74] ukloniti digitalni potpis. Ovakav dokument, koji više nije digitalno potpisan (ili nikad nije bio) predstavlja problem arhivističkoj struci ali nije predmet ovog istraživanja. Ovo implicira da će arhivski sustavi koji koriste ovdje predstavljeni model (ili na drugi način čuvaju digitalno potpisane zapise) morati trajno osigurati softversku i hardversku infrastrukturu potrebnu za njihovu reprodukciju. Ovo je zaseban problem koji, opet, nije predmet ovog istraživanja, ali može se kratko prokomentirati da je on, zbog širenja virtualizacijskih tehnologija i stabilizacije formata zapisa, danas značajno manje izražen (ali i dalje prisutan) nego što je bio na prijelazu iz prošlog u ovo stoljeće kada je provedena većina istraživanja ove problematike.

2.4. Zakonski i drugi propisi vezani uz digitalne arhive

U ovoj fazi istraživanja promotren je hrvatski Zakon o arhivskom gradivu i arhivima (NN 61/18, 98/19).[75] Prije izrade modela značajno je utvrditi postoje li neki specifični zahtjevi za ovakav sustav koji su propisani zakonom Republike Hrvatske. Model je predviđen za međunarodnu upotrebu ali bi usporedba nacionalnih zakona, čak i samo članica Europske unije, premašila opseg ovog istraživanja i prešla duboko u područje prava. Umjesto pregleda svih zakona koji bi mogli imati utjecaja kao reprezentativni primjer odabran je zakon matične države ustanove na kojoj je provedeno istraživanje. Osim specifičnih zahtjeva značajno je bilo proučiti i propisane rokove čuvanja arhivskog gradiva.

Hrvatski Zakon o arhivskom gradivu i arhivima ne postavlja nikakve posebne i konkretne zahtjeve informacijskim sustavima. U zakonu se dosta govori o procesu digitalizacije prethodno analognog gradiva i migracijama digitalnih podataka iz jednog sustava u drugi, ili kako to zakon zove "pretvorbi gradiva", ali i ta pravila su proceduralna i općenita te se uglavnom ograničavaju na upute poput članka 3. točke n:

> *"pretvorba gradiva je postupak prebacivanja gradiva iz jednog sustava u drugi, uz očuvanje autentičnosti, integriteta, pouzdanosti i iskoristivosti."* (Hrvatski sabor, 2018)[76]

[74] Blanchette, J.-F. (2016). The digital signature dilemma, n. dj.
[75] Hrvatski sabor. (2018). *Zakon o arhivskom gradivu i arhivima*, n. dj.
[76] Ibid.

Nažalost, gore navedeni pojmovi, poput autentičnosti i integriteta, koji su od izuzetne važnosti za ovo istraživanje u zakonu nisu detaljnije objašnjeni te se stoga nužno osloniti na prethodno utvrđene arhivske definicije.

Rokovi čuvanja arhivskog gradiva relevantni su za ovo istraživanje jer daju potvrdu (zakonskoj) potrebi za sustavom koji bi omogućio dugotrajnu pohranu digitalno potpisanoga gradiva. Hrvatski zakon je i u ovom pogledu vrlo općenit te propisivanje konkretnih rokova, u slučaju gradiva koje su oni stvorili, prepušta javnim tijelima u članku 10.[77] Ove rokove potvrđuje "nadležni državni arhiv". I bez konkretnih rokova indikativan je članak 10. stavak 1. istog zakona koji kaže:

> *"Javno arhivsko gradivo predaje se nadležnom državnom arhivu u roku koji u pravilu nije dulji od 30 godina od njegova nastanka."* (Hrvatski sabor, 2018)[78]

Ako je gradivo predano 30 godina nakon njegovog nastajanja sasvim je realno očekivati da su svi digitalni certifikati povezani uz takvo gradivo davno istekli. Njihovo održavanje prije ulaska u arhiv dužnost je javne ustanove koja je stvorila gradivo. Članak svejedno jasno ukazuje na potrebu za razvojem sustava koji će omogućiti dugotrajnu (to jest dugotrajniju) upotrebu digitalnih certifikata. Slični rokovi, koji ukazuju na potrebu za dugotrajnom pohranom mogu se naći u članku 18. stavku 3. koji navodi da klasificirani i drugi tajni podaci postaju dostupni tek nakon 40 godina od nastanka i u članku 19. stavku 1. koji navodi rok od 100 godina (da bi podaci postali dostupni) od rođenja osobe (ili odmah nakon smrti osobe) u slučaju osobnih podataka koji su dio arhivskog gradiva.

Na tragu ranije spomenutog članka 10. koji je odluku o konkretnim rokovima prepustio javim tijelima i nadležnom državnom arhivu, na web stranicama Nacionalnog Arhivskog Informacijskog Sustava[79] na koji autorska prava polažu Hrvatski Državni Arhiv[80] i Avicena Software d.o.o.[81] dostupan je "Opći popis gradiva s rokovima čuvanja" koji je 2012. izdalo Hrvatsko arhivsko vijeće (Hrvatsko arhivsko vijeće, 2012).[82] Dokument sam sebe opisuje na način:

---

[77] Hrvatski sabor. (2018). *Zakon o arhivskom gradivu i arhivima*, n. dj.
[78] Ibid.
[79] Hrvatski državni arhiv. (2021). Preuzeto 12. 12. 2021. s Nacionalni Arhivski Informacijski Sustav: http://arhinet.arhiv.hr/default.aspx
[80] URL Hrvatskog Državnog Arhiva: http://www.arhiv.hr/hr-hr/
[81] Autor disertacije nije uspio naći web sjedište firme Avicena Software d.o.o.
[82] Hrvatsko arhivsko vijeće. (2012). Opći popis gradiva s rokovima čuvanja. Preuzeto 12. 12 2021 s ArhiNET: http://arhinet.arhiv.hr/_Download/PDF/Opci_popis_gradiva_s_rokovima_cuvanja.pdf

"Opći popis s rokovima čuvanja je popis vrsta gradiva s rokovima čuvanja koji se odnosi na gradivo nastalo ili zaprimljeno obavljanjem administrativnih ili općih funkcija. To su funkcije koje obavlja svaki stvaratelj, bez obzira na to kojim se poslom bavi, kao što su na primjer upravljanje i organizacija, ljudski resursi, financije, upravljanje postrojenjima, nekretninama u vlasništvu, nabavom ili informacijskim sustavima. Prema odredbama članka 9. Pravilnika o vrednovanju te postupku odabiranja i izlučivanja arhivskog gradiva donosi ga Hrvatsko arhivsko vijeće na prijedlog Hrvatskog državnog arhiva." (Hrvatsko arhivsko vijeće, 2012)[83]

Za razliku od relevantnog zakona ovaj dokument navodi konkretne upute i rokove čuvanja arhivskog gradiva. Prema ovom dokumentu najduži propisani rok je 10 godina nakon zaključivanja spisa te su spisi koji spadaju u tu kategoriju označeni sa šifrom (Z+10). Zaključivanje spisa dokument definira kao godinu u kojoj je:

"...dokument (ugovor, odluka, pravilnik i sl.) prestao važiti ili je zamijenjen drugim odgovarajućim dokumentom." (Hrvatsko arhivsko vijeće, 2012)[84]

U slučaju ugovora ovo dodatno pojašnjeno:

"...rok čuvanja počinje teći istekom roka u kojem je potrebno osigurati mogućnost dokazivanja prava i obveza, odnosno u kojem je dopušteno osporavati ih." (Hrvatsko arhivsko vijeće, 2012)[85]

Na istom mjestu dokument Hrvatskog Arhivskog Vijeća navodi da u slučaju spora rok od 10 godina počinje teći po završetku spora. Bez ulaska u raspravu o rokovima za pobijanje ugovora i rokovima okončanja sudskih procesa moguće je tvrditi da je spise iz ove kategorije potrebno čuvati barem 10 godina. Ova (Z+10) kategorija obuhvaća sve skupa 18 vrsti spisa poput:

- ovlasti za korištenje žigova;
- dokumentacije u svezi s gubitkom, krađom, nestankom ili neovlaštenim korištenjem identifikacijske isprave;

[83] Hrvatsko arhivsko vijeće. (2012). Opći popis gradiva s rokovima, n. dj.
[84] Ibid.
[85] Ibid.

- dokumentacije o sporovima u svezi s vlasništvom i drugim stvarnim pravima na zemljištu;
- projektna dokumentacija građevinskih projekata;
- i druge.

S obzirom na rok od 10 (ili, s obzirom na sudske postupke, više godina) sasvim sigurno će isteći potpisni certifikati, a možda će doći i do povlačenja korištenih kriptografskih algoritama jer su se tijekom tog perioda pokazali kao nepouzdani (razlozi za ovo istraženi su u kasnijim poglavljima). Stoga smatram da su navedeni rokovi jasna indikacija potrebe za razvojem novog modela informacijskog sustava koji će riješiti problem dugoročne pohrane digitalno potpisanoga arhivskoga gradiva.

Osim ovog, općenitog arhivističkog zakona, za razumijevanje potrebe za ovim sustavom treba naglasiti i zakonsku mogućnost upotrebe digitalnih, to jest elektroničkih potpisa. U Republici Hrvatskoj upotreba elektroničkog potpisa regulirana je od 2005. kada je donesen Zakon o elektroničkoj ispravi (NN 150/05).[86] Zakon je 2018. zamijenjen Odlukom o proglašenju zakona o provedbi uredbe (EU) br. 910/2014 Europskog Parlamenta i Vijeća od 23. srpnja 2014. o elektroničkoj identifikaciji i uslugama povjerenja za elektroničke transakcije na unutarnjem tržištu i stavljanju izvan snage direktive 1999/93/EZ [87] koja je uvela novi zakon, temeljen na suvremenoj uredbi Europske Unije br. 910/2014.[88] Ova uredba poznatija je pod nazivom eIDAS (engl. *electronic IDentification, Authentication and trust Services*)[89] te je njome regulirana upotreba elektroničkih potpisa i povezanih servisa od povjerenja (engl. *trust services*), prije svega servisa za vremenske žigove.

Iako je danas stavljen izvan snage, Zakon o elektroničkoj ispravi je bio u upotrebi od 2005. do 2018. te je razumno očekivati da je značajna količina elektroničke arhivske građe koja se danas nalazi u arhivima ili će u njih doći nastala u ovom periodu. Zakon na općenit način

[86] Hrvatski Sabor. (2005). *Zakon o Elektroničkoj Ispravi.* Preuzeto 13. 12. 2021. s https://www.zakon.hr/z/272/Zakon-o-elektroni%C4%8Dkoj-ispravi
[87] Hrvatski Sabor. (2017). *Odluka o proglašenju zakona o provedbi uredbe (EU) br. 910/2014 Europskog Parlamenta i Vijeća od 23. srpnja 2014. o elektroničkoj identifikaciji i uslugama povjerenja za elektroničke transakcije na unutarnjem tržištu i stavljanju izvan snage direktive 199.* Preuzeto 12. 12. 2021. s https://narodne-novine.nn.hr/clanci/sluzbeni/2017_06_62_1430.html
[88] European Parliament and Council. (2014). REGULATION (EU) No 910/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Preuzeto 6. 2. 2020. s https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2014.257.01.0073.01.ENG
[89] European Parliament. (2014). eIDAS, n. dj.

definira elektroničke isprave i implicira upotrebu tehnologije bazirane na digitalnim certifikatima. Na primjer u članku 13. stavku 5.:

> *"Za svaku elektroničku ispravu mora u svim radnjama s elektroničkom ispravom postojati mogućnost provjere njene vjerodostojnosti, izvornosti i nepromjenjivosti."* (Hrvatski Sabor, 2005)[90]

Važniji dio zakona je onaj koji je omogućio upotrebu elektroničkih isprava na način na koji se koriste i papirnate isprave. Ovo je definirano člancima 5., 6. i 7. postavljanjem više zahtjeva na elektroničku ispravu kroz koje se ovdje neće detaljno prolaziti jer su u međuvremenu stavljeni izvan snage. Važan je početak 5. članka:

> *"Elektronička isprava ima pravnu valjanost kao i isprava na papiru…"* (Hrvatski Sabor, 2005)[91]

Ako je elektronička isprava uređena u skladu sa zahtjevima koji slijede ona je imala pravnu valjanost ekvivalentnu ispravi na papiru, što znači da je mogla biti korištena pri stvaranju spisa koji su arhivsko gradivo, uključujući one koji se po Hrvatskom arhivskom vijeću moraju čuvati barem 10 godina (Z+10 kategorija). Takvi spisi su nedavno ušli ili će tek ući u arhive te će dugotrajno očuvanje autentičnosti takvih spisa zahtijevati posebne mjere.

Od 2018. do danas na snazi je novi, ranije spomenuti, zakon o elektroničkim ispravama koji se temelji na uredbi Europske Unije br. 910/2014.[92] Ovim zakon detaljnije su definirani zahtjevi za elektroničke isprave i pružatelje usluge izdavanja takvih isprava. Ovi zahtjevi su, s obzirom na to da su aktualni i izravno primjenjivi na digitalne certifikate, detaljno razmotreni u idućem poglavlju. Tamo se detaljno objašnjava struktura i način funkcioniranja digitalnog certifikata pa tamo ima smisla i detaljno prikazati zahtjeve na takav sustav. U ovom poglavlju potrebno je samo napomenuti da je ova uredba, između ostaloga, definirala tri kategorije elektroničkog potpisa: jednostavni, napredni i kvalificirani. Kvalificirani elektronički potpis istovjetan je papirnatim identifikacijskim ispravama, u upotrebi je u Republici Hrvatskoj te više pružatelja usluga pruža usluge temeljene na njemu. Primjeri usluga koje se na neki način oslanjaju na kvalificirani elektronički potpis dostupni su od ovlaštenih pružatelja usluga poput Fine[93] (Financijska agencija) i AKD-a[94] (Agencija za komercijalnu djelatnost).

---

[90] Hrvatski Sabor. (2005). *Zakon o Elektroničkoj Ispravi*, n. dj.
[91] Ibid.
[92] European Parliament and Council. (2014). REGULATION (EU) No 910/2014, n. dj.
[93] Fina URL: https://www.fina.hr/
[94] AKD URL: https://www.akd.hr/

Fina nudi uslugu stvaranja kvalificiranog elektroničkog potpisa (i certifikata) koji se mogu koristiti na više načina, putem web aplikacije, desktop aplikacije, specijaliziranog portala ili individualiziranih rješenja u obliku programskih modula koji se ugrađuju u informacijski sustav korisnika[95].

AKD nudi, uz uobičajenije digitalne certifikate (koji su usklađeni sa zahtjevima na kvalificirani elektronički potpis) i usluge [96] koje uključuju NFC (engl. *near field communication*) kartice s kvalificiranim elektroničkim potpisom.[97]

Oba pružatelja usluga nude certifikate koji vrijede najviše dvije godine. Fina uz standardne kvalificirane potpise s trajanjem od dvije godine nudi i "soft" kvalificirani certifikat koji ne služi za potpisivanje već samo za pristup državnim servisima i traje pet godina.[98] Slična situacija je i s AKD certifikatima. Kvalificirani elektronički certifikati vrijede najviše dvije godine dok je NFC kartice moguće dobiti s trogodišnjim certifikatima[99].

Trajanje kvalificiranog certifikata od dvije godine, i zahtjev da (ovisno o vrsti spisa) njime potpisani spisi budu arhivirani najmanje 10 godine jasno ukazuju na potrebu istraživanja i razvoja novih modela dugotrajne pohrane koji će očuvati autentičnost.

Zadnji pravni akt koji je važan u ovom kontekstu je Opća uredba o zaštiti podataka, skraćeno zvan GDPR (engl. *General Data Protection Regulation*). GDPR je na razini Europske unije uveo mnogo novosti u načinu pohrane i upravljanja podacima, posebno osobnim podacima. U Republici Hrvatskoj on je reguliran Odlukom o proglašenju zakona o provedbi opće uredbe o zaštiti podataka[100] koja je u hrvatski zakon uvela i, u manjem dijelu proširila, hrvatski prijevod GDPR-a pod naslovom Opća uredba o zaštiti podataka. [101] Prijevod je istovjetan izvorniku[102] na engleskom pa će on biti korišten u ovoj raspravi.

---

[95] Financijska agencija. (2021). *Rješenja za elektronički potpis*. Preuzeto 13. 12. 2021. s Financijska Agencija: https://www.fina.hr/rjesenja-za-elektronicki-potpis
[96] Usluge su dostupne na URL:. https://www.id.hr/
[97] Agencija za komercijalnu djelatnost. (2021). *Proizvodi i usluge*. Preuzeto 13. 12. 2021. s id.hr: https://www.id.hr/hr/proizvodi-i-usluge
[98] Financijska agencija. (2021). *Osobni soft certifikat – FinaSoftCert*. Preuzeto 20. 12. 2021. s Financijska agencija: https://www.fina.hr/osobni-soft-certifikat-finasoftcert
[99] Financijska Agencija. (2021). *Cijene digitalnih certifikata i vremenskih žigova*. Preuzeto 13. 12. 2021. s Financijska Agencija: https://www.fina.hr/cijene
[100] Hrvatski Sabor. (2018). *Odluka o proglašenju Zakona o provedbi opće uredbe o zaštiti podatka.* Preuzeto 13. 12. 2021. s https://narodne-novine.nn.hr/clanci/sluzbeni/2018_05_42_805.html
[101] European Parliament and Council. (2018). *Uredba (EU) 2016/679*, n. dj.
[102] European Parliament and Council. (2016). *Regulation (EU) 2016/679 – General Data Protection Regulation.* Preuzeto 19. 11. 2022. s https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504

Većina GDPR-a nije važna za novi model koji je formuliran u ovom doktorskom radu, njime se uvode značajne proceduralne novosti vezane uz prikupljanje, obradu i pohranu osobnih podatka. Većina ovih pravila relevantna su za fazu stvaranja i korištenja osobnih podataka. Arhivi, pa tako i ovaj model, nastupaju kada je ova faza završena. Dio GDPR-a koji je od velike važnosti za novi model i kojega treba uvažiti je pravo na brisanje, to jest "pravo na zaborav" i pravo na ispravak. Ova prava definirana su člancima 16. i 17. Opće uredbe o zaštiti podatka.[103] S obzirom na to da ovi članci daju osobama pravo na izmjenu i brisanje evidentiranih podataka ona su sigurno relevantna za arhive, a posebno za sve sustave koji se temelje na nepromjenjivim podatkovnim strukturama, poput TrustChain modela koji je razvijen u sklopu ovog istraživanja.

Članak 16. daje osobama pravo na ispravak netočnih navoda. Ovo pravo nije ničim uvjetovano i potrebno ga je uvažiti.

TrustChain model u svoju nepromjenjivu podatkovnu strukturu uključuje minimalni set metapodataka, koji i nisu obavezni. Većina metapodataka pohranjuje se u pomoćnu bazu podataka koju je moguće mijenjati. Ipak, neki podaci ulaze u nepromjenjivu strukturu i za njih je potrebno osigurati mehanizme koji će omogućiti kasniji ispravak (u skladu s ovim zakonom). S obzirom na to da je ova podatkovna struktura nepromjenjiva problem je moguće riješiti ili isključivo izmjenama u pomoćnoj bazi ili stvaranjem novog zapisa u nepromjenjivoj podatkovnoj strukturi koji referencira stari zapis (koji sadrži grešku). Ova problematika dodatno je raspravljena u kasnijem poglavlju, koje detaljno objašnjava podatkovne strukture novog modela.

Članak 17. daje osobama pravo na brisanje podataka. Članak navodi da "ispitanik ima pravo od voditelja obrade (u ovom slučaju arhiv op.a.) ishoditi brisanje podatka…".[104] S obzirom na to da je ovo pravo uvjetovano ono predstavlja značajno manji problem nego pravo na ispravak (koje je bezuvjetno). Postoji više razloga zbog kojih osoba (ispitanik) ima pravo tražiti brisanje ali najvažnija od njih je navedena u članku 17. stavci 1. točci b koja navodi da se pravo na brisanje ostvaruje ako "ispitanik povuče privolu na kojoj se obrada temelji". Povlačenje privole nije ničim uvjetovano i može se dogoditi u bilo kojem trenutku. Ipak, GDPR, dozvoljava da pravo na brisanje nije ostvarivo u posebnim slučajevima. Članak 17. stavka 3. točka b navodi da se stavci 1. i 2. istog članka ne primjenjuju kada je obrada nužna iz razloga:

[103] Hrvatski Sabor. (2018). *Odluka o proglašenju Zakona o provedbi opće uredbe o zaštiti podatka*, n. dj.
[104] European Parliament and Council. (2018). *Uredba (EU) 2016/679*, n. dj.

"poštovanja pravne obveze kojom se zahtijeva obrada u pravu Unije ili pravu države članice kojem podliježe voditelj obrade ili za izvršavanje zadaće od javnog interesa ili pri izvršavanju službene ovlasti voditelja obrade." (European Parliament and Council, 2018)[105]

Točka d eksplicitno navodi arhivske svrhe:

"u svrhe arhiviranja u javnom interesu, u svrhe znanstvenog ili povijesnog istraživanja ili u statističke svrhe u skladu s člankom 89. stavkom 1. u mjeri u kojoj je vjerojatno da se pravom iz stavka 1. može onemogućiti ili ozbiljno ugroziti postizanje ciljeva te obrade." (European Parliament and Council, 2018)[106]

Prema ovim odredbama arhivi, posebno državni arhivi, izuzeti su od obveze poštovanja prava na zaborav. Svejedno, TrustChain model, zahvaljujući svojoj pomoćnoj bazi podataka, može funkcionirati bez pohrane ikakvih metapodataka (koji su osobni podaci), u nepromjenjivu strukturu podataka. Na ovaj način arhivi koji nisu državni mogu koristiti sustav, a i državni arhivi mogu, prema vlastitom nahođenju, u nekim slučajevima preskočiti zapis ovih metapodataka (u slučaju kada je moguće da će biti potrebno primijeniti pravo na brisanje).

2.5. Metapodaci TrustChain modela

Odabir metapodataka koji će se odnositi na zapise novog modela zadnji je dio istraživanja koji se odnosi na arhivistiku koja se bavi digitalnim zapisima, umjesto na tehnologije ili već postojeće sustave. Odabir skupa metapodataka koji će biti uključeni u TrustChain model od izuzetne je važnosti za razvoj modela pomoćne baze podataka čiji je cilj omogućiti brzu pretragu nepromjenjive podatkovne strukture, ulančanih zapisa, i izmjenu metapodataka, u skladu sa zahtjevima arhivske veze te pravu na ispravak i brisanje iz Opće uredbe o zaštiti podataka.

U ovoj fazi razvoja modela važno je odabrati minimalni skup metapodataka koji će ispuniti spomenute zahtjeve i stvoriti sustav koji je kompatibilan s postojećim arhivskim standardima. S obzirom na to da će se pomoćna baza podatka (koja služi za pohranu ovih metapodataka) temeljiti na postojećim tehnologijama za pohranu podatka koje omogućuju

[105] European Parliament and Council. (2018). *Uredba (EU) 2016/679*, n. dj.
[106] Ibid.

naknadne izmjene, to jest proširenja podatkovnih struktura nema potrebe od najranije faze razvoja definirati potpuni skup metapodataka (koji ispunjava sve zahtjeve). Umjesto toga u fazi razvoja modela potrebno je, kao što je navedeno, definirati minimalni skup metapodataka koji je koncipiran na način da kasnije, prema potrebama ustanova koje koriste model, može biti proširen u skladu s arhivskim standardima.

Rezultati istraživanja arhivskih standarda koje je prepoznalo minimalni set metapodataka pogodan za ovu svrhu objavili su autor i mentor ovog rada 2021. u članku u časopisu Computers.[107] Ovo poglavlje u potpunosti prihvaća rezultate ovog, nedavno provedenog, istraživanja te su oni ovdje preneseni ali i dodatno razrađeni s obzirom na to da doktorski rad, u usporedbi s člankom u časopisu, nije striktno prostorno ograničen te je moguće provesti detaljniju raspravu o razmatranim standardima i njihovim elementima.

Prvi korak u ovom razmatranju bio je prepoznati relevantne standarde koji definiraju metapodatke koji se koriste pri pohrani (digitalnih) objekata. Odabrana su tri arhivska standarda i jedan općeniti. Oni su:

1) ISAD(G) (engl. *General International Standard Archival Description*) međunarodni je standard koji definira elemente (metapodatke) koji moraju biti dio alata koji omogućuju pretragu arhivskog gradiva.[108] Osnovna svrha ISAD(G) standarda se podudara s svrhom TrustChain pomoćne baze podataka (pretraga ulančanih zapisa) te je njegovo uključivanje bio očigledan izbor. Standard je prihvatilo Međunarodno arhivsko vijeće (engl. *International Council on Archives,* ICA)[109] te je u upotrebi u digitalnim arhivima više arhivskih ustanova, uključujući Državni arhiv Ujedinjenog Kraljevstva Velike Britanije i Sjeverne Irske[110] i arhiv UNESCO-a (engl. *United Nations Educational, Scientific and Cultural Organization*).[111] Osim samog ISAD(G) standarda razmotren je i standard izveden iz njega: DACS.

[107] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.
[108] Brothman, B. (1992). Isad (g): general international standard archival description. *Archivaria, 34*, 17-32. Preuzeto 7. 1. 2022. s https://archivaria.ca/index.php/archivaria/article/view/11838/12790
[109] ICA URL: https://www.ica.org/en
[110] The National Archives. (2021). *Our digital cataloguing practices*. Preuzeto 14. 12. 2021. s The National Archives: https://www.nationalarchives.gov.uk/about/our-role/plans-policies-performance-and-projects/our-plans/our-digital-cataloguing-practices/
[111] UNESCO. (2021). *Organization of the Archives*. Preuzeto 14. 12. 2021. s UNESDOC Digital Library: https://unesdoc.unesco.org/archives/organization-of-the-archives

2) DACS (engl. *Describing Archives: A Content Standard*) [112] je prilagodba ISAD(G) standarda za upotrebu u arhivima Sjedinjenih Američkih Država, za prilagodbu i održavanje standarda brine se Društvo američkih arhivista.[113] U osnovi DACS je proširenje ISAD(G) standarda, svi elementi opisani u ISAD(G) prisutni su i u DACS sustavu, s iznimkom elementa "razine opisa" (engl. *level of description*). DACS je razmotren kao razrada ISAD(G) standarda.

3) PREMIS (engl. *Preservation Metadata: Implementation Strategies*) je standard "metapodataka sa svrhom očuvanja" (engl. *preservation metadata*).[114] PREMIS je specifično namijenjen dugotrajnom očuvanju digitalnih objekta (u ovom slučaju arhivskog gradiva), njegov cilj je definirati i standardizirati set metapodataka koji će omogućiti praćenje autentičnosti digitalnog objekta, koji će evidentirati eventualne izmjene u dokumentu nastale u procesu migracije s jednog digitalnog medija na drugi i druge metapodatke koji su specifični za digitalno gradivo. Iz ovog razloga metapodaci definirani PREMIS standardom su očigledan kandidat za TrustChain set metapodataka. Osim toga, PREMIS je izrastao iz OAIS (engl. *Open Archival Information System*) [115] referentnog modela koji je standardiziran on strane Međunarodne organizacije za standardizaciju. [116] OAIS je predmet velikog broj istraživanja, te je opće prihvaćen kao standard za razvoj (digitalnih) arhivskih sustava. Između ostalih OAIS je istraživan i u drugim disertacijama nastalima na Doktorskom studiju informacijskih i komunikacijskih znanosti Odsjeka za informacijske znanosti Filozofskog fakulteta Sveučilišta u Zagrebu,[117] prije svega u doktorskom radu Hrvoja Brzice, Koncept uspostave elektroničkoga arhiva u javnoj upravi, [118] koji je po više elementa preteča ovog istraživanja.

---

[112] Society of American Archivists. (2021). *Describing Archives: A Content Standard (DACS)*. Preuzeto 19.11.2022. s Society of American Archivists: https://www2.archivists.org/groups/technical-subcommittee-on-describing-archives-a-content-standard-dacs/describing-archives-a-content-standard-dacs-second-

[113] URL Društva Američkih Arhivista: https://www2.archivists.org/

[114] Caplan, P. (2009). Understanding Premis. Washington DC: USA: Library of Congress. Preuzeto 7. 1. 2022. s https://www.loc.gov/standards/premis/understanding-premis.pdf

[115] CCSDS. (2012). *REFERENCE MODEL FOR AN OPEN ARCHIVAL INFORMATION SYSTEM (OAIS).* Washington, DC, USA: Consultative Committee for Space Data Systems Secretaiat. Preuzeto 14. 12. 2021. s https://public.ccsds.org/Pubs/650x0m2.pdf

[116] International Organization for Standardization. (2018). *ISO 14721:2012*, n. dj.

[117] Odsjek za Informacijske i komunikacijske znanosti Filozofskog Fakulteta Sveučilišta u Zagrebu

[118] Brzica, H. (2018). *Koncept uspostave elektroničkoga arhiva*, n. dj.

4) Dublin Core (Metadata element set) skup je 15 elemenata, to jest metapodataka kojima je cilj opisati neki objekt. Dublin Core opisan je u dokumentu RFC5013[119] ali je na osnovu njega napisan i ISO standard 15836[120] i ANSI standard Z39.85.[121] Dublin Core je razmotren jer je riječ o univerzalnom standardu koji je, iako nije ograničen na digitalne objekte, našao primjenu u raznim informacijskim sustavima. Kompatibilnost TrustChain modela s ovim standardom svakako će olakšati upotrebu modela u arhivima koji primarno rade s digitalnim zapisima.

Usporedba ovih standarda, prema osnovnoj svrsi, ukupnom broju elemenata (metapodataka) i broju obaveznih elemenata prikazana je u tablici 1.

Tablica 1. Pregled razmotrenih standarda za metapodatke. Izvor: Stančić, Bralić. Digital Archives Relying on Blockchain: Overcoming the Limitations of Data Immutability. 2021

Standard	Arhivski standard?	Broj metapodataka	Broj obaveznih metapodataka
DACS	Da	25	10
Dublin Core	Ne	15	0
ISAD(G)	Da	26	6
PREMIS	Da	15	2

U tablici 2 prikazani su ovi elementi, to jest metapodaci. Svrha tablice je prikazati odnos ova četiri standarda kroz njihove obavezne elemente, to jest u slučaju da se obavezni metapodaci ne preklapaju pokušati pronaći ekvivalentne opcionalne elemente u drugim standardima.

[119] Weibel, S., & Baker, T. (2007). *RFC5013: Dublin Core Metadata for Resource Discovery.* Preuzeto 14. 12. 2021. s Internet Engineering Task Force: https://www.ietf.org/rfc/rfc5013.txt
[120] International Organization for Standardization. (2017). *ISO 15836-1:2017 Information and documentation — The Dublin Core metadata element set — Part 1: Core elements.* ISO. Preuzeto 14. 12. 2021. s https://www.iso.org/standard/71339.html
[121] ANSI. (2013). *ANSI/NISO Z39.85 – The Dublin Core Metadata Element Set.* Preuzeto 14. 12. 2021. s http://www.niso.org/publications/ansiniso-z3985-2012-dublin-core-metadata-element-set

Tablica 2. Pregled obaveznih metapodataka podataka razmotrenih standarda. Izvor: Stančić, Bralić. Digital Archives Relying on Blockchain: Overcoming the Limitations of Data Immutability. 2021

DACS	**Dublin Core**	**ISAD(G)**	**PREMIS**
Reference Code Element (M)	Identifier (O)	Reference code (M)	objectIdentifier (M)
X	Type (O)	X	objectCategory (M)
Title element (M)	Title (O)	Title (M)	originalName (O)
Name of creator element (M)	Creator (O)	Name of creator (M)	signatureInformation/ signer (O)
Date element (M)	Date (O)	Date of creation (M)	creatingApplication/date CreatedByApplication (O)
X	X	Level of description (M)	X
Extent element (M)	X	Extent of the unit description (M)	X
Scope and Content Element (M)	X	Scope and content (O)	X
Name and Location of Repository Element (M)	X	Existence and location of originals (O)	storage/contentLocation (O)
Conditions Governing Access Element (M)	X	Conditions governing access (O)	X
Languages and Scripts of the Material Element (M)	X	Language/scripts of material (O)	X
Rights Statements for Archival Description (M)	Rights (O)	X	linkingRightsStatement Identifier (O)

Uz naziv elementa u tablici su prisutni simboli:

- M (engl. *mandatory*) – ovime su označeni obavezni elementi standarda.
- O (engl. *optional*) – ovime su označeni opcionalni elementi koji su u tablicu uključeni u pokušaju mapiranja ekvivalentnih elemenata preko sva četiri standarda, to jest u pokušaju pronalaženja opcionalnih elemenata jednog standarda koji odgovaraju obaveznim elementima drugog.
- X – ovime su označeni elementi koji su u drugim standardima obavezni ali u standardu u čijem se stupcu nalaze nije pronađen ekvivalentni element.

Postoji očigledna razlika između ukupnog broja metapodataka između ISAD(G) i DACS standarda s jedne strane, te PREMIS i Dublin Core s druge. Ova razlike je još izraženija kada se gleda broj obaveznih metapodataka.

DACS i ISAD(G) standardi nisu odabrani kao početne točke za razvoj modela metapodataka TrustChain modela jer sadrže velik broj metapodataka koji su obavezni, a dio ovih metapodataka nije primjenjiv na digitalne zapise, ili nije od kritične važnosti pa po tome ne spada u minimalni set metapodataka, što je bio cilj ovog istraživanja objavljenog u časopisu Computers [122]. Jedan primjer ovakvog obaveznog metapodatka je DACS-ov "Name and Location of Repository Element". Lokacija repozitorija nije kritičan podatak u digitalnom arhivu, posebno ne u distribuiranom sustavu. Drugi primjer je ISAD(G)-ov obavezni "Level of description". Svrha ovog elementa je klasificirati razinu dostupnih metapodataka o konkretnom arhivskom zapisu. Iako je riječ o veoma smislenom podatku, on nije kritičan u digitalnom arhivu. U takvom okruženju veoma je lako automatski procijeniti do koje mjere je zapis opisan metapodacima. Oba sustava, ISAD(G) i DACS, su univerzalni, trude se biti primjenjivi i na klasični i na digitalni arhiv te iz ovog razloga uključuju metapodatke koji se nikako ne mogu svrstati u minimalni set nužnih za digitalno potpisano arhivskog gradivo (i PREMIS podržava ne-digitalno gradivo ali je puno fleksibilniji u odabiru metapodataka – on sadrži samo dva obavezna elementa). S druge strane, neki elementi, poput onih koji pokrivaju digitalni potpis, nedostaju. Iako ovi standardi nisu odabrani za primjenu u TrustChain modelu potrebno je naglasiti da oni mogu biti podržani kao što je to vidljivo iz tablice 2. njihove obavezne elemente je (uglavnom) moguće mapirati na elemente drugih standarda, a s obzirom da je pomoćnu bazu podataka moguće naknadno mijenjati, podršku za te standarde može se po potrebi kasnije dodati.[123]

Dublin Core uopće ne sadrži obavezne elemente, kompaktan je (u potpunosti se sastoji od 15 elementa), te je sve njegove elemente moguće preslikati u PREMIS i DACS standarde, dok "Rights" element nema ISAD(G) ekvivalent. U slučaju ovog standarda trebalo bi ga prihvatiti u cijelosti ali autori su se odlučili na prihvaćanje PREMIS standarda. S obzirom da se svi elementi Dublin Core-a izravno mapiraju u PREMIS elemente, može se reći da je i ovaj standard podržan u osnovnoj verziji TrustChain modela (iako on nije odabran za kao temelj modela).

Kao početna točka za razvoj modela metapodataka odabran je PREMIS standard. Zaključci izvornog istraživanja na temelju kojih je odabran PREMIS su citirani niže, nakon prijevoda na hrvatski jezik.

[122] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.
[123] Ibid.

"PREMIS se pokazao kao najprimjenjiviji standard za potrebe ovog istraživanja. Razlozi ovome su slijedeći. Prvo, PREMIS ima samo dva obavezna elementa ali kada se koristiti pri opisu digitalnih objekata, na primjer, datoteka ili nizova bitova, on zahtjeva još barem jedan dodatni element "objectCharacteristics/format". *Ovaj element je prikladan za TrustChain sustav jer opisuje način digitalnog zapisa koji koristi objekt. Standard je, kada se gledaju ostali elementi, vrlo širok i pokriva veliki broj metapodataka, od kojih je većina opcionalna. Čak i kada su elementi označeni kao obavezni, to uglavnom znači da su uvjetno obavezni (kao, na primjer, u slučaju opisanog* "/format" *elementa*[124]*), ili su primjenjivi samo u slučaju kada se koriste drugi elementi.*[125] *Ovo dozvoljava visoku razinu fleksibilnosti u primjeni standarda. Drugo, PREMIS uključuje metapodatke koji pokrivaju digitalne potpise, a to su, za TrustChain model, kritični podaci. Treće, PREMIS podržava pohranu metapodataka o arhivskoj vezi kroz svoje* "relationship" *elemente. Konačno, PREMIS je predviđen da se koristi u XML*[126] *notaciji, koja je vrlo slična JSON*[127] *notaciji. Ovo pojednostavljuje stvaranje podatkovnog modela temeljenog na MongoDB sustavu.*[128]*"* (Stančić & Bralić, 2021)[129]

Na osnovu ovog odabira i metapodataka koji su obuhvaćeni obaveznim i opcionalnim metapodacima PREMIS standarda definirani su metapodaci koji će obuhvatiti TrustChain model. Svi metapodaci iz izvornog istraživanja su prihvaćeni[130] i u ovom modelu, te su dodani neki novi, za kojima se pokazala potreba tijekom razvoja konačnog modela. Neki metapodaci su dodani i kao rezultat ranijeg razmatranja zakonskih propisa, prije svega Opće uredbe o zaštiti podatka. Iz istog razloga je, u odnosu na izvorni model, proširena upotrebna nekih metapodataka. Daljnji tekst sadrži parafraziran i, zbog navedenih razloga, proširen dio rada Stančića i Bralića iz 2021.

[124] Koji je obavezan samo u slučaju digitalnih objekata.
[125] PREMIS Editorial Committee. (2015). *Premis data dictionary for preservation metadata.* Preuzeto 19. 11. 2022. s USA: Library of Congress: https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf
[126] Engl. *extensible markup language*, standard zapisa podataka u ljudski čitljivom obliku
[127] Engl. *JavaScript Object Notation*, standard zapisa podataka u ljudski čitljivom obliku koji je proizašao iz programskog jezika JavaScript.
[128] NoSQL distribuirani sustav za pohranu podataka koji je odabran kao tehnološka osnova pomoćne baze podataka TrustChain modela.
[129] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.
[130] Ibid.

PREMIS sadrži samo dva obavezna metapodataka, ali s obzirom na to da model koji se razvija u ovom istraživanju služi za pohranu digitalno potpisanih (digitalnih) zapis, set obaveznih metapodataka je proširen na pet, oni su:

1) *objectIdentifier,* identifikator zapisa – jedinstveni identifikator zapisa u TrustChain sustavu.

2) *objectCategory,* kategorija zapisa – u PREMIS standardu ovaj metapodatak klasificira zapis. U TrustChain modelu ova klasifikacija se koristi da bi razlikovali zapise koji se odnose na potpisane datoteke i zapise koji se odnose na digitalne certifikate. Ova razlika rezultat je spajanja TrustChain 1 i 2 modela u jedinstveni model. Ovo je detaljnije objašnjeno u kasnijim poglavljima. U kombinaciji s, kasnije objašnjenim, metapodatkom 18, u ovom polju može biti pohranjena informacija koja označava zapis koji ne postoji u TrustChain sustavu.

3) *originalName*, naziv izvornika – naziv ili šifra (identifikator) zapisa prije ulaska u TrustChain sustav.

4) *dateCreatedByApplication,* vrijeme stvaranja – vrijeme stvaranja izvornog (digitalno potpisanoga) dokumenta ili digitalnog certifikata. Ovo vrijeme razlikuje se od vremena stvaranja TrustChain zapisa.

5) *Format,* oblik zapisa – podaci o obliku zapisa. Sadržaj ovog polja ovisi o sadržaju polja "objectCategory". U slučaju digitalno potpisane datoteke ovo se odnosi na njezin oblik (format), na primjer, .pdf ili .docx datoteka. U slučaju digitalnog certifikata ovo polje sadrži standard po kojem je certifikat izrađen. Danas je ovo gotovo sigurno x.509 ali u budućnosti će se možda koristiti i druge standarde.

Uz ovih pet metapodataka koji su preuzeti iz PREMIS standarda, kao obavezni dodana su još tri koji su specifični za funkcioniranje TrustChain modela. Oni su:

6) *TrustChainBlockID,* identifikator TrustChain bloka. Ovaj identifikator služi za oznaku bloka nepromjenjivog TrustChain zapisa u kojem su sadržani najvažniji elementi vezani uz ovaj zapis.

7) *TrustChainTimeStamp*, TrustChain vremenski žig. Ova kombinacija vremena i datuma označava trenutak u kojem je nastao TrustChain zapis. Ovo vrijeme se razlikuje od datuma stvaranja izvornog zapisa i uvijek mora biti nakon njega.

8) *TrustChainBlockCorrID*, identifikator TrustChain bloka koji sadrži ispravak informacija. Ovaj identifikator, kao i onaj pod brojem 6, identificira blok u kojem se nalaze ispravljeni podaci. Ovi ispravci su nužni (to jest u nekim slučajevima će biti nužni) da bi sustav bio usklađen s ranije objašnjenom Općom uredbom o zaštiti podatka – konkretno, s pravom na ispravak koje je u njoj uvedeno. Polje može sadržavati niti jedan ili niz identifikatora koji upućuju na TrustChain blokove s ispravljenim podacima.

S time završavaju obavezni elementi svakog zapisa. Idućih osam elemenata preuzeto je iz PREMIS standarda i njihova je funkcija opis digitalnih potpisa (i certifikata):

9) *signer*, potpisnik – ime osobe ili naziv ustanove ili tvrtke, ili neki drugi identifikator potpisnika koji se navodi u digitalnom potpisu.

10) *signatureMethod*, način potpisa – šifra, identifikator algoritama korištenih pri šifriranju i hashiranju digitalnog potpisa.

11) s*ignatureValidationRules,* pravila provjere potpisa – PREMIS ovaj element definira kao "postupci koje treba provesti da bi se potvrdio digitalni potpis". Ovo se odnosi na posebne korake koje je potrebno provesti pri provjeri digitalnog potpisa. Ovi posebni koraci mogu biti posljedica posebnih pravila unosa digitalnog zapisa u konkretni digitalni arhiv, migracije podatka, uvođenja novih tehnologija u sam digitalni potpis ili drugo.

12) *signatureProperties,* osobine potpisa – dodatni podaci o stvaranju digitalnog potpisa. Ovo polje u trenutnom modelu nema posebnu ulogu, ono je uključeno kao opcionalni element koji će se možda koristiti u budućnosti.

13) *signatureValue,* digitalni potpis – ovo polje sadrži sam potpis, u izvornom binarnom obliku.

14) *keyInformation*, javni ključ. Ovo polje sadrži javi ključ koji se koristi pri potvrdi potpisa, u izvornom binarnom obliku.

15) *signatureEncoding*, podaci o načinu potpisa – element sadrži podatke o načinu na koji su zapisani binarni podaci u polju 13, *"signatureValue"* i 14, *"keyInformation"*. Na primjer, ASCII,[131] Unicode[132] ili drugi.

Zadnja četiri, opcionalna, elementa preuzeta su iz PREMIS modula o odnosima (engl. *relationship*) zapisa. Njihova uloga u TrustChain modelu je očuvanje podataka o arhivskoj vezi. Oni su:

16) *relationshipType*, vrsta arhivske veze – prema PREMIS-u "visoka razina[133] kategorizacije vrste odnosa". Ovaj opcionalni element nosi vrstu arhivske veze prema kategorizaciji koju koristiti arhivska ustanova odgovorna za zapis.

17) *relationshipSubType*, naziv arhivske veze – PREMIS ovo polje definira kao podvrstu odnosa između zapisa. U kontekstu arhivske veze, ako tako detaljna kategorizacija vrsti (arhivskih) veza nije potrebna, polje može biti iskorišteno za naziv arhivske veze ili šifru kojom se konkretna veza identificira u arhivskoj ustanovi.

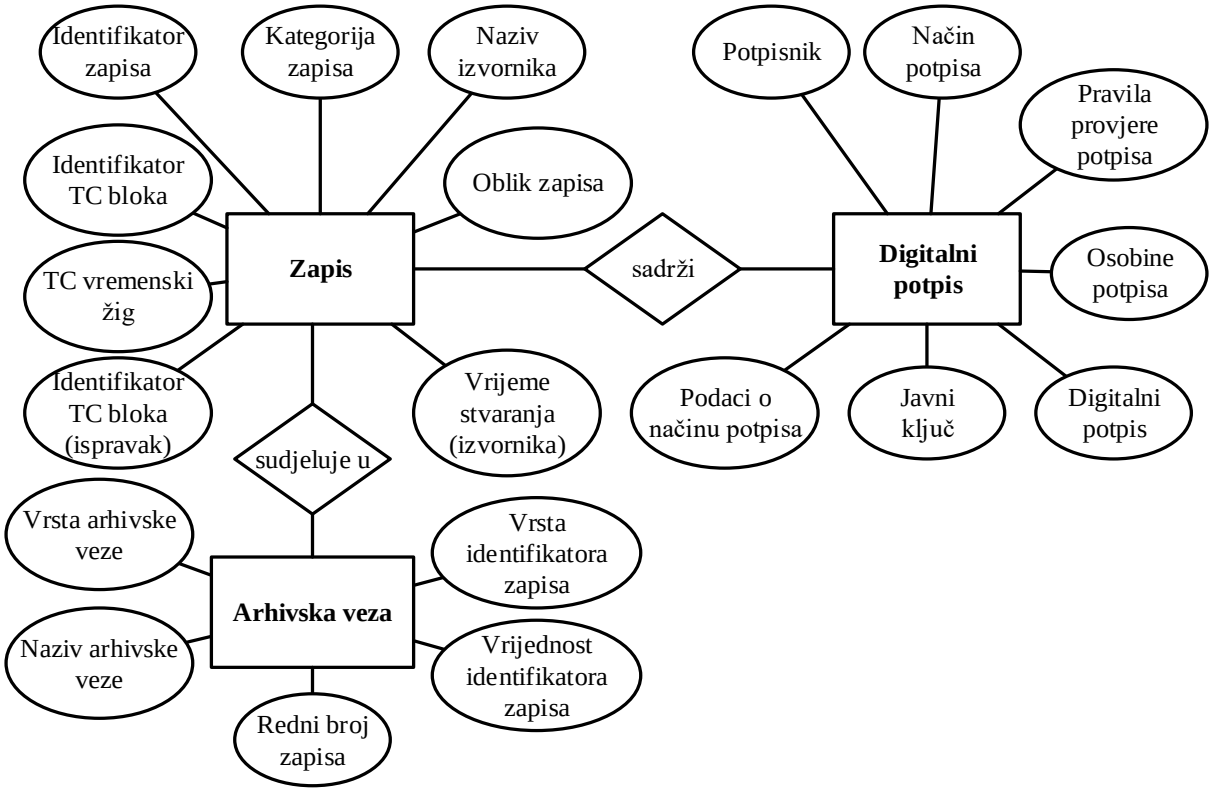
18) *relatedObjectIdentifierType*, vrsta identifikatora zapisa – prema PREMIS-u ovo polje definira domenu identifikatora zapisa u kojoj je identifikator jedinstven. Ovakvo polje nije potrebno za funkcioniranje TrustChain sustava, jer su svi identifikatori u sustavu su jedinstveni. Ipak, ovo opcionalno polje je uključeno jer dozvoljava stvaranja zapisa koji nisu dio TrustChain sustava i na taj način omogućava stvaranje opsežnih podataka o arhivskoj vezi. U ovom slučaju ovo polje navodi identifikator vanjskog sustava u kojem je sadržan relevantni sustav. Ovakva situacija rezultat je arhivske veze koja obuhvaća digitalno potpisane zapise i one koji to nisu, a možda ni nisu digitalni objekti (i onda ne mogu biti dio TrustChain sustava). Ovaj način pohrane podataka o arhivskoj vezi zapisa koji nisu sadržani u TrustChain sustavu značajno komplicira evidenciju zapisa te će gotovo sigurno rezultirati višestrukom pohranom metapodataka o istom zapisu. Upotreba ovog polja ostavljena je na izbor arhivskim ustanovama.

[131] ASCII engl. *American Standard Code for Information Interchange.* Standardni način zapisa u ekeltroničkoj komunikaciji.
[132] Unicode standard je proširen set znakova koji se standardno koristi u kompleksnijim zapisima, poput složenih tekstualnih datoteka.
[133] U odnosu na idući element, podvrstu kategorizacije.

19) *relatedObjectIdentifierValue*, vrijednost identifikator zapisa – identifikator povezanog zapisa. Ovo je polje koje može sadržavati više identifikatora zapisa iz TrustChain sustava (ili izvan njega, u slučaju upotrebe polja 18).

20) *relatedObjectSequence*, redni broj zapisa – redni broj ovog zapisa u odnosu na druge zapise koji sudjeluju u arhivskoj vezi.

Konačni korak u ovoj fazi istraživanja je izrada logičkog model prema Chenovoj notaciji[134]. Ova shema daje vizualni prikaz odnosa metapodataka i služi kao osnova za razvoj konkretnih implementacija u kasnijim poglavljima (slika 1).



Slika 1. Logički model metapodataka TrustChain modela. Izrađeno prema: Stančić, Bralić. Digital Archives Relying on Blockchain: Overcoming the Limitations of Data Immutability. 2021

[134] Chen, P. P.-S. (1976). The entity-relationship model—toward a unified view of data. *ACM transactions on database systems (TODS)*, 9-36. Preuzeto 14. 12. 2021. s https://dspace.mit.edu/bitstream/handle/1721.1/47638/entityrelationsh00chen.pdf?s

2.6. Zaključak

Cilj ovog poglavlja bio je postaviti temelje istraživanja definirajući zahtjeve koje novi model mora ispuniti. Osim toga poglavlje je i argumentiralo razloge za razvoj novog modela.

U prvom dijelu objašnjeni su temeljni pojmovi arhivistike i diplomatike: autentičnost i pouzdanost zapisa. Pokazano je da se, prema načelima diplomatike, smatra da je zapis pouzdan ako je stvoren u skladu s pravnim i kulturološkim zahtjevima podneblja u kojem je nastao, te ako je na njemu jasno naveden datum i potpis osobe koja ga je stvorila (koja mora biti ovlaštena za stvaranje takvog zapisa). S druge strane zapis je autentičan ako je od trenutka u kojem je stvoren ostao nepromijenjen. Ove dvije osobine zapisa (pouzdanost i autentičnost) zajedno dokazuju je zapis pravovaljan (engl. *genuine*) te se da se može koristiti kao dokaz u pravnom postupku, što je i izvorna svrha diplomatike. Objašnjeno je da se ove postavke diplomatike (discipline koja se bavi proučavanjem individualnih dokumenata) primjenjuju i na arhivistiku (znanost o arhivima, koja uzima u obzir cjelokupnu građu neke ustanove).

U idućem dijelu definiran je pojam arhivske veze kao "…mreže odnosa svakog zapisa s drugim zapisima iz istog skupa."[135] te je pokazano da je održavanje ove veze odnosa, iako je često zanemareno u sustavima za pohranu podataka koji nisu striktno arhivski, kritično za očuvanje pouzdanosti zapisa.

S obzirom na sve navedeno može se utvrditi da su zahtjevi arhivistike i diplomatike koji su primjenjivi na sustav za dugotrajno očuvanje digitalnih zapisa takvi da:

1) zapis sadrži vrijeme nastanka.
2) zapis sadrži potpis koji identificira autora.
3) zapis od trenutka potpisivanja treba ostati nepromijenjen i
4) zapis mora imati očuvanu arhivsku vezu.

Pokazano je da je točke od jedan do tri moguće privremeno riješiti upotrebom digitalnog potpisa. Digitalni potpis sadrži podatak koji identificira autora i vrijeme nastanka te, kao što će biti objašnjeno u idućem poglavlju, garantira integritet podataka. Na zahtjev broj četiri ne postoji izravno primjenjiva tehnologija te se on rješava na razini podatkovnih struktura modela. Ovo je tema kasnijeg poglavlja. Do raskola u razumijevanju termina integriteta i autentičnosti podatka između arhivistike i računarstva dolazi zbog nedostatka potrebe da dokazivanje

[135] Duranti, L. (1997). The archival bond, n. dj.

autentičnosti u računarstvu bude (dugo)trajno. Digitalni certifikati su kratkotrajni te mnogi protokoli za provjeru integriteta podataka uopće ne uzimaju u obzir identitet autora. Provjera je li zapis stvoren u skladu s pravnim i drugim zahtjevima vremena i podneblja je zahtjev diplomatike koji nije moguće riješiti na univerzalan način. Stručnjak mora proučiti svaki individualni zapis.

Uz ovo raspravljeno je na koji način digitalni potpis rješava (možda je bolje reći zaobilazi) problematiku vezanu uz činjenicu da digitalni zapisi nisu striktno vezani uz medij na kojem su pohranjeni, problem koji je bio izvor velikog broja istraživanja o dokazivanju autentičnosti digitalnih zapisa.

U dijelu poglavlja čija su tema zakonski i drugi propisi pokazano je da je digitalni potpis na području Europske unije, kada je stvoren u skladu s relevantnim uredbama i zakonima, istovjetan ručnom potpisu ili žigu. Ovime je potvrđena osnovna pretpostavka diplomatike, da se dokazivanjem autentičnosti i pouzdanosti zapisa (što se u ovom slučaju postiže digitalnim potpisom) stvara mogućnost da se zapis koristi kao pravni dokaz.

Osim toga objašnjen je temeljni problem upotrebe digitalnih potpisa kao mehanizma dugotrajnog očuvanja autentičnosti arhivskog gradiva. Pokazano je da propisi Republike Hrvatske i pravila nadležnih arhiva traže da se određeni zapisi čuvaju barem 10 godina od kada je završen postupak tijekom kojeg su nastali. Pokazano je i da digitalni certifikati, mehanizam koji omogućava potvrdu identiteta autora potpisa (i potpisanog zapisa), istječu nakon dvije godine. Ovaj nesrazmjer između vremena tijekom kojeg je nužna dokazivost autentičnosti i pouzdanosti zapisa s jedne strane i životnog vijeka mehanizma koji to omogućava s druge strane razlog su ovog istraživanja. Cilj istraživanja je, u skladu sa svim rečenim, utvrditi postoji li mogućnosti dugotrajnog očuvanja autentičnosti i pouzdanosti digitalno potpisnog arhivskog gradiva upotrebom tehnologije ulančanih zapisa.

Osim toga, cilj istraživanja je utvrditi postoji li mogućnost upotrebe iste tehnologije za dugotrajno očuvanje arhivske veze (koja je također nužna za očuvanje pouzdanosti i autentičnosti zapisa). Temelji podatkovne strukture koja će omogućiti očuvanje arhivske veze (zahtjeva broj 4) postavljeni su u zadnjem dijelu poglavlja koje proučava metapodatke koji će biti uključeni u model. Na osnovu postojećih standarda prikazan je minimalni set metapodataka koji su potrebni za očuvanje podataka o arhivskoj vezi i podataka koji će omogućiti učinkovitu pretragu TrustChain ulančanih zapisa.

3. Digitalni potpis i digitalni certifikat

Digitalni potpis prisutan je u računarstvu već više od pola stoljeća.[136] Sam potpis odlično funkcionira u kontekstu računarstva, za čije je potrebe i stvoren. Ove potrebe i svrha digitalnog potpisa se često odnose na potpuno tehničke, krajnjem korisniku često i nepoznate procese poput digitalnog potpisa koje koristi TSL mrežni protokol[137] ili potpisivanja digitalnih datoteka.[138] Od 1989. digitalni potpis se počinje probijati kao alat za autentifikaciju korisnika (a ne samo digitalnih servisa) kada se pojavljuje prvi komercijalni set alata (poput kalendara, e-mail servisa, adresara i drugih kolaboracijskih alata) koji je omogućio digitalni potpis: Lotus Notes.[139] Od tada do danas digitalni potpis probio se u razne, korisniku izravno korisne procese, poput potpisivanja kratke elektroničke poruke (engl. *e-mail*) ili digitalnog dokumenta, poput PDF datoteka[140] ili autorizacije elektroničkih servisa, poput servisa e-građani.[141] S obzirom na sve navedene upotrebe te širenje interneta i povezanih servisa rasla je i potreba za digitalnim potpisom, a s njom i broj dokumenata koji su njime potpisani. Posljedično, ovo podrazumijeva i značajan porast digitalno potpisanih dokumenta u raznim arhivskim ustanovama.

Suvremeni digitalni potpis obuhvaća veliki broj različitih standarda koji omogućuju rješenje navedenih procesa. Da bi se postupak digitalnog potpisivanja mogao odvijati na standardiziran način bilo je potrebno razviti više kriptografskih algoritama, prije svega hash funkcije i algoritme javnog ključa, te definirati niz standarda koji se na njih oslanjaju. Ovo uključuje standarde koji opisuju ranije navedene algoritme, standarde koji opisuju način dodavanja potpisa u digitalni zapis ili dokument, distribuciju ključeva, distribuciju i povlačenje

[136] Diffie, W., & Hellman, M. (1976). New directions in cryptography. *IEEE transactions on Information Theory*, 22(6), 644-654. Preuzeto 20. 12. 2021. s https://caislab.kaist.ac.kr/lecture/2010/spring/cs548/basic/B08.pdf
[137] Das, M., & Samdaria, N. (2014). On the security of SSL/TLS-enabled applications. *Applied Computing and informatics*, 68-81. doi: https://doi.org/10.1016/j.aci.2014.02.001
[138] Cooper, D., Regenscheid, A., Souppaya, M., Bean, C., Boyle, M., Cooley, D., & Jenkins, M. (2018). *Security considerations for code signing.* NIST. Preuzeto 20. 12. 2021. s https://www.encryptionconsulting.com/wp-content/uploads/2020/01/NIST-code-sign-whitepaper.pdf
[139] IBM. (2007). *The History of Notes and Domino.* IBM. Preuzeto 20. 12. 2021. s https://www.notesmail.com/home.nsf/ls-NDHistory-pdf.pdf
[140] Brzica, H., Herceg, B., & Stančić, H. (2013). Long-term Preservation of Validity of Electronically Signed Records. *INFuture2013: Information Governance* (str. 147-158). Zagreb: Filozofski Fakultet Sveučilišta u Zagrebu. Preuzeto 20. 12. 2021. s http://darhiv.ffzg.unizg.hr/id/eprint/8291/1/4-03%20Brzica%2C%20Herceg%2C%20Stancic%2C%20LTP%20of%20Validity%20of%20Electronically%20Signed%20Records.pdf
[141] Financijska agencija. (2021). *Osobni soft certifikat*, n. dj.

certifikata i ostale. Većina ovih standarda je otvorenog tipa i pod pokroviteljstvom IEEE,[142] NIST[143] ili ETSI[144] organizacije.

Uz koncept digitalnog potpisa usko je vezan i digitalni certifikat. Termini se često, pogrešno, koriste kao sinonimi ili se smatra da jedan od ova dva pojma obuhvaća drugi. U izvornom obliku digitalnog potpisa, digitalni certifikat ne postoji, on je sveden na privatni i javni ključ. Kroz ovu disertaciju pojam digitalni potpis označava krajnji rezultat procesa digitalnog potpisivanja, to jest dodavanja digitalne oznake nekom digitalnom zapisu ili dokumentu. Ova oznaka je jedinstvena za vlasnika digitalnog certifikata koji je korišten u procesu. Digitalni certifikat je skup podataka u digitalnom obliku koji je jedinstven i čijom je upotrebom moguće stvoriti digitalni potpis te koji jednoznačno identificira vlasnika certifikata, to jest autora potpisa. Najvažniji primjer standarda koji definira digitalni certifikat je x.509.[145] Ovaj standard i po njemu nazvani x.509 digitalni certifikat temelj je većine sustava koji koriste suvremene digitalne potpise.

Ovo poglavlje dati će tehnički pregled koncepata digitalnog potpisa i digitalnog certifikata, kriptografskih algoritama koji su ih omogućili i njihovih primjena te na taj način objasniti drugo od dva područja (prvo se odnosilo na zahtjeve diplomatike i arhivistike) koja su pretpostavka za razumijevanje problematike i modela koje ova disertacija predstavlja. Cilj ovog poglavlja je dokazati da se digitalni potpis može koristiti za dokazivanje autentičnosti i pouzdanosti arhivskih zapisa u skladu s ranije postavljenim zahtjevima. Ovo će se postići kritičkom analizom korištenih tehnologija.

3.1. Kriptografska osnova digitalnog potpisa

Za razumijevanje digitalnog potpisa i certifikata nužno je prvo razumjeti kriptografske koncepte koji su ih omogućili te je to i tema ovog poglavlja. Poglavlje pruža tehnički pregled digitalnog potpisa, digitalnog certifikata i povezanih pojmova iz područja kriptografije. Ovo su prije svega, proces šifriranja i dešifriranja, to jest enkripcije (zakrivanja) i dekripcije

[142] IEEE URL: https://www.ieee.org/
[143] NIST URL: https://www.nist.gov/
[144] ETSI URL: https://www.etsi.org/
[145] Adams, C., Cain, P., Pinkas, D., & Zuccherato, R. (2001). *RFC 3161: Internet X.509 Public Key Infrastructure Time-Stamp Protocol (TSP).* Preuzeto 6. 17. 2019. s Internet Engineering Task Force (IETF): https://www.ietf.org/rfc/rfc3161.txt

(raskrivanja) podataka upotrebom sustava s javnim ključem i kriptografskog sažetka, ili preciznije *hash* (engl.) funkcije.

U kontekstu kriptografije razlikujemo simetričnu i asimetričnu kriptografiju.[146] Simetrična kriptografija je poznata kroz velik dio ljudske povijesti i obuhvaća pojmove poput Cezarove šifre te ostale supstitucijske i transpozicijske sustave šifriranja. Problem ovakvih sustava je oslanjanje na jedinstveni ključ za šifriranje i dešifriranje poruka. Vjerojatno najvažniji primjer simetričnog algoritma kojim je moguće dokazivati autentičnost zapisa je takozvani, engl. *keyed hash message authentication code*, skraćeno HMAC algoritam.[147] Iako u računarstvu postoje primjeri takve primjene, kao što je to u slučaju TSL protokola, algoritmi bazirani na simetričnom ključu , zbog njegove nesigurnosti – ključ koji je formirala jedna strana mora biti dostavljen drugoj strani prije početka komunikacije šifriranim porukama – u pravilu se smatraju nepogodnim za digitalne potpise te sustavi za digitalni potpis temeljeni na njima do danas nisu zaživjeli.

Suvremeni digitalni potpis omogućen je razvojem asimetrične kriptografije koja se često koristi i kao sinonim za sustave javnog ključa,[148] engl. *Public Key Infrastructure* ili skraćeno PKI.

Asimetrična kriptografija i digitalni potpis su izvorno omogućeni upotrebom RSA kriptosustava,[149] koji je prisutan već više od 40 godina (od 1977.godine) i koji se detaljnije razmatra u kasnijem poglavlju. Suvremeni digitalni potpisi se oslanjaju na sustav infrastrukture javnog ključa.

Suvremeni sustavi infrastrukture javnog ključa omogućavaju dokazivanje identiteta autora i integriteta potpisanih podataka u digitalnom okruženju. S obzirom na to da pod pojmom integritet podataka podrazumijevamo da su oni ostali nepromijenjeni možemo reći da sustav omogućava primjenu termina pouzdanosti, putem dokazivanja identiteta autora, i autentičnosti, putem dokazivanja integriteta podataka, kako su definirani u sklopu diplomatike.

---

[146] Paar, C., & Pelzl, J. (2009). *Understanding cryptography: a textbook for students and practitioners.* Springer Science & Business Media. doi: https://doi.org/10.1007/978-3-642-04101-3

[147] Krawczyk, H., Bellare, M., & Canetti, R. (Februray 1997). *RFC 2014 – HMAC: Keyed-Hashing for Message Authentication.* Preuzeto 9. 5. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc2104

[148] Paar, C., & Pelzl, J. (2009). *Understanding cryptography*, n. dj.

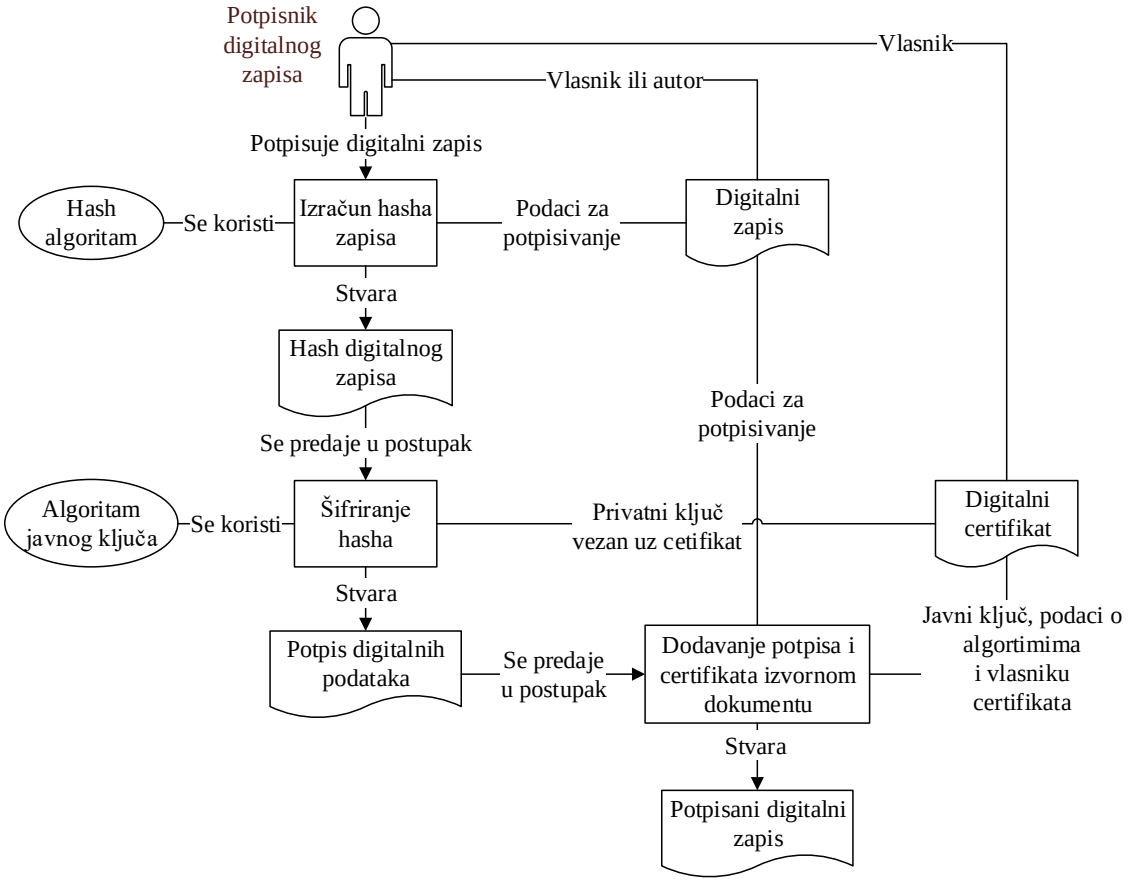
[149] Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1996). *Handbook of Applied Cryptography.* CRC Press. Preuzeto 20. 12. 2021. s http://labit501.upct.es/~fburrull/docencia/SeguridadEnRedes/old/teoria/bibliography/HandbookOfAppliedCryptography_AMenezes.pdf

Osnova ovakvog sustava su dva ključa, privatni ključ koji se koristi u procesu stvaranja digitalnog potpisa i javni ključ koji se koristi u procesu potvrde ispravnosti digitalnog potpisa te kriptografski algoritmi koji reguliraju oba procesa.

U sustavu infrastrukture javnog ključa proces potpisivanja digitalnog zapisa prati sljedeće korake:

1) Potpisnik računa hash zapisa koji je potrebno potpisati.
2) Potpisnik šifrira izračunati hash zapisa upotrebom svog privatnog (tajnog) ključa.
3) Potpisnik zapisu dodaje šifrirani hash zapisa i svoj digitalni certifikat (koji sadrži javni ključ).

Ovaj proces karakterističan je za sve PKI sustave i detaljnije je opisan na slici 2. Proces je identičan bez obzira potpisuje li dokument fizička ili pravna osobe ili podatke potpisuje automatizirani sustav.

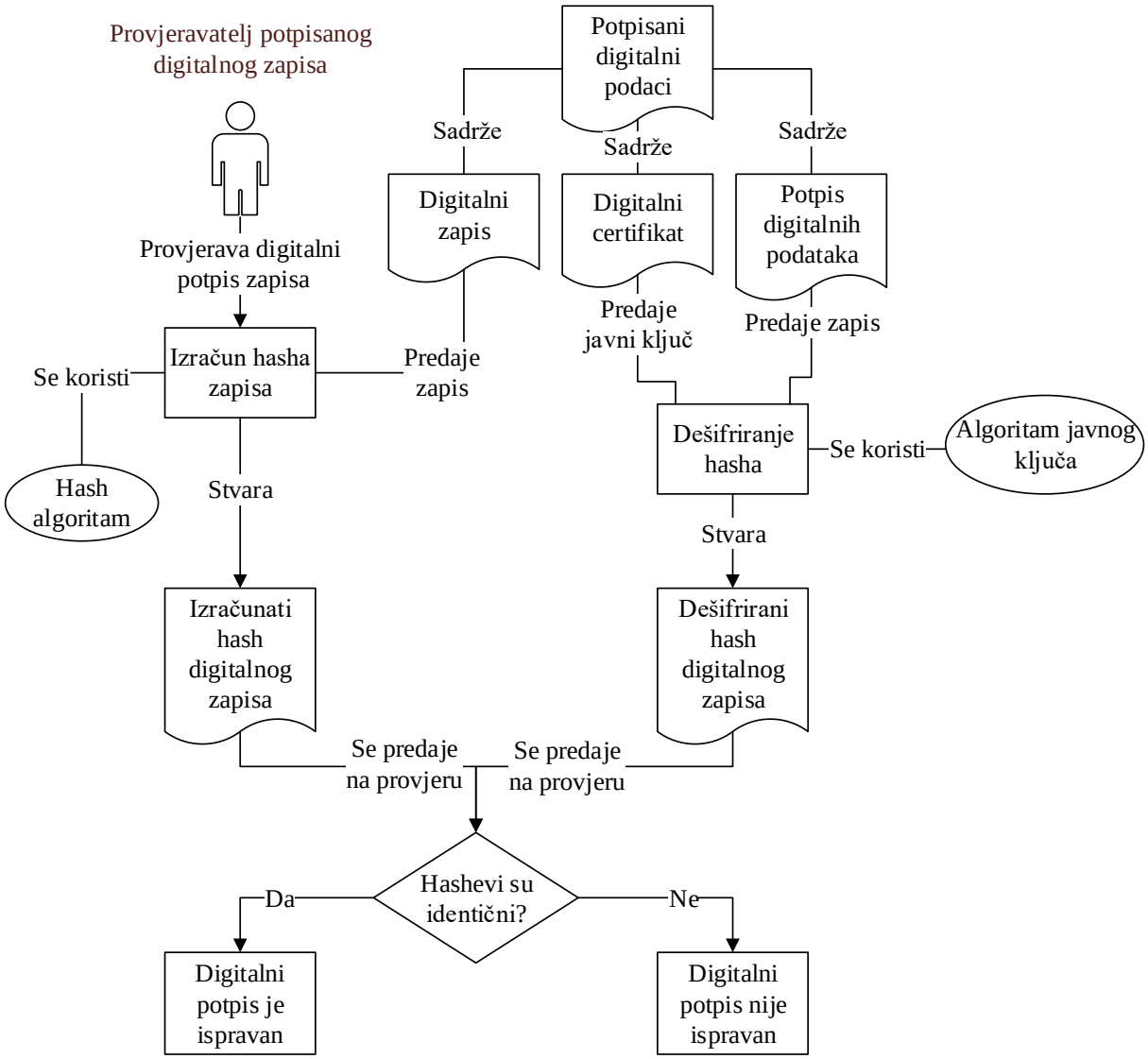


Slika 2. Postupak potpisivanja dokumenta digitalnim certifikatom

Pri provjeri ovakvo potpisanog zapisa potrebno je pratiti sljedeću proceduru:

1) Primatelj računa hash zapisa čiju je autentičnost potrebno provjeriti.
2) Primatelj dešifrira hash zapisa koji dodan poruci koristeći javni ključ koji je dio digitalnog certifikata (i pridružen je zapisu).
3) Primatelj uspoređuje dešifrirani i upravo izračunati hash. Ako su hashevi identični smatra se da je zapis potpisao vlasnik digitalnog certifikata.

Ova procedura opisana je i slikom 3.



Slika 3. Postupak provjere ispravnosti digitalnog potpisa

Iz gore prikazanih procesa jasno je da su ranije navedeni procesi: šifriranje, dešifriranje i izračuna hasha centralni elementi procesa digitalnog potpisivanja i dokazivanja ispravnosti

potpisa. Sva tri pojma smatraju se osnovnim pojmovima u kriptografiji. Pod pojmom šifriranja podrazumijeva se postupak koji mijenja podatke na način da oni više nisu razumljivi bez posebnog znanja, a postupak dešifriranja je povratak podataka iz šifriranog u izvorno stanje.

PKI sustavi posredstvom certifikacijskog autoriteta (engl. *Certificate Authority* ili skraćeno CA) povezuju javni ključ s fizičkom ili pravnom osobom i na taj način omogućuju upotrebu digitalnog potpisa za dokazivanje autentičnosti i integriteta potpisanog zapisa.

U ostatku poglavlja objašnjavaju se tehnologije potrebne da se ovaj, ukratko opisani, postupak detaljno razradi te da se nakon njegovog razumijevanja može sa sigurnošću tvrditi da on može biti iskorišten u svrhu dokazivanja autentičnosti i pouzdanosti dokumenata zapisa.

3.1.1. Hash algoritmi

Hash algoritmi, to jest hash funkcija je ključan pojam u polju kriptografije i jedna od pretpostavki digitalnog potpisa te je iz tog razloga ovo istraživanje mora razmotriti i objasniti. Nadalje, hash funkcije su od izuzetne važnosti za ovo istraživanje jer nisu ključne samo za digitalni potpis već su i pretpostavka za podatkovnu strukturu poznatu kao ulančani blokovi (engl. *blockchain*).

S obzirom na važnost hash funkcija za dva centralna pojma ovog istraživanja (digitalni potpis i ulančani blok) i na to da je istraživanje teorijsko, hash funkcije koje smatram primjenjivim na novi model, i koje se koriste u digitalnom potpisu, će biti raspravljene do razine konkretnih algoritama.

Engleska riječ *hash* prema Cambridge Online rječniku[150] je imenica koja označava:

1. Jelo koje je pripremljeno od mješavine mesa, krumpira i povrća koji su narezani te kuhani ili pečeni.
2. Neformalni naziv za opojno sredstvo.
3. Simbol # na tipkovnici računala ili telefona.
4. Neuspjeli pokušaj.

Ovaj rječnik, uopće ne spominje kriptografsko značenje riječi koje je u smislu izrade najsličnije definiciji pod brojem 1. Hash funkcija rezultira ispremještanim, šifriranim, podacima

[150] URL: https://dictionary.cambridge.org/dictionary/english/hash

koji više uopće ne sliče svom izvornom obliku, slično kao i jelo koje je opisano u rječniku Sveučilišta Cambridge. U kontekstu informacijskih tehnologija, u hrvatskom jeziku "hash" se često prevodi kao "sažetak". Vjerojatno iz razloga što je hash često (ali ne uvijek) kraći od izvornika i što se važan dio hash algoritma zove engl. *compression function*, što se prevodi kao "funkcija za sažimanje". Hrvatski školski rječnik[151] koji je izdao Institut za hrvatski jezik i jezikoslovlje[152] definira riječ sažetak kao:

> *"*im. m. *⟨G sažétka;* mn. *N sažétci, G sàžētākā⟩ tekst u kojemu je na kratak način izneseno sve što je bitno u sadržaju nekoga većeg teksta ili djela [~ predavanja]"*

Osobno smatram da upotreba riječi sažetak nije odgovarajući prijevod riječi hash u kontekstu informacijske tehnologije, to jest kriptografije (a i šire) iz dva razloga:

1. Rezultat hash funkcije uopće ne mora biti kraći od izvornika. Moguće je hashirati bilo kakvu poruku, čak i samo jedan znak, a rezultat hash funkcije mora biti točno određene duljine. Ovo znači da će u slučaju hashiranja jednog znaka (i drugih kratkih poruka), "sažetak" biti duži od izvornika!
2. Rezultat hash funkcije ne smije prenositi informacije o izvorniku. Ovo se odnosi na zahtjev da hash funkcije budu jednosmjerne. Trebalo bi biti nemoguće pročitati ovakav "sažetak" i iz njega izvući bilo koju informaciju koju se može povezati s izvornikom. Ovo je u izravnoj suprotnosti s navedenom definicijom riječi sažetak u hrvatskom jeziku koja implicira da je sačuvano i preneseno "sve što je bitno u sadržaju".

Kao dokaz gornjih teza u nastavku je prikazan rezultat hashiranja kratice "FFZG", upotrebom SHA-512 algoritma:

```
6bc238e669953ae1ce98b7ef5f5b1ac000562cf686aabd737488c252005b327
3650126611904819dde0f8eca45d647c0ee37a96d0f5994679a615a787544a4
97
```

i CRC32 algoritma, koji rezultira jednim od najkraćih hasheva:

```
af16930a
```

[151] URL: http://rjecnik.hr/search/?q=sa%C5%BEetak
[152] URL: http://ihjj.hr/

Jasno je da je u oba slučaj sažetak ne samo duži od izvornika već nikako ne "iznosi sve što je bitno u sadržaju" izvorne poruke - upravo suprotno. Iz ovih razloga u ovoj disertaciji se ne koristi hrvatski, iz navedenih razloga neodgovarajući, prijevod riječi hash već se ona koristi u izvornom obliku.

Engleski termin *hash function*, osim što se ponekad (potpuno krivo) prevodi kao funkcija za sažimanje, u nekim slučajevima je preveden i kao funkcija raspršivanja.[153] Ovo je puno prikladniji prijevod od funkcije za sažimanje koji djelomično odgovara postupku koji ove funkcije provode ali ne pomaže u slučaju samog *hasha*, rezultata hash funkcije. Stoga ostajem pri upotrebi strane riječi.

Iako se kriptografija s latinskog prevodi doslovno kao "skriveno pisanje",[154] i obično se doživljava kao znanost koja se bavi skrivanjem podataka, preciznije je reći da je to znanost kontrolirane promjene podataka. Hash funkcija je primjer ovakve upotrebe kriptografije, ona ne skriva podatke već ih po određenim pravilima, jednosmjerno i nepovratno, mijenja.

Ova promjena podataka, u slučaju hash funkcije rezultira novim podacima koji su točno određene veličine. Dakle, hash funkcija se može definirati kao "funkcija koja prima podatke proizvoljne dužine, a vraća podatke određene dužine" [155] ili kao funkcija koja reducira neograničenu domenu na ograničenu kodomenu. U kontekstu hash funkcija ulazni podaci obično se zovu poruka (engl. *message*), a izlazni hash poruke, to jest sažetak poruke. U engleskom jeziku kao sinonim za hash poruke često se koristi engl. *message digest* – ovo je još jedan termin, to jest kolokacija bez dobrog prijevoda na hrvatski jezik.

Jednostavan primjer hash funkcije je osnovni oblik funkcija koje se koriste pri postupku direktnog adresiranja u računarstvu. Direktno adresiranje[156] je postupak u kojem se izravno iz podataka, npr. iz OIB-a osobe, računa adresa pretinca polja u koje će biti pohranjen podatak. Ako znamo da je podatkovna struktura zvana polje neprekinut memorijski prostor[157] podijeljen na *n* pretinca jednake dužine, da možemo u algoritamskom vremenu konstantne složenosti

---

[153] Npr. Miroslav Kiš, Informatički rječnik. Englesko-hrvatski / hrvatsko-engleski, Naklada Ljevak, Zagreb, 2000., str. 446-447.
[154] Na modernom latinskom *"cryptographia"* ili na tradicionalnom grčkom *"kryptos"* i *"graphia"* prema: https://www.etymonline.com/word/cryptography
[155] Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1996). *Handbook of Applied Cryptography*, n. dj.
[156] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms.* Massachusetts Institute of Technology. Preuzeto 7. 12. 2021. s http://139.59.56.236/bitstream/123456789/106/1/Introduction%20to%20Algorithms%20by%20Thomas%20%20H%20Coremen.pdf
[157] IST. (2021). *array*. Preuzeto 20. 12. 2021. s Dictionary of Algorithms and Data Structures: https://xlinux.nist.gov/dads/HTML/array.html

izračunati fizički položaj bilo kojeg pretinca (na primjer, ako je svaki pretinac velik 2 bajta onda se *n-ti* pretinac nalazi *n* * 2 bajta iza početka polja) i da građana Republike Hrvatske ima 4 milijuna, a OIB ima 11 znamenki[158] suočavamo se s pitanjem koliko mora biti veliko polje za pohranu OIB-a svih građana: Da bi ovo polje bilo racionalno treba domenu od 11 znamenki, to jest prostor od 10^{11} memorijskih pretinaca svesti na racionalnu veličinu, na primjer, $6 * 10^6$ je već 50% više pretinaca nego što je potrebno za pohranu podataka 4 milijuna građana Republike Hrvatske. Hash funkciju koja će omogućiti racionalno adresiranje moguće je jednostavno izvesti upotrebnom matematičke operacije *modulo*. Modulo je operacija koja daje ostatak cjelobrojnog dijeljenja dva broja. Vrijedi da je:

$$x : y = r + o$$

Gdje je *r* rezultat cjelobrojnog dijeljenja (količnik) *x* i *y*, a *o* ostatak iste operacije, onda za operaciju modulo (operator modulo se označavamo s riječi mod), vrijedi:

$$x \bmod y = o$$

Ili preciznije:

$$x \bmod y = x - (x : y) * y$$

Jasno je da ostatak cjelobrojnog dijeljenja, to jest rezultat operacije modulo nikad ne može biti veći od djelitelja, to jest nazivnika. Ovo uopće ne ovisi djeljeniku (brojniku) te on može biti bilo koji broj. Na ovaj način izračunom modula, to jest ostatka cjelobrojnog dijeljenja gdje je djeljenik OIB građana (lišen vodećih nuli), a djelitelj broj pretinaca (na primjer, $6 * 10^6$) dobivamo adresu pretinca koju možemo koristiti pri postupku direktnog adresiranja.[159] Ovaj postupak zorno prikazuje svrhu hash funkcija. Funkcija:

$$f(x) = x \bmod n$$

neograničenu domenu (bilo koji broj) funkcije svodi na točno omeđenu kodomenu:

$$f(x) = \{ y \mid y \in N, y < n \}$$

te je kao takvu možemo je smatrati karakterističnim primjerom hash funkcije. Hash algoritmi koji se koriste u kriptografiji su značajno složeniji i kasnije će biti detaljno razmotren primjer porodice algoritama koja se najčešće koristi u digitalnim potpisima te je izvrstan kandidat za

[158] OIB je jedinstveni identifikacijski broj od 11 znamenki prema: https://www.porezna-uprava.hr/HR_OIB/Stranice/default.aspx
[159] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms*, n. dj.

hash funkciju kojom se realizira struktura povezanih blokova koja čini temelj TrustChain modela.

U prošlom poglavlju navedeno je da postupak digitalnog potpisivanja počinje računanjem hash-a izvornih podataka – poruke. Razlog ovome je učinkovitost i zahtjev shema za digitalni potpis da kao ulaz prime podatke jednake dužine. Podaci koji se digitalno potpisuju mogu biti i izravno predani algoritmu javnog ključa ali to će rezultirati s dužim vremenom šifriranja te će, u slučaju većine shema za digitalni potpis, biti nužno podatke podijeliti u manje blokove, što predstavlja sigurnosti rizik.[160] S obzirom na to da je hash poruke, premda definiciji, uvijek točno određene dužine, postupak računanja hash poruke prije enkripcije i pohrane u digitalni potpis omogućava da se potpis brzo generira te da bude predvidljive dužine. Predvidljiva dužina je korisna osobina u svim primjenama digitalnog potpisa ali je i nužna u slučaju kada je digitalni potpis dio mrežnog protokola poput TLS i SSL protokola.

Iako digitalni potpis može koristiti bilo koji hash algoritam u ovom radu je detaljnije opisana SHA porodica algoritam. Algoritmi uz ove porodice (SHA-1,SHA-2 i SHA-3) su, uz MD i CRC porodice, jedni od najkorištenijih algoritama za izračun hasha čija primjena obuhvaća ali i prelazi digitalni potpis. Na primjer, eIDAS zahtjevi za interoperabilnošću sustava koji koriste TLS zahtijevaju upotrebu algoritama iz SHA porodice.[161]

3.1.1.1. Ranjivosti hash algoritama

Iako se oslanjanju na činjenicu da samo mala izmjena u izvornoj poruci rezultira velikim izmjenama u rezultatu, to jest u izračunatom hashu, svi hash algoritmi podložni su napadima u kojima se u razumnom vremenskom razdoblju pokušava naći druga izvorna poruka koja rezultira istim hashem. Na primjer, niz znakova:

```
Filozofski fakultet Sveučilišta u Zagrebu
```

Upotrebom SHA-1 algoritma pretvara se u niz:

```
67dfa71e2e9b3eadfffe99b44fd78c7c621f446c
```

[160] Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1996). *Handbook of Applied Cryptography*, n. dj.
[161] eIDAS eID Technical Subgroup. (2019). eIDAS Cryptographic Requirements for the Interoperability Framework. Preuzeto 20. 12. 2021. s https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiP3uPQl_L0AhV5S_EDHeP1BS4QFnoECAwQAQ&url=https%3A%2F%2Fec.europa.eu%2Fcefdigital%2Fwiki%2Fdownload%2Fattachments%2F82773108%2FeIDAS%2520Cryptographic%2520Requirement%

Ali niz u kojem je izmijenjen samo jedan znak, veliko "Z" pretvoreno je u malo "z":

```
Filozofski fakultet Sveučilišta u zagrebu
```

Rezultira s potpuno drukčijim nizom:

```
cff953cb713f29a968c3b65c997a6497acfad478
```

Ovu osobinu hash algoritama kriptografija naziva efekt lavine (engl. *avalanche effect*) te je ona karakteristična za sve hash funkcije i sustave šifriranja blokova.[162] Termin efekt lavine u ovom kontekstu je 1973. godine prvi upotrijebio Horst Feistel.[163] Od tada do danas termin i definicija lavine u kriptografiji su dalje razvijani. Webster i Tavares definirali su 1985. godine "striktni kriterij lavine" kao zahtjev da ako se samo jedan bit ulaznih podataka kriptografske funkcije promijeni svaki od izlaznih bitova ima 50% šanse da se i on promijeni.[164]

Na ovoj promjenjivosti rezultata, to jest efektu lavine, temelji se sigurnost hash algoritma. Ideja je da ako hashiramo, na primjer neki kupoprodajni ugovor, tekstualni dokument od 1000 riječi, i izmijenimo samo jednu riječ, na primjer prezime kupca, dobiti ćemo drastično drukčiji rezultat. I ovo je uglavnom istina, pa čak i ako je moguće naći dva izvorna zapisa koji rezultiraju istim hashem oni će, vrlo vjerojatno, biti toliko različiti da će iz konteksta biti očigledno koji od njih je falsifikat. Ipak, ovo može izazvati razne pravne poteškoće, a ne pomaže ni u situacijama kada je izvorna poruka jako kratka. Na primjer, možda se nije hashirao cijeli kupoprodajni ugovor već samo iznos tijekom kasnije bankovne transakcije.

[162] Kriptografski algoritam koji radi na nizu blokova podataka jednake dužine, najčešće je riječ o algoritmima s simetričnim ključem. Na primjer, DES (engl. Data Encryption Standard) i AES (engl. Advanced Encryption Standard).
[163] Feistel, H. (1973). Cryptography and computer privacy. *Scientific american, 228*(5), 15-23. Preuzeto 15. 5. 2020. s http://www.apprendre-en-ligne.net/crypto/bibliotheque/feistel/
[164] Webster, A. F., & Tavares, S. E. (1985). On the design of S-boxes. *Conference on the theory and application of cryptographic techniques* (str. 523-534). Berlin, Heidelberg: Springer. Preuzeto 15. 5. 2020. s https://link.springer.com/content/pdf/10.1007/3-540-39799-X_41.pdf

Kada govorimo o napadima na hash algoritme u kriptografiji razlikujemo dva različita pojma[165]:

- napade kolizijom, engl. *collision attacks*.
- napade preslikom, engl. *preimage attacks* i drugom preslikom, engl. *second preimage attack.*

Napadi kolizijom podrazumijevaju da postoje dvije ili više različitih izvornih poruka koje se hashiraju u isti rezultat. Ako je m poruka, a f hash funkcija, moguće (a često i sigurno) je da postoji takav m' da vrijedi:

$$f(m) = f(m')$$

Napad kolizijom svodi se na pronalaženje ovakvog para. Prateći raniji primjer jednostavne hash funkcije za direktno adresiranje sasvim je jasno da su u takvom sustavu kolizije česte. U slučaju hash funkcija koje se koriste u kriptografske svrhe situacija je nešto povoljnija ali ova prijetnja i dalje postoji.

Ovaj napad je izuzetno nezgodan za digitalne potpise jer jednom kada je napadač našao prikladan par m i m', upotrebom digitalnog potpisa više nije moguće dokazati koja od ove dvije poruke je potpisana. Na primjer, zamislimo situaciju u kojoj osoba A i B potpisuju i digitalnim potpisom ovjeravaju ugovor. Zlonamjerna osoba, osoba B, prije potpisivanja napiše legitimni ugovor, m, i drugu varijantu ugovora, n, koju osoba A nikad ne bi potpisala. Nakon izrade ova dva ugovora, osoba B istražuje moguće kolizije na način da prijedlog legitimnog ugovora mijenja tako da dodaje, miče i pomiče interpunkcijske znakove, mijenja riječi u sinonime, mijenja raspored riječi bez izmjene sadržaja i slično, dakle mijenja ugovor ali ne na način da on osobi A postane neprikladan za potpisivanje. Ove promjene se rade dok se ne nađe varijanta dokumenta m, m' koja rezultira istim hashem kao i neprikladni dokument n. Naravno, tijekom ove potrage za kolizijom moguće je mijenjati (bez značajnih promjena u značenju) i dokument n. Od ovog napada se brani na način da se koristi hash koji je dovoljno dug da s postojećom tehnologijom nije moguće naći par m i m' u razumnom vremenskom periodu. Ovakvi napadi su iz perspektive arhivistike manje značajni jer se moraju odigrati u trenutku stvaranja dokumenta. Na ovaj način nije moguće naknadno napasti već potpisane i arhivirane dokumente,

[165] Maetouq, A., Daud, S., Ahmad, N., Maarop, N., Sjarif, N., & Abas, H. (2018). Comparison of hash function algorithms against attacks: A review. *International Journal of Advanced Computer Science and Applications*, *9*(8). Preuzeto 21. 12. 2021. s https://pdfs.semanticscholar.org/6ed2/50d11a5c80f550bd8efcc673606c3cae34b7.pdf

to jest dokument i njegov potpis su morali biti komprimirani već u trenutku stvaranja te kao takvi eventualno ući u arhiv. U takvom slučaju arhiv ne može biti odgovoran za takav napad, jer to i nije uspjeli napad na arhiv, napad se dogodio prije arhiviranja dokumenta. Arhiv je, sukladno ranije definiranim terminima, primio autentičan ali nepouzdan dokument.

Napad kolizijom dolazi i u drugoj varijanti, napad određenim prefiksom, engl. *chosen-prefix collision attack*. Napad podrazumijeva sve već navedeno, dakle da postoji m i m' za koje vrijedi $f(m) = f(m')$, ali olakšava napad na način da se poruci dodaju prefiksi n i n' koji stvaraju situaciju u kojoj vrijedi:

$$f(n||m) = f(n'||m')$$

Gdje f označava hash funkciju, a || operaciju spajanja (engl. *concatenation*). U ovom slučaju poruke m i m' uopće ne moraju biti slične, i s obzirom na to da se n i n' naknadno dodaju puno ih je jednostavnije izračunati. Ovaj napad specifičan je za Merkle–Damgård[166,167] hash algoritme koji uključuju SHA-1 i SHA-2 algoritme.

Napadi preslikom je vrlo sličan koncept napadu kolizijom, ali podrazumijeva da već imamo hash i pokušavamo naći neki izvorni dokument koji mu odgovara. Ovaj postupak bi trebao biti nemoguć pri upotrebi dobro napisane hash funkcije i ključa odgovarajuće dužine. S obzirom na to da se oba napada svode na pronalaženje dokumenta koji rezultiraju specifičnim hashem, otpornost na napad preslikom implicira i otpornost na napad kolizijom. Otpornost na napad preslikom se dijeli na dvije, specifične, otpornosti:[168]

- Otpornost na napad preslikom, engl. *preimage attack resistance*, podrazumijeva da je za konkretni hash *h* nemoguće naći poruku *m* za koju vrijedi *f(m) = h* gdje je *f* hash funkcija.
- Druga otpornost na napad preslikom, engl. *second preimage resistance*, podrazumijeva da je nemoguće naći takve izvorne poruke m i m', da ako je f hash funkcija vrijedi: $f(m) = f(m')$

[166] Merkle, R. C. (1979). *Secrecy, authentication, and public key systems.* Stanford University. Preuzeto 5. 5. 2020. s http://www.merkle.com/papers/Thesis1979.pdf
[167] Damgård, I. B. (1989). A design principle for hash functions. *Conference on the Theory and Application of Cryptology* (str. 416-427). New York, NY: Springer. Preuzeto 5. 5. 2020. s https://link.springer.com/content/pdf/10.1007/0-387-34805-0_39.pdf
[168] Maetouq, A., Daud, S., Ahmad, N., Maarop, N., Sjarif, N., & Abas, H. (2018). Comparison of hash function algorithms, n. dj.

Važno je napomenuti da svi navedeni napadi predstavljaju značajno manji rizik za ulančane zapise kakve koristi TrustChain od rizika koji predstavljaju za individualne dokumente. TrustChain blokovi su striktno definirani i sadrže niz hasheva. U sklopu ovog istraživanja to nije dokazano ali s razumnom sigurnošću možemo pretpostaviti da će bilo kakve izmjene u striktno određenoj podatkovnoj strukturi koje bi mogle rezultirati istim hashem stvoriti situaciju u kojoj je iz konteksta jasno koji blok je izvoran, a koji je falsifikat. Ovo je posljedica činjenice da će napadač morati izmijeniti podatke više zapisa da bi došao do željenog rezultata (ako je on uopće moguć), a napad određenim prefiksom uopće nije moguć. Ovakvi napadi su, s druge strane, vrlo mogući u slučaju manje striktnih oblika zapisa, poput PDF datoteka.

3.1.1.2. SHA-1

SHA-1 (engl. *Secure Hash Algorithm*) se smatra prvim predstavnikom SHA algoritma te čini osnovu cijelog niza algoritama koji spadaju u ovu porodicu. SHA-1 algoritmu prethodio je SHA algoritam, retroaktivno nazvan i SHA-0, ali zbog sigurnosnih propusta koji su uočeni tijekom razvoja taj algoritam nikad nije zaživio te se SHA-1, koji nije ništa više od revizije SHA-0 koja izbjegava uočeni propust,[169] smatra *de facto* prvim algoritmom u ovoj porodici algoritama. Iako se na nekim mjestima još uvijek može naći u upotrebi, SHA-1 je, zbog sigurnosnih propusta, uglavnom povučen i zamijenjen SHA-2 algoritmom. National Institute of Standards and Technology je definirao SHA-1 standard u izvornom Secure Hash Standard dokumentu, FIPS PUB 180-1, iz 1995.[170] Struktura SHA-2 algoritma, koji se i dalje koristi te je relevantan kandidat za upotrebu i u modelima koji su tema ovog istraživanja, se temelji na SHA-1 te je iz tog razloga važno razumjeti funkcioniranje izvorne funkcije. SHA-2 definiran je u FIPS PUB 180-2 dokumentu.[171]

[169] Wang, X., Yu, H., & Yin, Y. (2005). Efficient collision search attacks on SHA-0. *Annual International Cryptology Conference* (str. 1-16). Berlin: Springer. Preuzeto 3. 1. 2022. s https://link.springer.com/content/pdf/10.1007/11535218_1.pdf

[170] National Institute of Standards and Technology. (1995). *NIST FIPS 180-1 Secure Hash Standard.* doi: https://doi.org/10.6028/NIST.FIPS.180-1

[171] National Institute of Standards and Technology. (2002). *NIST FIPS 180-2 Secure Hash Standard (SHS).* Preuzeto 15. 5. 2020. s National Institute of Standards and Technology: https://csrc.nist.gov/publications/detail/fips/180/2/archive/2002-08-01

SHA-1 funkcija rezultira hashem od 20 bajta, to jest 160 bita te se njeno funkcioniranje može podijeliti u dvije faze, pretprocesiranje i izračun hasha, to jest poziv funkcije za sažimanje. Nadalje, algoritam se može detaljnije raščlaniti na tri faze:

1. Faza pretprocesiranja. U ovoj fazi se izvorna poruka priprema za obradu. Ovo znači da se poruka proširuje do 512 bita ili se dijeli na više 512-bitnih dijelova.
2. Faza obrade 512-bitnih dijelova poruke. Ova faza se ponavlja za svaki dio koji je rezultat 1. faze i sastoji se od 4 koraka:
    a. Pripremi 16 32-bitne riječi.
    b. Na osnovu tih 16 riječi stvori još 64 (ukupno 80 riječi).
    c. Za svaku od 80 riječi odradi obradu prema zadanim funkcijama.
    d. Dodaj rezultat obrade rezultat prethodnog dijela.
3. Faza konačne prilagodbe rezultata iz prethodne faze (izračun konačnog hasha).

U daljnjem tekstu ovaj postupak se detaljno opisuje.

Pretprocesiranje, to jest priprema za izračun hasha ima zadaću pripremiti poruku na način da je poveća do veličine koja je višekratnik broja 512 u bitovima.[172] Dakle, ako veličinu podataka označavamo s *d*, onda ona mora zadovoljavati sljedeći uvjet: $d \; mod \; 512 = 0$. Ovaj postupak zovemo dopunjavanje podataka (engl. *padding*). Ako imamo poruku *d*, dužine *l* bitova, onda je postupak pretprocesiranja opisan idućim algoritmom:

1. Na kraj poruke *d* se dodaje bit *o* koji je jednak "1".

$$d_2 = d + o$$

2. Računa se broj *k* na način:

$$k = 512 - 64 - 1 - l$$

3. Na kraj poruke d_2 se dodaje niz *ko* koji se sastoji od *k* "0" bitova.

$$d_3 = d_2 + ko$$

[172] Paar, C., & Pelzl, J. (2009). *Understanding cryptography*, n. dj.

4. Na kraj poruke d_3 se dodaje broj l u 64 bitnom binarnom zapisu.

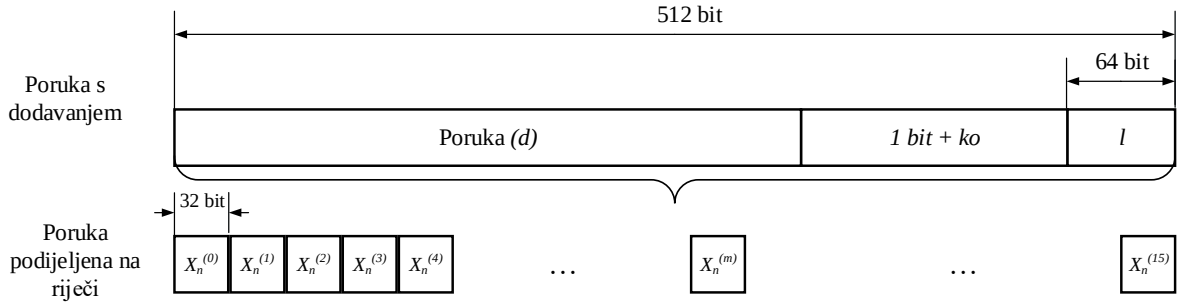
$$d_4 = d_3 + l$$

5. Zapis d_4 (koji je dužine 512*n bitova) se dijeli u n blokova $X_1, X_2, X_{2\ldots}, X_n$ (koji su dužine 512 bitova).

6. Svaki od X_n blokova se dijeli na 16 riječi veličine 32-bita, tako da je

$$X_n = X_n^{(0)}, X_n^{(1)}, X_n^{(2)}, \ldots, X_n^{(15)}$$

Rezultat ovog postupka prikazan je na slici 4. Ako je izvorna poruka d duža od 512 bita formira se n ovakvih bloka te samo zadnji prolazi kroz postupak paddinga.



Slika 4. SHA-1 priprema podataka za kompresiju

Ako bi ovaj postupak htjeli primijeniti na poruku "FFZG" prvo moramo slova prikazati u njihovim 8 bitnim binarnim ASCII vrijednostima. Za poruku "FFZG" to su: F=01000110, Z=01011010 i G=01000111 te je potpuni binarni niz prikazan na slici 5.

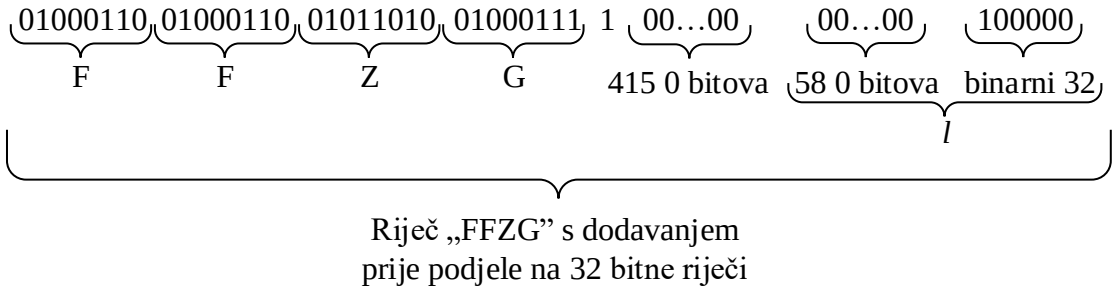
$$\underbrace{01000110}_{F}\,\underbrace{01000110}_{F}\,\underbrace{01011010}_{Z}\,\underbrace{01000111}_{G}$$

Slika 5. Riječ "FFZG" u 8 bitnim binarnim ASCII vrijednostima

Da bi mogli odraditi pretprocesiranje potrebno je izračunati k sukladno koraku broj 2 iz prethodnog algoritma. Ako je dužina naše poruke $l=32$ (jer imamo 4 slova po 8 bitova) onda je k:

$$k = 512 - 64 - 1 - 32 = 415$$

Nakon izračuna *k* poruka može proći postupak dopunjavanja. Potpuna struktura prikazana je na slici 6.



Slika 6. Riječ "FFZG" u konačnoj fazi SHA-1 pretprocesiranja

Nakon uspješnog pretprocesiranja, algoritam prelazi u fazu izračuna hasha, to jest sažimanja. Ovaj algoritam se odvija u *n* faza, gdje je *n* broj blokova od 512 bitova koji su bili rezultat pretprocesiranja. Svaki od ovih blokova sažima se individualno u 80 iteracija, to jest u 4 kruga od 20 koraka.

U prvom koraku stvara se spremnik (engl. *buffer)* u koji će se pohraniti izračunata 160 bitna hash vrijednost. Ovaj *spremnik* se sastoji od 5 32-bitnih pretinaca koji pohranjuju 5 riječi (*A, B, C, D, E*) te u početnom stanju poprimaju zadane vrijednosti koje u heksadecimalnom obliku glase:

```
A₀  = 67452301
B₀  = EFCDAB89
C₀  = 98BADCFE
D₀  = 10325476
E₀  = C3D2E1F0
```

U idućem koraku izvodi se 80 riječi, W_i, iz trenutnog bloka – X_t. U prvom koraku ove riječi se tvore od 16 32-bitnih riječi stvorenih u fazi pretprocesiranja $X_{0\text{-}15}$. Ove riječi se izvode na slijedeći način:

$$W_i = \begin{cases} X_n^i & za\ 0 \ \leq\ i\ \leq 15 \\ bs_1(W_{i-3} \oplus W_{i-8} \oplus W_{i-14} \oplus W_{i-16}) & za\ 16 \leq\ i\ \leq 79 \end{cases}$$

Gornja operacija rezultirat će s 80 riječi W_i koje su spremne za daljnju obradu. Funkcija bs_1 označava operaciju bitnog cirkularnog lijevog pomaka za 1 mjesto (engl. *circular left bit shift)*. Ova operacija, koja je u programiranju često označena s << *n* (lijevi pomak) ili >> *n* (desni pomak), u slučaju lijevog cirkularnog pomaka podrazumijeva odbacivanje n bitova

počevši s lijeve strane (najvažniji bitovi) i njihovo dodavanje na desnu stranu (najmanje važni bitovi)[173]. Desni pomak funkcionira obrnuto. Na primjer:

S = 1001 0110

S << 1 = 0010 1101

S << 5 = 1101 0010

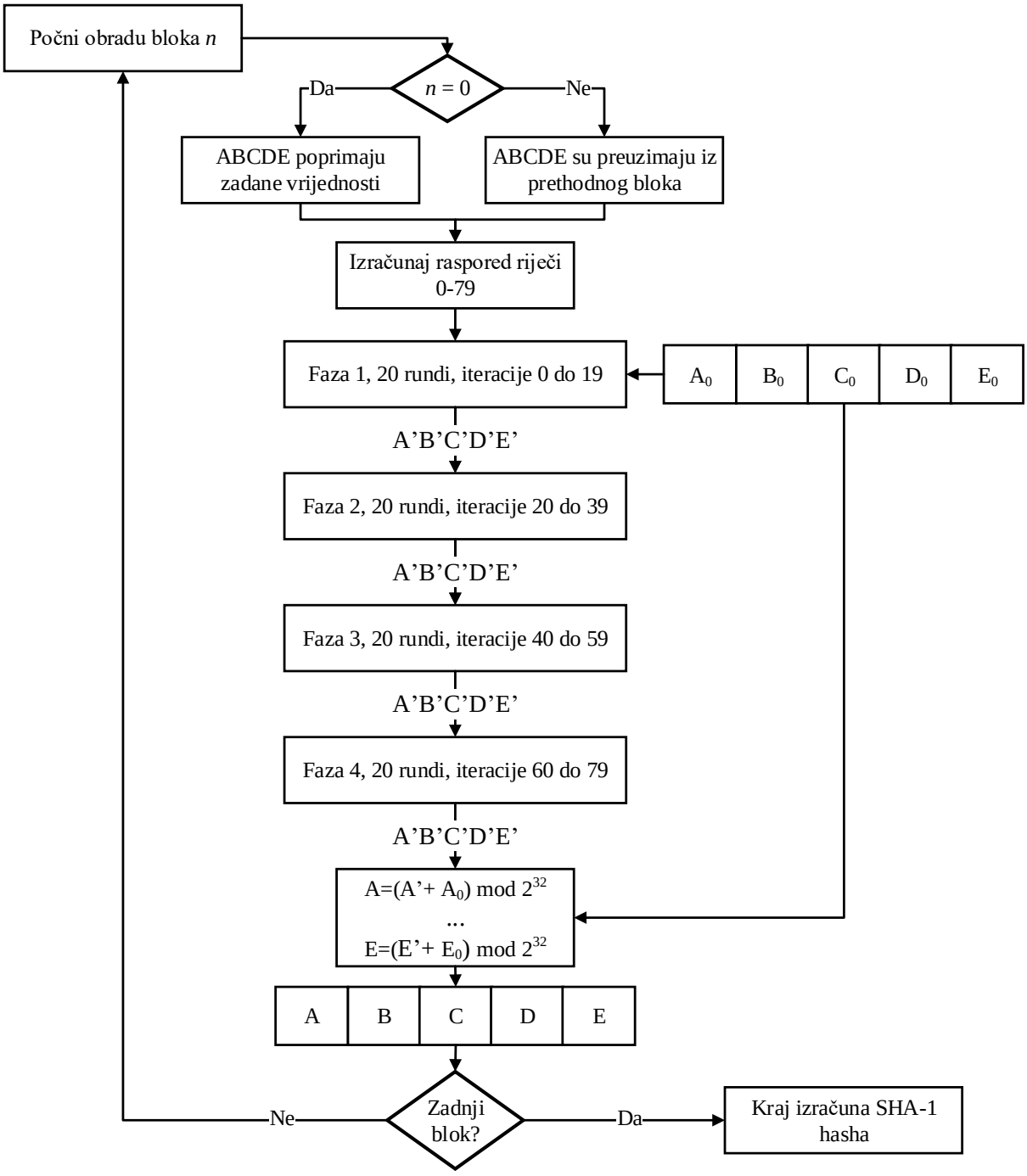
Simbol ⊕ označava XOR digitalna logička vrata, to jest operaciju ekskluzivne disjunkcije ili isključivog "ili" (*engl. eXclusive OR)*. XOR funkcija prima dva binarna ulaza te vraća 1 kada je samo jedan od dva ulaza jednak 1. Funkcioniranje ove operacije opisano je u tablici 3.

Tablica 3. XOR operacija

X	Y	X ⊕ Y
0	0	0
1	0	1
0	1	1
1	1	0

Nakon pripreme ovih 80 riječi (engl. *word schedule*) počinje izvršavanje 80 iteracija, *i,* za blok X_t. Ovaj postupak prati dijagram toka prikazan na slici 7.

[173] Yordzhev, K. Y. (2009). An example for the use of bitwise operations in programming. *Proceedings of the Thirty Eighth Spring Conference of the Union of Bulgarian Mathematicians*, (str. 196-202). Borovetz. Preuzeto 21. 12. 2021. s https://www.researchgate.net/publication/51978936_An_Example_for_the_Use_of_Bitwise_Operations_in_Programming



Slika 7. Struktura izvođenja SHA-1 algoritma

Slika 7 dijeli iteracije u 4 faze, od kojih svaka ima 20 rundi, jer se na podatke primjenjuju različite funkcije, ovisno o trenutnoj iteraciji. Ove funkcije opisane su niže, nakon opisa dijela iteracija koji je isti u svim slučajevima.

Ako postojeći spremnik (engl. *buffer*) riječi označimo s A,B,C,D,E, a novi s A',B',C',D',E' onda svaka iteracija uzima postojeći spremnik i nad njim provodi:

$$A' = (E + f_i(B, C, D) + bs_5(A) + W_i + K_i)$$

$$B' = A$$

$$C' = bs_{30}(B)$$

$$D' = C$$

$$E' = D$$

U gornjim formulama W_i označava trenutnu izvedenu riječ, a bs_5 i bs_{30} označavaju operaciju bitnog kružnog lijevog pomaka (engl. *circular left bit shift*) za 5 i 30 položaja.

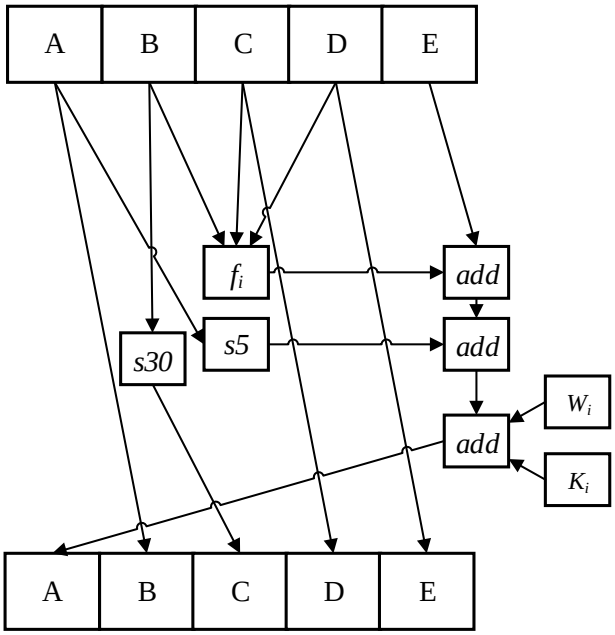
Funkcija f_i i konstanta K_i su varijabilne i ovise o rundi, to jest o iteraciji koja se trenutno izvodi. Moguće funkcije i konstante prikazane su u tablici 4.

Tablica 4. Funkcije i konstante SHA-1 algoritma

Iteracija i	Funkcija f_i	Konstanta K_i
$0 \leq i \leq 19$	(B ∧ C) ∨ (¬(B) ∧ D)	5A827999
$20 \leq i \leq 39$	B ⊕ C ⊕ D	6ED9EBA1
$40 \leq i \leq 59$	(B ∧ C) ∨ (B ∧ D) ∨ (C ∧ D)	8F1BBCDC
$60 \leq i \leq 79$	B ⊕ C ⊕ D	CA62C1D6

Operacije u tablici odgovaraju standardnim logičkim veznicima primijenjenim nad nizovima bitova: ∧ – konjunkciji, ∨ – disjunkciji, ¬ – negaciji i ⊕ – ekskluzivnoj disjunkciji, to jest XOR operaciji.

Postupak svake od 80 iteracija, u kontekstu 160-bitnog spremnika, prikazan je i na slici 8. Funkcija f_i na slici odnosi se na funkcije prikazane u tablici 4.



Slika 8. Runda izračuna SHA-1 algoritma

Po završetku ovih 80 iteracija dolazimo do kraja rada SHA-1 algoritma. Korak koji je preostao je povezivanje zadnji 5 riječi (*A*, *B*, *C*, *D*, *E*) u hash poruke. Ključna operacija u ovom postupku je bitna disjunkcija – ∨. Ako H označava konačni hash, a bs_x operacije bitnog kružnog lijevog pomaka onda zadnji korak ovog algoritma glasi:

$$H = bs_{128}(A) \vee bs_{96}(B) \vee bs_{65}(C) \vee bs_{32}(D) \vee E$$

Ovdje prikazani algoritam više nije, to jest ne bi trebao biti u upotrebi ali to nije bio slučaj još u prošlom desetljeću. Prvi konkretni teorijski dokaz napada kolizijom na SHA-1 algoritme pojavio se je još 2013. i opisan je u radu Marca Stevensa,[174] a na temelju ovih teorijskih spoznaja Google je 2017. dao prvi konkretni primjer napada kolizijom na SHA-1 algoritam.[175] U tom članku kao rješenje problema Stevens i suradnici predlažu migraciju na

[174] Stevens, M. (2013). New collision attacks on SHA-1 based on optimal joint local-collision analysis. *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (str. 245-261). Berlin, Heidelberg: Springer. Preuzeto 27. 12. 2021. s https://link.springer.com/content/pdf/10.1007/978-3-642-38348-9_15.pdf

[175] Stevens, M., Bursztein, E., Karpman, P., Albertini, A. M., Bianco, A. P., & Baisse, C. (2017). *Announcing the first SHA1 collision*. Preuzeto 27. 12. 2022. s Google Security Blog: https://security.googleblog.com/2017/02/announcing-first-sha1-collision.html

SHA-2 (SHA-256) ili SHA-3 algoritme. Tu migraciju (na SHA-2) predložio je NIST još 2011. godine,[176,177] prije konkretnih dokaza ove ranjivosti.

Usprkos činjenici da se više ne koristi, SHA-1 je osnova SHA-2 algoritma koji je danas preporučeni standard te, samo po imenu, prethodnik SHA-3 algoritma koji je najvjerojatniji nasljednik SHA-2 nakon što se i on, s vremenom i s jačanjem procesorske snage, pokaže ranjivim. Osim toga, arhitektura SHA-1 hash funkcije vrlo je slična i onoj koju koristi (još stariji) MD-5 hash algoritam pa je smatram (najviše) karakterističnom hash funkcijom te je zbog toga u ovom poglavlju ona detaljno objašnjena.

3.1.1.3. SHA-2

SHA-2 algoritam odgovor je na dokazane ranjivosti njegovog prethodnika, SHA-1 algoritma. Iako je SHA-2 algoritam baziran na istoj strukturi kao i SHA-1, Merkle–Damgård strukturi,[178,179] nije dokazano da dijeli ranjivosti SHA-1 algoritma.[180] SHA-2 algoritam, koji je 2001.[181] razvila Nacionalna Sigurnosna Agencija Sjedinjenih Američkih Država,[182] još uvijek je u upotrebi iako je njegov nasljednik, SHA-3 već razvijen. SHA-2 razvijan je od 2002. do 2015. (kada se pojavila najnovija ali možda ne i konačna verzija standarda) te je prezentiran u tri NIST[183] dokumenta: FIPS PUB 180-2, FIPS PUB 180-3 i FIPS PUB 180-4. Zadnji od ovih dokumenata definira sve verzije SHA-2 algoritma koje su danas u uporabi.[184]

Hash funkcije bazirane na SHA-2 algoritmu dolaze u više varijanti, prvenstveno ovisno o dužini rezultirajućeg hasha. Ove varijante su, kronološkim redom dodavanja u NIST-ov Secure Hash Standard:

---

[176] Zadnja iteracija ovog dokumenta, koji donosi općenite preporuke za migraciju na nove kriptografske algoritme i dužine ključeva objavljena je 2019.
[177] National Institute of Standards and Technology. (2019). *Transitioning the Use of Cryptographic Algorithms and Key Lengths.* Preuzeto 27. 12. 2021. s https://csrc.nist.gov/publications/detail/sp/800-131a/rev-2/final
[178] Merkle, R. C. (1979). *Secrecy, authentication, and public key systems*, n. dj.
[179] Damgård, I. B. (1989). A design principle for hash functions, n. dj.
[180] Mendel, F., Nad, T., & Schläffer, M. (2011). Finding SHA-2 characteristics: searching through a minefield of contradictions. *International Conference on the Theory and Application of Cryptology and Information Security* (str. 288-307). Berlin, Heidelberg: Springer. Preuzeto 5. 5. 2020. s https://link.springer.com/content/pdf/10.1007/978-3-642-25385-0_16.pdf
[181] Penard, W., & van Werkhoven, T. (2008). On the secure hash algorithm family. *Cryptography in Context*, (str. 1-18). Preuzeto 5. 5. 2020. s https://web.archive.org/web/20160330153520/http://www.staff.science.uu.nl/~werkh108/docs/study/Y5_07_08/infocry/project/Cryp08.pdf
[182] Engl. United States National Security Agency, skraćeno NSA
[183] Engl. National Institute of Standards and Technology
[184] National Institute of Standards and Technology. (2015). *NIST FIPS 180-4 Secure Hash Standard (SHS).* doi: https://doi.org/10.6028/NIST.FIPS.180-4

1. SHA-256, izvorno definirana u FIPS PUB 180-2[185] –ova funkcija rezultira hashem od 256 bita.
2. SHA-512, izvorno definirana u FIPS PUB 180-2[186] – ova funkcija rezultira hashem od 512 bita.
3. SHA-384, izvorno definirana u FIPS PUB 180-2[187] – ova funkcija rezultira hashem od 384 bita. Ovo je verzija SHA-512 funkcije sa skraćenim izlazom.
4. SHA-224, izvorno definirana u FIPS PUB 180-3[188] – ova funkcija rezultira hashem od 224 bita. Ovo je verzija SHA-256 funkcije sa skraćenim izlazom.
5. SHA-512/224, izvorno definirana u FIPS PUB 180-4[189] – ova funkcija rezultira hashem od 224 bita. Ovo je verzija SHA-512 funkcije sa skraćenim izlazom.
6. SHA-512/256, izvorno definirana u FIPS PUB 180-4[190] – ova funkcija rezultira hashem od 256 bita. Ovo je verzija SHA-512 funkcije sa skraćenim izlazom.

Iz gornjeg popisa vidljivo je da iako SHA-2 porodica danas sadrži 6 različitih funkcija, četiri od njih su varijante sa skraćenim izlazom dvije osnove SHA-2 funkcije: SHA-256 i SHA-512. Struktura ove dvije funkcije je gotovo identična te se razlikuje samo po duljini riječi s kojima radi (u slučaju SHA-256 to je 32 bita, a u slučaju SHA-512 64 bita), vrijednostima nekih konstanti i broju rundi u kojima se podaci obrađuju. Obje funkcije baziraju se na ranije prikazanoj SHA-1 strukturi.

U odnosu na SHA-1, SHA-2 algoritam na značajno složeniji način obrađuje blokove koji su pripremljeni u fazi pretprocesiranja. Proces pretprocesiranja je gotovo isti te rezultira 512-bitnim blokovima koji se predaju na daljnju obradu.

[185] National Institute of Standards and Technology. (2002). *NIST FIPS 180-2*, n. dj.
[186] National Institute of Standards and Technology. (2002). *NIST FIPS 180-2*,n. dj.
[187] Ibid.
[188] National Institute of Standards and Technology. (2008). *NIST FIPS 180-3 Secure Hash Standard (SHS).* Preuzeto 15. 5 2020 s National Institute of Standards and Technology: https://csrc.nist.gov/publications/detail/fips/180/3/archive/2008-10-31
[189] National Institute of Standards and Technology. (2015). *NIST FIPS 180-4*, n. dj.
[190] Ibid.

SHA-2 algoritam svaki od ovih 512-bitnih blokova obrađuje na sljedeći način:

1. Blok se dijeli na 16 32-bitne riječi, ovim riječima se dodaje još 48 riječi koje obično počinju kao niz nula. Ukupno na obradu odlazi 64 riječi od kojih prvih 16 sadrži podatke iz bloka koji se trenutno obrađuje.

2. Ako svaku riječ označimo s r_i, gdje je i redni broj riječi, a brr_x i brs_x označavaju bitni desni cirkularni i necirkularni pomak za x mjesta, onda se za svaku riječ od rednog broja 16 do kraja niza (prazne riječi) računaju s_0 i s_1 na način:

$$s_0 = brr_7(r_{(i-15)}) \oplus brr_{18}(r_{(i-15)}) \oplus brs_{13}(r_{(i-15)})$$

$$s_1 = brr_{17}(r_{(i-2)}) \oplus brr_{19}(r_{(i-2)}) \oplus brs_{10}(r_{(i-2)})$$

   s_0 i s_1 se koriste da bi izračunali ranije prazne riječi na način:

$$r_i = r_{(i-16)} + s_0 + r_{(i-7)} + s_1$$

3. Riječi koje se koriste u daljnjoj obradi, A – H (ovdje ih je 8, SHA-1 je koristio 5), postavljaju se na vrijednosti konstanti:

```
A₀:= 0x6a09e667
B₀:= 0xbb67ae85
C₀:= 0x3c6ef372
D₀:= 0xa54ff53a
E₀:= 0x510e527f
F₀:= 0x9b05688c
G₀:= 0x1f83d9ab
H₀:= 0x5be0cd19
```

4. Nakon pripreme svih 64 riječi SHA-2 funkcija za sažimanje može početi, prije početka obrade svih riječi (64 runde umjesto 80 rundi u SHA-1) potrebno je postaviti polje s konstantama k. Svaka runda koristi drugu od sljedećih konstanti:

```
k[0..63] =   0x428a2f98, 0x71374491, 0xb5c0fbcf, 0xe9b5dba5,
             0x3956c25b, 0x59f111f1, 0x923f82a4, 0xab1c5ed5,
             0xd807aa98, 0x12835b01, 0x243185be, 0x550c7dc3,
             0x72be5d74, 0x80deb1fe, 0x9bdc06a7, 0xc19bf174,
             0xe49b69c1, 0xefbe4786, 0x0fc19dc6, 0x240ca1cc,
             0x2de92c6f, 0x4a7484aa, 0x5cb0a9dc, 0x76f988da,
             0x983e5152, 0xa831c66d, 0xb00327c8, 0xbf597fc7,
             0xc6e00bf3, 0xd5a79147, 0x06ca6351, 0x14292967,
             0x27b70a85, 0x2e1b2138, 0x4d2c6dfc, 0x53380d13,
             0x650a7354, 0x766a0abb, 0x81c2c92e, 0x92722c85,
             0xa2bfe8a1, 0xa81a664b, 0xc24b8b70, 0xc76c51a3,
             0xd192e819, 0xd6990624, 0xf40e3585, 0x106aa070,
             0x19a4c116, 0x1e376c08, 0x2748774c, 0x34b0bcb5,
             0x391c0cb3, 0x4ed8aa4a, 0x5b9cca4f, 0x682e6ff3,
             0x748f82ee, 0x78a5636f, 0x84c87814, 0x8cc70208,
             0x90befffa, 0xa4506ceb, 0xbef9a3f7, 0xc67178f2
```

Osim ovih konstanti potrebno je napraviti i kopije početnih riječi A_0-H_0 u *A*-*H*.

SHA-2 funkcija za sažimanje provodi petlju od 64 koraka koja za svaku ranije pripremljenu riječ, r_i, računa t_0 i t_1 iz riječi *A*-*H* stvorenih u prethodnoj rundi. Ako, kao i ranije, brr_x i brs_x označavaju bitni desni cirkularni i necirkularni pomak za *x* mjesta onda je t_0 izveden iz riječi *A, B i C* na način:

$$s_0 = brr_2(A) \oplus brr_{13}(A) \oplus brs_{22}(A)$$

$$v = (A \wedge \mathrm{B}) \oplus (A \wedge \mathrm{C}) \oplus (B \wedge \mathrm{C})$$

$$t_0 = s_0 + v$$

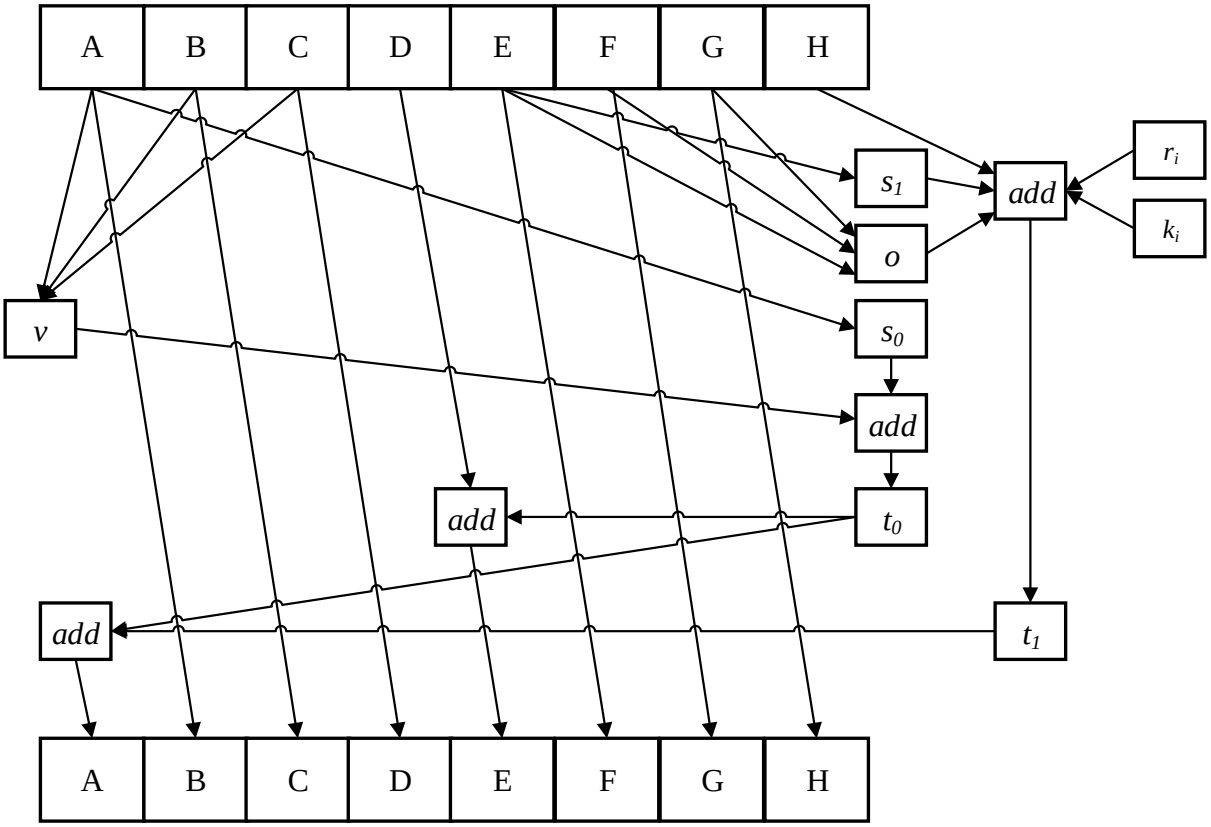
Izračun t_1, osim riječi *E*, *F*, *G* i *H* koristi i ranije pripremljene riječi, r_i, i konstante, k_i, u kojima i označava redni broj. Izračun se provodi na sljedeći način:

$$s_1 = brr_6(E) \oplus brr_{11}(E) \oplus brs_{25}(E)$$

$$o = (E \wedge \mathrm{F}) \oplus (\neg(E) \wedge \mathrm{G})$$

$$t_1 = H + s_1 + o + k_i + r_i$$

S izračunatim t_0 i t_1 mogu se provesti transformacije riječi: $H = G$, $G = F$, $F = E$, $E = D + t_0$, $D = C$, $C = B$, $B = A$, $A = t_0 + t_1$. Ove transformacije, to jest postupak izračuna koji se ponavlja u svakoj rundi prikazan je na slici 9. Ova shema, u usporedbi sa shemom na slici 8 jasno ilustrira složenost SHA-2 algoritma u odnosu na prethodnika.



Slika 9. Runda izračuna SHA-2 algoritma

5. Sažeti podaci se dodaju u riječi A_0-H_0 na način:

$$A_0 = A_0 + A$$

Na isti način računaju se i preostale riječi (B_0-H_0).

6. Konačna hash vrijednost se izvodi dodavanjem (engl. *append*) svih 8 riječi (*A*-*H*). Ako operaciju dodavanja označimo s >>, a hash s *h*, onda se konačni korak SHA-2 algoritma provodi na način:

$$h = A_0 >> B_0 >> C_0 >> D_0 >> E_0 >> F_0 >> G_0 >> H_0$$

Prikazani postupak izračuna hasha je najbolji kandidat za hash funkciju koja će provoditi izračune hash vrijednosti u novom TrustChain modelu. Napredniji, SHA-3 algoritam, je vjerojatni nasljednik ove funkcije nakon što postane standardiziran.

3.1.1.4. SHA-3

SHA-3 najmlađi je član Secure Hash Algorithm porodice i po načinu rada značajno se razlikuje od svojih prethodnika, SHA-1 i SHA-2 algoritama (koji je bio samo nadogradnja SHA-1 algoritma). SHA-3 definiran je 2015. u NIST-ovom[191] standardu FIPS PUB 202.[192] Algoritam se oslanja na Keccak funkciju koja je 2012. osvojila NIST-ovo natjecanje za temelj, tada još budućeg, SHA-3 algoritma.[193] SHA-3 je vrlo mlad algoritam i još nije zaživio kao standardni dio protokola koji koriste hash funkcije. NIST tek planira preporučiti njegovu upotrebu u sklopu standarda i protokola koji trenutno koriste funkcije iz SHA-2 porodice jer je ispunio jedan od osnovnih zahtjeva na natjecanju - interoperabilnost s SHA-2 funkcijama.[194] Upravo zbog svoje relativne mladosti i, pretpostavljeno, visoke razine sigurnosti predstavlja dobrog kandidata za hash funkciju koja će biti korištena u kasnije predloženim modelima ali s obzirom na zahtjeve za kompatibilnošću s trenutno aktualnom SHA-2 funkcijom, njegova upotreba u TrustChain modelu može početi službenu standardizaciju. Svejedno, SHA-3 je posebno važan hash algoritam s obzirom na to se od ovog algoritma očekuje otpornost na napade preslikom i nakon razvoja kvantnog računarstva. Chen i ostali autori izvještaja koji je NIST naručio 2016. navode otpornost SHA-2 i SHA-3 algoritama na napade izvedene kvantnim računalima uz obavezno povećanje izlaznog hasha.[195] Ovo je u skladu s procjenama u kojima su iste godine istraživači Matthew Amy i ostali pokazali da je za uspješan napada na SHA-2 i SHA-3 hash u 256 bitnoj varijanti potrebno 2^{166} procesorskih operacija što je svakako moguće upotrebom kvantnog računala.[196] Srećom, veličina hasha može, i u budućnosti će morati, biti veća, no već i danas oba algoritma mogu proizvesti hash od 512 bita.

---

[191] Engl. National Institute of Standards and Technology, institut za standardizaciju Sjedinjenih Američkih Država. URL: https://www.nist.gov/

[192] Dworkin, M. J. (2015). *SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions.* doi: https://doi.org/10.6028/NIST.FIPS.202

[193] Bertoni, G., Daemen, J., Peeters, M., & Van Assche, G. (2013). Keccak. *Annual international conference on the theory and applications of cryptographic techniques* (str. 313-314). Berlin, Heidelberg: Springer. Preuzeto 14. 5. 2020. s https://link.springer.com/content/pdf/10.1007/978-3-642-38348-9_19.pdf

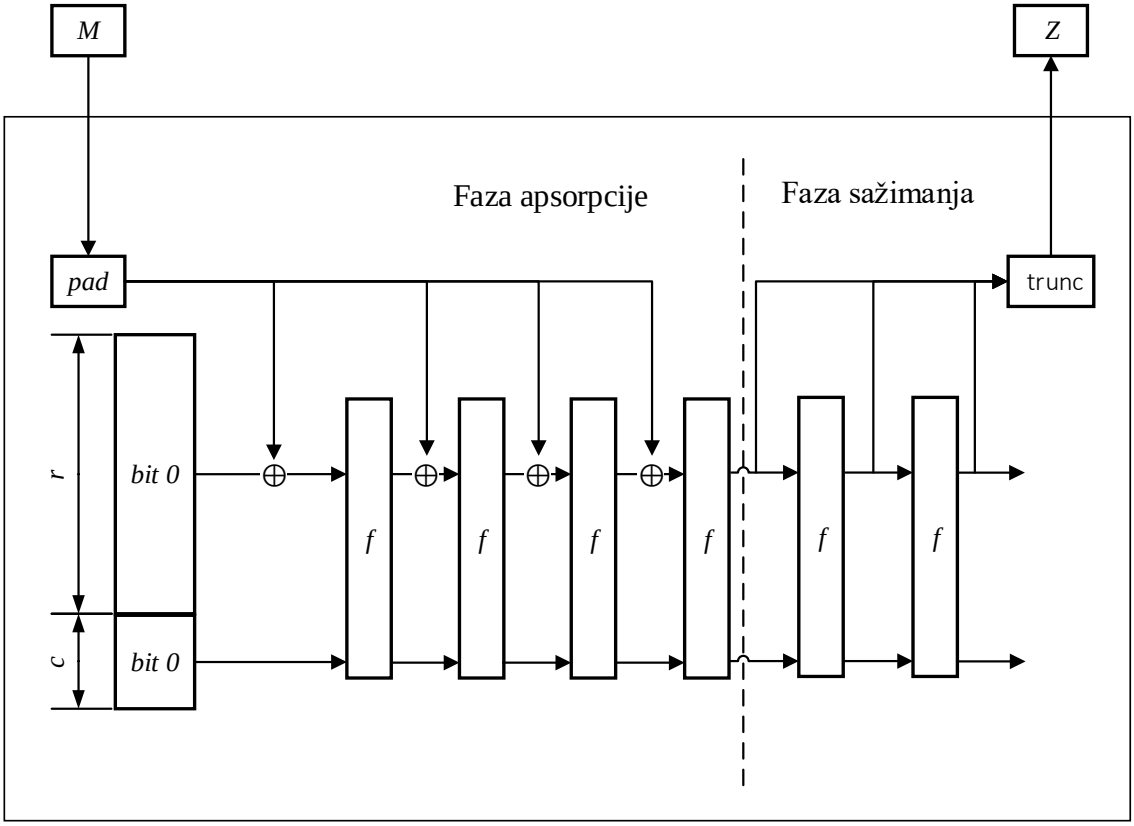
[194] Kayser, R. F. (2007). *Announcing request for candidate algorithm nominations for a new cryptographic hash algorithm (SHA-3) family.* Preuzeto 14. 5. 2020. s Federal Register: https://www.federalregister.gov/documents/2007/11/02/E7-21581/announcing-request-for-candidate-algorithm-nominations-for-a-new-cryptographic-hash-algorithm-sha-3

[195] Chen, L., Jordan, S., Liu, Y.-K., Moody, D., Peralta, R., Perlner, R., & Smith-Tone, D. (2016). *NISTIR 8105 – Report on Post-Quantum Cryptography.* National Institute of Standards and Technology, U.S. Department of Commerce. National Institute of Standards and Technology. doi: http://dx.doi.org/10.6028/NIST.IR.8105

[196] Amy, M., Di Matteo, O., Gheorghiu, V., Mosca, M., Parent, A., & Schanck, J. (2016). Estimating the cost of generic quantum pre-image attacks on SHA-2 and SHA-3. *International Conference on Selected Areas in Cryptography* (str. 317-337). Cham: Springer. Preuzeto 14. 5. 2020. s https://arxiv.org/pdf/1603.09383.pdf

SHA-3 algoritam temeljen je Keccek funkciji.[197] Keccak hash funkcija ima drastično različitu strukturu u odnosu na ranije opisane SHA funkcije. Srž ove funkcije je spužvasta konstrukcija (engl. *sponge construction*) koja dobila naziv zbog načina funkcioniranja funkcije pri kojem u prvoj fazi podaci upijaju u funkciju te se kasnije ispuštaju i smanjuju na unaprijed određenu veličinu (poput spužve koja upija i ispušta tekućinu).

Kao što je rečeno, spužvasta konstrukcija Keccak funkcije iterira kroz obradu podataka u dvije distinktivne faze, kao što je prikazano na slici 10.



Slika 10. Spužvasta struktura Keccak funkcije[198]

Općenito, možemo reći da faza apsorpcije uzima ulazne podatke te na njih primjenjuje funkciju f koja rezultira povećanom podatkovnom strukturom. U fazi sažimanja upotrebom iste funkcije veličina se smanjuje na početnu. Keccak funkcija rezultira izlaznim podacima jednake dužine ulaznim.[199]

[197] Bertoni, G., Daemen, J., Peeters, M., & Van Assche, G. (2013). Keccak, n. dj.
[198] Prema Team Keccak – The sponge and duplex constructions. URL: https://keccak.team/sponge_duplex.html
[199] Paar, C., & Pelzl, J. (2010). SHA-3 and The Hash Function Keccak. U C. Paar, & J. Pelzl, *Understanding Cryptography-A Textbook for Students and Practitioners.* Springer.

Funkcija *f* je skup funkcija od 5 funkcija zvan Keccak permutacije. Ove permutacije se koriste u obje faze Keccak funkcije ali na različite načine.[200] Keccak permutacije tvore slijedeće funkcije kojima su dana imena prema grčkim slovima te se njihovo funkcioniranje dijeli u četiri faze, one su:

1) Faza Theta (θ) funkcije.

Theta funkcija ulazne podatke dijeli u 5 nizova od kojih se svaki sastoji od 5 podnizova, dužine 64 bita (za ukupnu dužinu predanih podataka od 1600 bitova),[201] to jest u dvodimenzionalnu matricu nizova. Prema Paar i Pelzl ako svaki od 64 bitna podniza nazovemo *A[x,y]*,[202] a *C[x]* i *D[x]* međukoraci, operaciju desne rotacije za 1 bit nazovemo brs_1 onda Theta funkciju možemo opisati na način:[203]

$$C[x] = A[x,0] \oplus \mathrm{A[x,1]} \oplus \mathrm{A[x,2]} \oplus \mathrm{A[x,3]} \oplus \mathrm{A[x,4]} \quad za\ x = 0,1,2,3,4$$

$$D[x] = C[x-1] \oplus bsr_1(C[x+1]) \quad za\ x = 0,1,2,3,4$$

$$A[x] = A[x,y] \oplus D[x] \quad za\ x,y = 0,1,2,3,4$$

Faza rezultira s 25 *A[x]* riječi koje se predaju na daljnju obradu.

2) Faza Rho (ρ) i Pi (π) funkcija.

U ovoj fazi uzimaju se 25 riječi iz prethodne faze te se rotiraju[204] za određeni broj mjesta (ρ funkcija) i novi rotirani nizovi pomiču se na novo mjesto u matrici (π funkcija).[205] Ove promjene mogu se opisati jedinstvenom formulom:

$$B[y, 2x+3y] = brs_{r[x,y]}(A[x,y]) \quad za\ x,y = 0,1,2,3,4$$

Desni bitni pomak ($brs_{r[x,y]}$) provodi se prema broju mjesta određenom koordinatama x i y i sljedećom matricom:[206]

[200] Paar, C., & Pelzl, J. (2010). SHA-3 and The Hash Function Keccak, n. dj.
[201] Ovu veličinu (1600 bitova) zovemo stanje Keccak funkcije.
[202] Gdje su x i y koordinate niza u dvodimezionalnoj matrici nizova
[203] Paar, C., & Pelzl, J. (2010). SHA-3 and The Hash Function Keccak, n. dj.
[204] Bitni desni pomak
[205] Paar, C., & Pelzl, J. (2010). SHA-3 and The Hash Function Keccak, n. dj.
[206] Ibid.

Tablica 5. Vrijednosti bitnog pomaka Rho permutacije Keccek funkcije

	x = 3	x = 4	x = 0	x = 1	x = 2
y = 2	25	39	3	10	43
y = 1	55	20	36	44	6
y = 0	28	27	0	1	62
y = 4	56	14	18	2	61
y = 3	21	8	41	45	15

3) Faza Chi (χ) funkcije.

χ funkcija mijenja prethodno stvorene nizove (B) na način:

$$A[x,y] = B[x,y] \oplus (\neg(\mathrm{B}[\mathrm{x}+1,\mathrm{y}]) \wedge \mathrm{B}[\mathrm{x}+2,\mathrm{y}]) \qquad za\ x,y = 0,1,2,3,4$$

Rezultat ove promjene je nova matrica (1600 btino) stanje *A*.

4) Faza Iota (ι) funkcije.

Konačni korak u Keccak transformacijama je i najjednostavniji. ι funkcija provodi operaciju ekskluzivnog ili (XOR operacija) s RC konstantom i nizom A[0,0] (Paar & Pelzl, 2010)[207]:

$$A[0,0] = A[0,0] \oplus RC[i]$$

Vrijednost RC konstante ovisi o rundi transformacije, moguće vrijednosti prikazane su u tablici 6.

Tablica 6. Vrijednosti RC konstante Iota funkcije Keccek permutacija

RC[0]	0x0000000000000001	RC[12]	0x000000008000808B
RC[1]	0x0000000000008082	RC[13]	0x800000000000008B
RC[2]	0x800000000000808A	RC[14]	0x8000000000008089
RC[3]	0x8000000080008000	RC[15]	0x8000000000008003
RC[4]	0x000000000000808B	RC[16]	0x8000000000008002
RC[5]	0x0000000080000001	RC[17]	0x8000000000000080
RC[6]	0x8000000080008081	RC[18]	0x000000000000800A
RC[7]	0x8000000000008009	RC[19]	0x800000008000000A
RC[8]	0x000000000000008A	RC[20]	0x8000000080008081
RC[9]	0x0000000000000088	RC[21]	0x8000000000008080
RC[10]	0x0000000080008009	RC[22]	0x0000000080000001
RC[11]	0x000000008000000A	RC[23]	0x8000000080008008

[207] Paar, C., & Pelzl, J. (2010). SHA-3 and The Hash Function Keccak, n. dj.

Osim faza apsorpcije i sažimanja, za SHA-3 karakterističnih faza, prema autorima Keccak funkcije[208] rad algoritma zahtjeva još dvije faze u kojima se obavlja priprema za rad spužvaste funkcije. To su faza dopunjavanja i faza određivanja stanja (engl. *state size*). Cijeli algoritam prati sljedeće korake:

1) Faza dopunjavanja. Dopunjavanje je izravno ovisno o veličini izlaza (hasha). Kao što je bio slučaj i s SHA-2 funkcijom, SHA-3 je u stanju proizvesti hasheve različite duljine. U tablici 7 prikazane su moguće veličine izlaza. Dobro je odmah primijetiti da ove veličine odgovaraju veličinama koje proizvodi SHA-2 funkcija. Ovo je jedan od osnovnih zahtjeva za funkciju koji garantira kompatibilnost sa sustavima temeljenim na SHA-2 funkcijama. Za TrustChain model ovo znači da, u slučaju da se u prvoj fazi implementacije modela koristi SHA-2 funkcija, naknadni prijelaz na SHA-3 neće zahtijevati nikakve izmjene u podatkovnim strukturama (uključujući nepromjenjivi lanac blokova).

Tablica 7. Parametri SHA-3 funkcije prema veličini izlaza. Izvor: https://keccak.team/keccak.html

SHA-3 vrsta	Veličina izlaza	Stopa (r)	Kapacitet (c)
SHA-3/224	224	1152	448
SHA-3/256	256	1088	512
SHA-3/384	384	832	786
SHA-3/512	512	576	1024

Svrha SHA-3 dopunjavanja je da poruku dovede do veličine koja je višekratnik vrijednosti u stupcu Stopa (engl. *rate*). Samo dopunjavanje provodi se u dvije faze:

a) Dodaje se poseban niz bitova koji je karakterističan za veličinu izlaza, prema tablici:

Tablica 8. SHA-3 početak paddinga u usporedbi s veličinom izlaza (Paar & Pelzl, 2010)[209]

Veličina izlaza	Početak paddinga
224	11001
256	11101
384	11001
512	11101
Varijabilna (svi ostali izlazi)	1111

[208] Bertoni, G., Daemen, J., Hoffert, S., Peeters, M., Van Assche, G., & Van Keer, R. (2021). *Keccak,* n. dj.
[209] Paar, C., & Pelzl, J. (2010). SHA-3 and The Hash Function Keccak, n. dj.

b) Dodaje se niz koji počinje i završava s bitom 1, a između njih se dodaje potreban broj 0 bitova.

Ako je m_p poruka s paddingom, a n broj blokova koji će biti predani na obradu možemo reći da vrijedi:

$$m_p = n \times r$$

Na kraju ove faze, u daljnju obradu predaje se n blokova dužine r bitova.

2) Faza određivanja stanja.

Svrha ove faze je određivanje broja rundi u kojim će se obrađivati prethodno pripremljeni blokovi. Broj rundi izravno ovisi o veličini stanja, to jest o veličini podataka koji se predaju na obradu. Iako je Keccek funkcija u stanju raditi s više različitih veličina stanja, standardna veličina stanja za SHA-3 funkciju je 1600 bitova, što rezultira s 24 runde obrade podataka.[210]

3) Faza Keccak permutacija – apsorpcija. U fazi apsorpcije primjenjuju se ranije opisane Keccek permutacije na način da se svaki ranije pripremljeni blok, označen s Pi, obrađuje prema sljedećoj formuli:[211]

$$S[x, y] = S[x, y] \oplus Pi[x + 5y]$$

$$S = KP(S)$$

Gornja formula počinje s praznom matricom S[212] te funkcija KP označava ranije opisane Keccek transformacije.

[210] Paar, C., & Pelzl, J. (2010). SHA-3 and The Hash Function Keccak, n. dj.
[211] Bertoni, G., Daemen, J., Hoffert, S., Peeters, M., Van Assche, G., & Van Keer, R. (2021). *Keccak*, n. dj.
[212] Matrica 5 puta 5 nizova koji su popunjeni bitovima 0.

4) Faza Keccak permutacija–sažimanje. U konačnoj fazi podaci se opet predaju Keccak permutacijama ali rezultati se ovaj puta spajaju u jedinstveni niz. Ako operaciju spajanja (konkatenacije) označimo s || onda je ova faza opisana sljedećim formulama. (Bertoni, i dr., Keccak specifications summary, 2021)[213]

$$Z = Z \,||\, S[x, y]$$

$$S = KP(S)$$

Po završetku obrade Z sadrži hash izvornih podataka.

3.1.2. RSA kriptosustav

RSA kriptosustav, nazvan po prezimenima svojih autora Ron Rivesta, Adi Shamira i Leonarda Adlemana, predstavljen je 1977.,[214] te uz hash funkcije, predstavlja ključni kriptografski element nužan za uspostavu sustava javnog ključa, to jest za digitalni potpis.

Prema riječima autora RSA algoritma:

"Operacija podizanja broja na unaprijed određenu potenciju (op. a. privatni ili javni ključ) te izračun modula sa složenim modulusom se pokazao kao dovoljan za implementaciju "digitalnog potpisa"; načina za stvaranje prepoznatljivog, otpornog na falsificiranje i za dokument jedinstvenog digitalnog potpisa čiju autentičnost autor ne može kasnije opovrgnuti." (Rivest, Shamir, & Adleman, 1977)[215]

[213] Bertoni, G., Daemen, J., Hoffert, S., Peeters, M., Van Assche, G., & Van Keer, R. (2021). *Keccak,* n. dj.
[214] Rivest, R., Shamir, A., & Adleman, L. (1977). *On Digital Signatures and Public-Key Cryptosystems.* MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTER SCIENCE. Preuzeto 4. 1. 2022. s https://apps.dtic.mil/sti/pdfs/ADA039036.pdf
[215] Rivest, R., Shamir, A., & Adleman, L. (1977). *On Digital Signatures*, n. dj.

Kao i ostali algoritmi za asimetričnu kriptografiju, RSA se oslanja na dva posebna ključa, te je algoritam za generiranje ključeva ključan za funkcioniranje RSA kriptosustava. Ovaj algoritam prati sljedeće korake:[216]

1. Odabiru se dva različita prosta broja: p i q.

   p i q se biraju nasumično i predlaže se razlika u veličini od nekoliko znamenki. Ovi brojevi nisu ključevi ali moraju ostati tajni.

2. Računa se n. $n = p * q$.

   n je modulus javnog i tajnog ključa te se distribuira kao dio javnog ključa.

3. Računa se Carmichaelov tocijent[217] od n, $\lambda(n)$

4. Odabire se javni ključ e.

   Za e mora vrijediti $1 < e < \lambda(n)$ i da najveći zajednički djelitelj od e i $\lambda(n)$ mora biti 1, to jest e i $\lambda(n)$ moraju biti prosto relativni.

5. Odabire se privatni ključ d.

   Za d mora vrijediti $(d * e) \% \lambda(n) = 1$.

Rezultat gornjeg algoritma je par ključeva, javni ključ (n, e) i privatni ključ (n, d).

Koristeći ovaj par ključeva moguće je šifrirati i dešifrirati poruku. Školski primjeri algoritma često odmah prelaze na postupak šifriranja i dešifriranja ali stvarni sustavi nužno dodaju i korak dopunjavanja koji je u slučaju RSA algoritma potreban ne iz praktičnih već iz sigurnosnih razloga [218] te se dopunjavanje u kontekstu RSA algoritma ponekad zove i oklopljivanje (engl. *armouring*) poruke.[219] S obzirom na to da za razumijevanje potrebe za dopunjavanjem treba u potpunosti razumjeti RSA algoritam, ta problematika je objašnjena kasnije, iako po kronološkom slijedu izvođenja RSA algoritma spada na njen početak.

[216] Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1996). *Handbook of Applied Cryptography*, n. dj.
[217] Najmanji prirodni broj (n) za koji vrijedi $x^m \equiv 1 \ (mod\ n)$ ako je x prosto relativan s n. Izvor: https://mathworld.wolfram.com/CarmichaelFunction.html. Na primjer, za m=2 Charmicaelov tocijent je 8, $1^2 \% 8 = 1, 3^2 \% 8 = 1, 5^2 \% 8 = 1, \ldots$
[218] Paar, C., & Pelzl, J. (2009). *Understanding cryptography*, n. dj.
[219] Lawson, N. (2009). *Why RSA encryption padding is critical*. Preuzeto 23. 4. 2020. s rdist: https://rdist.root.org/2009/10/06/why-rsa-encryption-padding-is-critical/

Ako je *m* poruka koju treba šifrirati, a *c* rezultat šifriranja, onda se postupak šifriranja provodi na sljedeći način:

$$c = m^e \bmod n$$

Dešifriranje iste poruke provodi se na način:

$$m = c^d \bmod n$$

Ovo je uobičajeni postupak upotrebe RSA algoritma koji vrijedi kada je svrha upotrebe algoritma skrivanje podataka u prijenosu. Ako se RSA koristi u digitalnom potpisu uloga javnog i privatnog ključa je obrnuta. U tom slučaju javni ključ *(n, e)*, koji služi za šifriranje se drži tajnim, a privatni ključ *(n, d)*, koji služi za dešifriranje, se distribuira s digitalnim certifikatom.

Kao pomoć objašnjenju rada RSA algoritma prikazan je programski kod 1 u jeziku Python koji sadrži klasu (*RSA*) koja realizira funkcionalnosti RSA algoritma prema gornjem algoritmu (bez dopunjavanja). Programski kod je rad autora disertacije, osim funkcije *getGCD*, koja je opće poznata iterativna implementacija Euklidovog algoritma. Osnovna struktura klase uključuje četiri metode:

- *primeGen* metoda generira dva nasumična prosta broja.
- *keyGenPP* metoda generira par RSA ključeva, to jest provjerava zadovoljavaju li predani prosti brojevi uvjete RSA javnog i privatnog ključa.
- *encrypt* metoda prima privatni ključ i poruku koju šifrira.
- *decrypt* metoda prima javni ključ i poruku koju dešifrira.

RSA klasa je funkcionalna u stanju u kojem je dana, komentari koji opisuju programski kod nalaze se u linijama koje počinju s *#*.

Programski kod 1. Implementacija RSA algoritma u programskom jeziku Python

```
class RSA:
    def primeGen(self,n):
        # Metoda primeGen bira dva nasumična i različita prosta broja:
        # p i q, metoda se sastoji od dvije funkcije. chkPrime provjerava
        # je li predani broj n prost a funkcija getRandPrimes generira
        # nasumične proste brojeve. Ova funkcija realizira korak broj 1 u
        # ranijem opisu RSA algoritma.
        def chkPrime(n):
            for i in range(2,n):
```

```
            if n%i==0:
                return False
        return True
    def getRandPrimes(n):
        minn=100
        from random import randint
        # Nasumični odabir koristi randint funkciju random modula koji
        # dolazi sa standardnim distribucijama Pythona
        pC=False
        # Bira se prvi nasumični broj iz raspona od 100 do 10^n
        while not pC:
            p=randint(minn,10**n)
            pC=chkPrime(p)
            # Pozivom funkcije chkPrime provjerava se je li nasumično
            # odabrani broj prost
        randVar=randint(2,4)
        qC=False
        # Bira se drugi  nasumični broj iz raspona od 100 do 10^n*10^(x),
        # gdje je x nasumični broj između 2 i 4.
        while not qC:
            q=randint(minn,10**n*10**randVar)
            # Pozivom funkcije chkPrime provjeravamo je li nasumično
            # odabrani broj prost
            qC=chkPrime(q)
        return (p,q)
    return getRandPrimes(n)

def keyGenPP(self,p,q):
    # Metoda keyGenPP generira par RSA ključeva na osnovu dva
    # nasumično odabrana prosta broja. Prvi korak je generiranje modulusa
    # javnog i tajnog ključa n (točka 2 u ranijem opisu)
    n=p*q
    # U idućem koraku računa se tocijent prostih brojeva. Ovaj korak
    # odgovara točki broj 3 iz ranijeg opisa
    tot=(p-1)*(q-1)
    # Funkcija getGCD je standardna iterativna implementacija Euklidovog
    # algoritma za pronalazak najvećeg zajedničkog djelitelja dva broja, ova
    # funkcionalnost nužna je za kasnije korake.
    def getGCD(x,y):
        while y != 0:
            temp = y
            y = x % y
            x = temp
        return x
    # Prema točki 4 bira se javni ključ e. e mora biti prosto relativan s
    # odabranim tocijentom.
    for x in range(2,tot):
      if getGCD(x,tot)==1:
          e=x
          break
```

```
        # Prema točki 5 bira se privatni ključ d. d mora biti takav da modulus
        # umnoška d i e mora biti 1.
        for x in range(1,10):
            y = 1 + x*tot
            if y % e==0:
                d=int(y/e)
                break
        # Metoda vraća par ključeva (e,d) i modulus n
        self.n=n
        self.e=e
        self.d=d
        return [["public",n,e],["private",n,d]]
    # Metode encrypt i decrypt šifriraju i dešifriraju predanu poruku koristeći javni
    # i privatni ključ sadržan u klasi: self.e, self.d, self.n
    def encrypt(self,m):
        c=m**self.e%self.n
        return c
    def decrpyt(self,c):
        m=c**self.d%self.n
        return m
```

Paar i Pelzl u svoj knjizi uočavaju i objašnjavaju četiri sigurnosna nedostatka RSA algoritma:[220]

1. RSA enkripcija je deterministička (filozofska teza prema kojoj su svi događaji predodređeni događajima iz prošlosti ili prirodnim zakonitostima). Ovo znači da se za svaki ključ svaka konkretna poruka uvijek šifrira u točno isti šifrirani niz. Zbog ove osobine moguće je:
    a. statističkom analizom (više) šifriranih poruka otkriti podatke o izvornoj, nešifriranoj poruci,
    b. usporedbom parova šifriranih i nešifriranih poruka iz šifrirane poruke, bez upotrebe ključa, izvući djelomične podatke o nešifriranoj poruci (koja nije dostupna).
2. Nešifrirane vrijednosti 0, 1 i -1 ostaju iste nakon šifriranja, iz razloga što je rezultat potenciranja i modulusa ovih brojeva uvijek jednak njima samima.
3. Upotreba malog eksponenta javnog ključa e i kratkih poruka je ranjiva bez uporabe dopunjavanja.

[220] Paar, C., & Pelzl, J. (2009). *Understanding cryptography*, n. dj.

4. RSA algoritam je rastezljiv (engl. *malleable*). Pod ovim pojmom podrazumijevamo svojstvo da je šifrirani tekst moguće izmijeniti na način da on bude istovjetan rezultatu šifriranja drugog izvornog teksta bez upotrebe tajnog ključa. Ako se poruka *m* šifrirala na uobičajeni RSA način: $c = m^e \bmod n$ onda se ranjivost očituje na način:

    a. Odabire se cijeli broj *x*.

    b. Šifrirana poruka *c* se mijenja u *c′* na način $c' = x^e c$, gjde je *e* eksponent javnog ključa.

    c. Dešifrirana poruka m', koju računamo na način $m' = c'^d \bmod n$, biti će jednaka izvornoj vrijednosti pomnoženoj s *x*, to jest: $m = {m'}/{x}$.

Svojstvo opisano pod brojem 4 nije kritično kada se RSA algoritam koristi za skrivanje podataka ali je izuzetno opasno u slučaju upotrebe algoritma za digitalno potpisivanje jer dozvoljava napadaču da bez tajnog ključa falsificira potpis, to jest da generira potpis predvidljivog izvornog dokumenta bez upotrebe tajnog ključa.

Sva četiri navedena problema moguće je riješiti upotrebom posebnih algoritama za dopunjavanja poruke. U računarstvu se dopunjavanje često provodi na način da se poruka jednostavno produži do potrebne dužine dodajući bitove 0.[221] Za potrebe kriptografije razvijeni su posebni algoritmi za dopunjavanje koji u dodane podatke ubacuju nasumične elemente. PCKS#1 dozvoljava da se RSA dopunjavanje provede upotrebom više algoritama, neki od njih su OAEP i PSS.[222]

OAEP algoritam, engl. *Optimal Asymmetric Encryption Padding* ili u prijevodu: algoritam za optimalno dopunjavanje u asimetričnoj enkripciji, prvi puta su autori Mihir Bellare i Phillip Rogaway predstavili 1994.,[223] a suvremeni RSA, koji je opisan dokumentom RFC 817, koristi sebi prilagođenu verziju.

[221] Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1996). *Handbook of Applied Cryptography*, n. dj.
[222] Moriarty, K., B., K., Jonsson, J., & R Rusch, A. (2016). *RFC 8017 – PKCS #1: RSA Cryptography Specifications Version 2.2.* Preuzeto 23. 4. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc8017
[223] Bellare, M., & Rogaway, P. (1994). Optimal asymmetric encryption. *Workshop on the Theory and Application of of Cryptographic Techniques – EUROCRYPT 1994: Advances in Cryptology — EUROCRYPT'94* (str. 92-111). Berlin, Heidelberg: Springer. doi: https://doi.org/10.1007/BFb0053428

Za razumijevanje PCKS#1 varijante OAEP algoritma prvo objasnimo imena elemenata[224]:

- poruku označimo s *M*,
- dužinu poruke u bajtovima s |*M*|,
- opcionalnu labelu s *L*,
- dužinu RSA modulusa u bajtovima s *k*,
- dužina izlaza hash funkcije s |*H*|.

S gornjim elementima možemo započeti OAEP algoritam:[225]

1. Generira se niz *PS* koji se sastoji od nuli dužine *k* - |*M*| - 2|*H*| - 2 bajtova.
2. Generira se blok *DB* dužine *k* - |*M*| - 1 bajtova na način da spojimo *Hash(L), PS*, bajt heksadecimalne vrijednosti 0x01 i poruku *M*:

$$DB = Hash(L) \,||\, PS \,||\, 0x01 \,||\, M$$

3. Generira se nasumični bajt string *S* dužine |*H*|.
4. Generira se *DBmaska* na način:

$$DBmaska = MGF(S, k - |H| - 1)$$

   Funkcija za generiranje maske, *MFG*, često je *hash* funkcija poput SHA-1.

5. Generira se *DBmaskirani*:

$$DBmaskirani = DB \oplus DBmaska$$

6. Generira se *Smaska*:

$$Smaska = MGF(DBmaskirani, |H|)$$

7. Generira se *Smaskirani*:

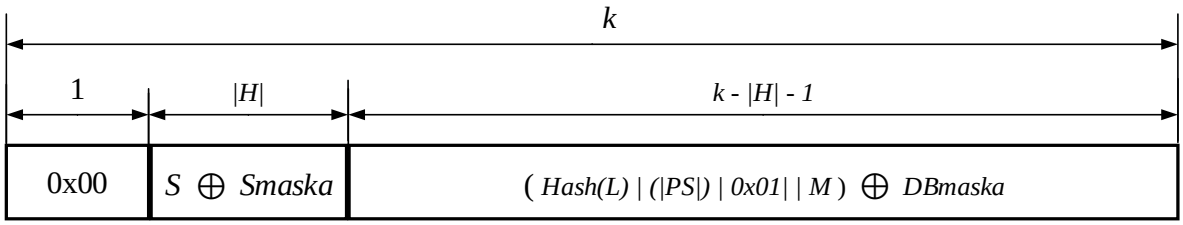
$$Smaskirani = S \oplus Smaska$$

[224] Moriarty, K., B., K., Jonsson, J., & R Rusch, A. (2016). *RFC 8017*, n. dj.
[225] Paar, C., & Pelzl, J. (2009). *Understanding cryptography*, n. dj.

8. Stvara se nova poruka *EM* dužine *k* bajtova na način:

$$EM = 0x00||Smaskirani||DBmaskirani$$

Konačni rezultat je dopunjena poruka prikazana na slici 11.[226]



Slika 11. OAEP poruka prije RSA enkripcije

Ovako pripremljena poruka predaje se ranije opisanom RSA algoritmu te je konačni rezultat šifrirana poruka koja se, ovisno o odabiru distribuira li se ključ za šifriranje ili dešifriranje, može koristiti u svrhu skrivanja sadržaja ili dokazivanja autentičnosti poruke (digitalni potpis).

3.1.3. X.509 digitalni certifikat

Ranije prikazani kriptografski mehanizmi omogućavaju dokazivanje autentičnosti zapisa na način da su u stanju garantirati da se podaci nisu izmijenili od trenutka njihovog potpisivanja. Kao što je rečeno u zaključku drugog poglavlja, arhivistika (i diplomatika), osim ovoga zahtjeva da je dokument potpisan na način koji identificira autora te da nisu izgubljeni podaci o arhivskim vezama u kojima sudjeluje zapis. Arhivska veza je arhivistički koncept koji računarstvo uglavnom ignorira te rješenje tog problema ne možemo naći u općeprihvaćenim tehnologijama koje su opisane u ovom poglavlju. U ovom poglavlju raspravit će se standardiziran način zapisa identiteta autora potpisa koji, uz javni ključ generiran RSA algoritmom, čini digitalni certifikat – x.509 standard.

Digitalni certifikat je opisan dokumentom x.509[227] kojega održava Internet Engineering Task Force (skraćeno IETF). Dokument opisuje sadržaj i organizaciju podataka u certifikatu, načine njihove distribucije i opoziva (revokacije, engl. *revocation*), algoritamski je agnostičan i u kombinaciji s ranije opisanim algoritmima za izračun hash vrijednosti i enkripciju javnim

[226] Paar, C., & Pelzl, J. (2009). *Understanding cryptography*, n. dj.
[227] Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., & Polk, W. (2008). *RFC 5280: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile.* Preuzeto 6. 17. 2019. s Internet Engineering Task Force (IETF) : https://tools.ietf.org/html/rfc5280

ključem čini temelj suvremenog digitalnog potpisa na način da standardizira način zapisa zadnjeg elementa potrebnog za funkcionalni digitalni potpis – identitet autora.

Tablica 9 prikazuje podatke sadržane u x.509 certifikatu.[228]

Tablica 9. Sadržaj X.509 digitalnog certifikata

X.509 certifikat	Verzija certifikata	
	Serijski broj certifikata	
	ID potpisnog algoritma	
	Naziv certifikacijskog autoriteta	
	Period valjanosti	Ne prije
		Ne nakon
	Naziv (ime) vlasnika certifikata	
	Podaci o javnom ključu	Naziv algoritma javnog ključa
		Javni ključ vlasnika certifikata
	Jedinstveni identifikator certifikacijskog autoriteta	
	Jedinstveni identifikator vlasnika certifikata	
	Dodatni podaci (opcionalno)	Lokacija distribucijske točke opozivnih listi (CDP)
		Lokacija pristupne točke za podatke o certifikacijskom autoritetu(AIA)
		Druga polja…
	Naziv algoritma potpisa certifikata (certifikacijskog autoriteta)	
	Javni ključ certifikacijskog autoriteta	

Osim sadržaja certifikata, x.509 definira sadržaj opozivne liste certifikata – CRL (engl. *Certificate Revocation List*).[229] Opozivne liste su posebni popisi certifikata koji su prestali vrijediti prije predviđenog vremena isteka (onog koje je navedeno u samom certifikatu). Razlog opoziva najčešće je gubitak (krađa) privatnog ključa vezanog uz certifikat, to jest uz javni ključ koji on sadrži.

Osim opozivnih listi status individualnog X.509 certifikata moguće je provjeriti i upotrebom OCSP (engl. *online certificate status protocol*). Ovaj protokol opisan u IETF (engl. *Internet Engineering Task Force*) dokumentu RFC 6960.[230]

[228] Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., & Polk, W. (2008). *RFC 5280,* n. dj.
[229] Ibid.
[230] Santesson, S., Myers, M., Ankney, R., Malpani, A., Galperin, S., & C., A. (2013). *RFC 6960: X.509 Internet Public Key Infrastructure Online Certificate Status Protocol – OCSP*. Preuzeto 11. 19 2021 s Internet Engineering Task Force (IETF): https://tools.ietf.org/html/rfc6960

X.509 opozivna lista prikazana je u tablici 10.

Tablica 10. Sadržaj X.509 opozivne liste certifikata

<table>
<tr><td rowspan="13">x.509 opozivna lista (CRL)</td><td colspan="3">Verzija opozivne liste</td></tr>
<tr><td colspan="3">Potpisni algoritam (certifikacijskog autoriteta)</td></tr>
<tr><td colspan="3">Naziv certifikacijskog autoriteta</td></tr>
<tr><td colspan="3">Vrijeme ažuriranja (ove liste)</td></tr>
<tr><td colspan="3">Vrijeme idućeg ažuriranja (opcionalno)</td></tr>
<tr><td rowspan="10">Podaci o opozvanim certifikatima</td><td rowspan="3">Opozvani certifikat 1</td><td>Serijski broj certifikata</td></tr>
<tr><td>Datum opoziva</td></tr>
<tr><td>CRL dodaci zapisa (opcionalno)</td></tr>
<tr><td rowspan="3">Opozvani certifikat 2</td><td>Serijski broj certifikata</td></tr>
<tr><td>Datum opoziva</td></tr>
<tr><td>CRL dodaci zapisa (opcionalno)</td></tr>
<tr><td>…</td><td>…</td></tr>
<tr><td rowspan="3">Opozvani certifikat n</td><td>Serijski broj certifikata</td></tr>
<tr><td>Datum opoziva</td></tr>
<tr><td>CRL dodaci zapisa (opcionalno)</td></tr>
<tr><td colspan="3">CRL dodaci (opcionalno)</td></tr>
<tr><td colspan="3">Potpis certifikacijskog autoriteta</td></tr>
</table>

Tablice 9 i 10 prikazuju sadržajnu shemu x.509 certifikata. Ovo je apstraktna shema, da bi omogućili prijenos podataka potrebno ih je zapisati na način koji nije čitak čovjeku, ali koji omogućava jednostavan prijenos između različitih uređaja. Postupak pretvaranja podataka iz ljudski čitljivog oblika u ovakav oblik zovemo serijalizacija podataka (i deserijalizacija, u suprotnom pravcu) ili kodiranje (i dekodiranje, u suprotnom pravcu) podataka, to jest računalnog zapisa.

PCKS standard[231] kao standard kodiranja X.509 certifikata navodi ASN.1 standard koji održava International Telecommunication Union pod nazivima X.680 (pravila kodiranja) i X.690 (specifikacija notacije).[232] ASN.1 navodi više načina na koje je moguće kodirati podatke ali PCKS predviđa upotrebu samo dva načina:

[231] OASIS Open. (2015). *PKCS #11 Cryptographic Token Interface Base Specification Version 2.4.* Preuzeto 28. 4. 2020. s OASIS | Advancing open standards for the information society: http://docs.oasis-open.org/pkcs11/pkcs11-base/v2.40/os/pkcs11-base-v2.40-os.html

[232] ITU. (2022). *X.680 : Information technology – Abstract Syntax Notation One (ASN.1): Specification of basic notation.* Preuzeto 5. 1. 2022. s International Communication Union: https://www.itu.int/rec/T-REC-X.680

- osnovnih pravila kodiranja, engl. *Basic Encoding Rules* (skraćeno BER) i
- karakterističnih pravila kodiranja, engl. *Distinguished Encoding Rules* (skraćeno DER).

Većina suvremenih certifikata se distribuira u DER formatu. U slučaju kada je certifikat samostalan, to jest nije dio neke digitalne datoteke (na primjer, PDF datoteke) često je DER kodiranje popraćeno i PEM notacijom. PEM definira na koji način se označava početak i kraj digitalnog certifikata. PEM (engl. *Privacy-Enhanced Mail*), je definiran u dokumentu RFC7468.[233]

U ostatku poglavlja dan je praktičan primjer dekodiranja PEM RSA certifikata i podataka o privatnom ključu upotrebom programskog jezika Python. Primjeri certifikata i ključa preuzeti su s FM4DD stranica za sigurnost informacijskih sustava.[234]

Niže je prikazan primjer PEM RSA certifikata s ključem od 1024 bita u svom kodiranom obliku.

Programski kod 2. PEM RSA digitalni certifikat

```
-----BEGIN CERTIFICATE-----
MIICVjCCAb8CAg37MA0GCSqGSIb3DQEBBQUAMIGbMQswCQYDVQQGEwJKUDEOMAwG
A1UECBMFVG9reW8xEDAOBgNVBAcTB0NodW8ta3UxETAPBgNVBAoTCEZyYW5rNERE
MRgwFgYDVQQLEw9XZWJDZXJ0IFN1cHBvcnQxGDAWBgNVBAMTD0ZyYW5rNEREIFdl
YiBDQTEjMCEGCSqGSIb3DQEJARYUc3VwcG9ydEBmcmFuazRkZC5jb20wHhcNMTIw
ODIyMDUyNzIzWhcNMTcwODIxMDUyNzIzWjBKMQswCQYDVQQGEwJKUDEOMAwGA1UE
CAwFVG9reW8xETAPBgNVBAoMCEZyYW5rNEREMRgwFgYDVQQDDA93d3cuZXhhbXBs
ZS5jb20wgZ8wDQYJKoZIhvcNAQEBBQADgY0AMIGJAoGBAMYBBrx5PlP0WNI/ZdzD
+6Pktmurn+F2kQYbtc7XQh8/LTBvCo+P6iZoLEmUA9e7EXLRxgU1CVqeAi7QcAn9
MwBlc8ksFJHB0rtf9pmf8Oza9E0Bynlq/4/Kb1x+d+AyhL7oK9tQwB24uHOueHi1
C/iVv8CSWKiYe6hzN1txYe8rAgMBAAEwDQYJKoZIhvcNAQEFBQADgYEAASPdjigJ
kXCqKWpnZ/Oc75EUcMi6HztaW8abUMlYXPIgkV2F7YanHOB7K4f7OOLjiz8DTPFf
jC9UeuErhaA/zzWi8ewMTFZW/WshOrm3fNvcMrMLKtH534JKvcdMg6qIdjTFINIr
evnAhf0cwULaebn+lMs8Pdl7y37+sfluVok=
-----END CERTIFICATE-----
```

[233] Josefsson, S., & Leonard, S. (2015). *RFC 7468: Textual Encodings of PKIX, PKCS, and CMS Structures.* Preuzeto 25. 4. 2020. s Internet Engineering Task Force (IETF): https://tools.ietf.org/html/rfc7468

[234] FM4DD. (2020). *X509 certificate examples for testing and verification – Public Key Infrastructure and Digital Certificates*. Preuzeto 26. 4. 2020. s FM4DD.com – Information Technology Security, Programming, Software: http://fm4dd.com/openssl/certexamples.htm

Ovakav certifikat jednostavno je dekodirati upotrebom Pythonovog SSL modula i njegove, još uvijek eksperimentalne klase _test_decode_cert (programski kod 3).

Programski kod 3. Funkcija za dekodiranje PEM digitalnog certifikata

```
#RSA cert key PEM read test

def certRead(path="d:\\",file="rsa-test.pem"):

    import ssl
    cert=ssl._ssl._test_decode_cert(path+file)
    return cert
```

Funkcija `certRead()` vratit će rječnik[235] sa sadržajem certifikata. Za jednostavniji ispis podataka rječnika (programski kod 4) iskorišten je u Python ugrađeni PrettyPrinter modul.[236]

Programski kod 4. Ispis x.509 certifikata upotrebom pprint modula

```
from pprint import pprint

pprint(certRead())
```

Rezultat gornjeg programa je raniji primjer certifikata (programski kod 2) u ljudski čitljivom obliku (programski kod 5):

Programski kod 5. x.509 certifikat u ljudski čitljivom obliku

```
{'issuer': ((('countryName', 'JP'),),
            (('stateOrProvinceName', 'Tokyo'),),
            (('localityName', 'Chuo-ku'),),
            (('organizationName', 'Frank4DD'),),
            (('organizationalUnitName', 'WebCert Support'),),
            (('commonName', 'Frank4DD Web CA'),),
            (('emailAddress', 'support@frank4dd.com'),)),
 'notAfter': 'Aug 21 05:27:23 2017 GMT',
 'notBefore': 'Aug 22 05:27:23 2012 GMT',
 'serialNumber': '0DFB',
 'subject': ((('countryName', 'JP'),),
             (('stateOrProvinceName', 'Tokyo'),),
             (('organizationName', 'Frank4DD'),),
             (('commonName', 'www.example.com'),)),
 'version': 1}
```

[235] engl. *dicitonary*, Pythonova podatkovna struktura za pohranu parova vrijednost-ključ: *dict*
[236] Python PrettyPrinter dokumentacija dostupna je na: https://docs.python.org/3/library/pprint.html

Primjer ključa preuzet iz istog izvora prati PCKS#1 standard, definiran u dokumentu RFC 2437.[237] Primjer u sirovom PEM obliku prikazan je programskim kodom 6.

Programski kod 6. RSA privatni ključ u kodiranom obliku

```
-----BEGIN RSA PRIVATE KEY-----
MIICWwIBAAKBgQDGAQa8eT5T9FjSP2Xcw/uj5LZrq5/hdpEGG7XO10IfPy0wbwqP
j+omaCxJlAPXuxFy0cYFNQlangIu0HAJ/TMAZXPJLBSRwdK7X/aZn/Ds2vRNAcp5
av+Pym9cfnfgMoS+6CvbUMAduLhzrnh4tQv4lb/AkliomHuoczdbcWHvKwIDAQAB
AoGAXzxrIwgmBHeIqUe5FOBnDsOZQlyAQA+pXYjCf8Rll2XptFwUdkzAUMzWUGWT
G5ZspA9l8Wc7IozRe/bhjMxuVK5yZhPDKbjqRdWICA95Jd7fxlIirHOVMQRdzI7x
NKqMNQN05MLJfsEHUYtOLhZE+tfhJTJnnmB7TMwnJgc4O5ECQQD8oOJ45tyr46zc
OAt6ao7PefVLiW5Qu+PxfoHmZmDV2UQqeM5XtZg4O97VBSugOs3+quIdAC6LotYl
/6N+E4y3AkEAyKWD2JNCrAgtjk2bfF1HYt24tq8+q7x2ek3/cUhqwInkrZqOFoke
x3+yBB879TuUOadvBXndgMHHcJQKSAJlLQJAXRuGnHyptAhTe06EnHeNbtZKG67p
I4Q8PJMdmSb+ZZKP1v9zPUxGb+NQ+z3OmF1T8ppUf8/DV9+KAbM4NI1L/QJAdGBs
BKYFObrUkYE5+fwwd4uao3sponqBTZcH3jDemiZg2MCYQUHu9E+AdRuYrziLVJVk
s4xniVLb1tRG0lVxUQJASfjdGT81HDJSzTseigrM+JnBKPPrzpeEp0RbTP52Lm23
YARjLCwmPMMdAwYZsvqeTuHEDQcOHxLHWuyN/zgP2A==
-----END RSA PRIVATE KEY-----
```

Dekodiranje ključa odrađeno je uz pomoć PyASN1 modula[238] i b64decode funkcije ugrađenog modula base64 (programski kod 7):[239]

Programski kod 7. Funkcija za dekodiranje RSA ključa

```
def keyRead(path="d:\\",file="rsa-keypair-test.pem"):

    with open(path+file,"r") as f:
        raw=f.read()[31:-30].replace("\n","")

    from base64 import b64decode

    DERkey=b64decode(raw)

    from pyasn1.codec.der.decoder import decode as der_decoder
    from pyasn1_modules import  rfc2437

    pkDef=rfc2437.RSAPrivateKey()

    RSAkey, add = der_decoder(DERkey, asn1Spec=pkDef)

    return   RSAkey, add
```

[237] Kaliski, B., & Staddon, J. (1998). *RFC 2437: PKCS #1: RSA Cryptography Specifications.* Preuzeto 26. 4. 2020. s Internet Engineering Task Force (IETF): https://tools.ietf.org/html/rfc2437
[238] Etingof, I. (2020). *ASN.1 modules for Python*. Preuzeto 26. 4. 2020. s GitHub: https://github.com/etingof/pyasn1-modules
[239] Python base63 dokumentacija dostupna je na: https://docs.python.org/3/library/base64.html

Gornji programski kod rezultirat će ispisom sadržaja privatnog RSA 1024 bitnog ključa (programski kod 8):

Programski kod 8. 1024 bitni RSA privatni ključ

```
RSAPrivateKey:
version=0

modulus=139043143640772622771388283044214826965667264427324172021986
633294306758315262236778983673330655705620167317073681229374461016425
016559632748800061643148093109919036725276250166884695479145020476584
808080671072691794919672879117563446037295443325307265043866528912693
788429412141749405485503929725519120166699

publicExponent=65537

privateExponent=668770003790902027099150320905708191649261258415568694
092963240159204168048119730453214750032697792939735063851838980029235
637501521266984310657404390365924863566194701849859857828221578295108
910607777909141600923888487632239328000127701470255725340618645079480
314989901923361030424276956112258973644572742 57

prime1=132312258264253660078884120457748249193102514793722130476014816
689721363285386622280202429607360468351783851106551275419192012349134
908927078893741483982 63

prime2=105087121529643956237039724759299660425373981941607478756880528
338477846331526969166769218783030774181470983689966800412544028236226
371062122670955955745 73

exponent1=487643664817550834726241214205174361202099293883999014223669
359647760777697366140433039273194712109582218767967352717773954234444
818131776654958181150207 7

exponent2=609513963651803723771663948235424064587838290059719468774362
977127762927011307071115130620013208689470997304121796545048083311917
761999605323247688754414 5

coefficient=387423447930555299993790644944901056561884534895710976493
242109245311126156510801601341234759007584629887538375774052084325820
923927992027152900753719701 6
```

Rezultat dekodiranja privatnog ključa su sam modulus *n* (`modulus`), javni eksponent (ključ) *e* (`publicExponent`), privatni eksponent (ključ) *d* (`privateExponent`), prim broj p (`prime1`), prim broj q (`prime2`), prim eksponent $d \bmod (p-1)$ (`exponent1`), prim eksponent $d \bmod (q-1)$ (`exponent2`) i koeficijent $q^{-1} \bmod p$ (`coefficient`). Ovakav

sadržaj privatnog ključa je u skladu sa specifikacijom definiranom u Kriptografskoj specifikaciji za sučelja tokena (engl. Cryptographic Token Interface Base Specification), to jest u standardu PKCS #11.[240]

U danom primjeru zorno je prikazan proces ekstrakcije i dekodiranja podataka iz računalnog zapisa certifikata, no treba biti svjestan da je u ovom primjeru prikazan proces dekodiranja podataka koji su već bili samostalni. Ipak, proces ekstrakcije potpisa iz digitalnog dokumenta poput PDF datoteke je značajno složeniji te će biti razmotren u kasnijem poglavlju. Javni ključ u ovakvom, samostalnom, obliku vidimo pri upotrebi digitalnog potpisa u mrežnim protokolima, na primjer, u slučaju SSL (engl. *Secure Socket Layer*) i TLS (engl. *Transport Layer Security*) protokola.

3.2. Vrste primjene i standardi digitalnog potpisa

Digitalni potpis se koristi u gotovo svim suvremenim informacijskim sustavima koji kao dio svoje funkcionalnosti uključuju prepoznavanje konkretne osobe, ustanove ili sustava. Općenito govoreći upotrebu digitalnog potpisa možemo podijeliti na:

1. digitalno potpisivanje dokumenata i zapisa,
2. digitalno potpisivanje elektroničke pošte,
3. digitalno potpisivanje podataka u prijenosu (SSL/TSL protokol) te
4. digitalno potpisivanje izvršnih datoteka.

Digitalni arhivi, a s njima i ovo istraživanje, su prije svega usredotočeni na prve dvije kategorije dok digitalni potpis funkcionira identično u sva četiri slučaja te se sve navedeno, uključujući i kasnije predložene modele informacijskih sustava, jednako primjenjuje na sve kategorije potpisanih podatka. Ipak, s obzirom na opseg istraživanja, primjene navedene pod 3 i 4 neće biti detaljnije proučene u ovom istraživanju.

Točke 1 i 2, digitalno potpisivanje dokumenata, zapisa i elektroničke pošte u Europskoj Uniji su regulirane Uredbom eIDAS (EU N°910/2014), engl. *Electronic Identification, Authentication and Trust Services*.[241] eIDAS postavlja zahtjeve za digitalni potpis i sustave vezane uz njih te između ostaloga uvodi pojam naprednog elektroničkog potpisa i

[240] OASIS Open. (2015). *PKCS #11*, n. dj.
[241] European Parliament and Council. (23. 7 2014). REGULATION (EU) No 910/2014, n. dj.

kvalificiranog elektroničkog potpisa. Napredni elektronički potpis mora ispuniti sljedeće uvjete:

1. potpisnik mora biti jedinstveno identificiran i povezan s potpisom;
2. potpisnik mora imati potpunu kontrolu nad podacima koji su korišteni za stvaranje naprednog elektroničkog potpisa,
3. potpis mora moći prepoznati jesu li potpisani podaci mijenjani nakon trenutka potpisivanja,
4. u slučaju da su potpisani podaci mijenjani potpis mora postati nevažeći.

U ovom poglavlju već je pokazano kroz opis rada digitalnog potpisa baziranog na RSA sustavima da oni ispunjavaju ove uvjete ili su oni više organizacijske nego tehničke prirode. Na primjer, točka 2 uopće nije tehničko pitanje koje se rješava unutar sustava za digitalni potpis. Sve što ona znači jest da certifikacijski autoritet koji dodjeljuje certifikat mora u svojem arhivu čuvati samo javnu komponentu PKI ključa, a čuvanje privatne komponente mora osigurati vlasnik potpisa. Točka 4 je dodana kao osnova za buduće zakone koji će se oslanjati na napredni digitalni potpis, to jest kao način da se jasno da do znanja da je integritet potpisanih podataka ključan za valjanost potpisa.

Točka 1 je vezana uz certifikat i certifikacijski autoritet te se nadovezuje na drugi važan pojam uveden Uredbom eIDAS: kvalificirani elektronički potpis, engl. *Qualified Electronic Signature*, skraćeno QES. Kao što je pokazano u drugom poglavlju, kvalificirani elektronički potpis je zakonski ekvivalent vlastoručnog potpisa. Ovakav potpis mora zadovoljiti uvjete naprednog elektroničkog potpisa koji zahtijevaju da je:

1. stvoren upotrebom kvalificiranog digitalnog certifikata te
2. stvoren na kvalificiranom uređaju za potpisivanje, engl. *Qualified Signature Creation Device*, skraćeno QSCD.[242]

[242] European Parliament and Council. (23. 7 2014). REGULATION (EU) No 910/2014, n. dj.

Da bi certifikacijski autoritet zadovoljivo točku 1, to jest da bi mogao izdati kvalificirani digitalni certifikat, on mora:

1. jasno navesti vrijeme valjanosti svih certifikata,
2. povući certifikate kojima je isteklo vrijeme valjanosti i poništiti povezane potpise,
3. zapošljavati isključivo pravilno obrazovane djelatnike i
4. koristiti softver i hardver koji je pouzdan i u stanju spriječiti falsificiranje certifikata.

Drugi zahtjev za kvalificirani elektronički potpis, da je potpis stvoren upotrebom kvalificiranog uređaja za potpisivanje QSCD-a, iako značajno podiže razinu sigurnosti, predstavlja i značajnu tehnološku, organizacijsku i logističku barijeru upotrebi digitalnog potpisa u svakodnevnom životu. Zahtjev podrazumijeva da svi korisnici naprednog elektroničkog potpisa (u idealnoj situaciji to su svi punoljetni građani neke države) imaju poseban hardverski uređaj koji omogućuje njegovu upotrebu. Izvan konteksta EU regulative, ovi uređaji poznati su i kao "sigurni uređaji za stvaranje potpisa", engl. *Secure Signature Creation Device*, skraćeno SSCD.

eIDAS u svom aneksu II definira zahtjeve za ove uređaje.[243] Ovi zahtjevi su dobra definicija končanih i potpunih mjera koje moraju biti osigurane da bi se sustav koji omogućava digitalno potpisivanje mogao smatrati sigurnim te onda biti i uzet u obzir kao dokaz na sudovima, što je slučaj s kvalificiranim elektronskim potpisima u Europskoj Uniji. Iz tog razloga ovi zahtjevi su u nastavku preneseni u potpunosti.

1. Uređaji za kvalificirani elektronski potpis moraju osigurati, kroz prikladne tehnološke i proceduralne načine, da je:
    a. povjerljivost podataka korištenih za stvaranje elektroničkog potpisa razumno osigurana,
    b. podaci koji se koriste za stvaranje elektroničkog potpisa se mogu pojaviti samo jednom (certifikati i ključevi su jedinstveni),

[243] European Parliament and Council. (23. 7 2014). REGULATION (EU) No 910/2014, n. dj.

c. podaci koji se koriste za stvaranje elektroničkog potpisa se, s razumnom sigurnošću, ne mogu izvesti (iz potpisa ili drugog izvora) te je potpis pouzdano zaštićen od falsificiranja upotrebom trenutno dostupne tehnologije,

d. pravni vlasnik može razumno zaštititi podatke koji se koriste za stvaranje elektroničkog potpisa radi sprječavanja neovlaštene uporabe.

2. Uređaji za kvalificirani elektronski potpis ne smiju mijenjati podatke koji se potpisuju ili na bilo koji način spriječiti prikazivanje tih podataka potpisniku prije potpisivanja.

3. Samo kvalificirani pružatelj usluga povjerenja, engl. *qualified trust service provider* (ovlašteni certifikacijski autoritet) može u ime potpisnika stvarati i upravljati podacima koji se koriste za potpisivanje.

4. Bez predrasuda prema točki (d) stavke 1, kvalificirani pružatelj usluga povjerenja koji pruža uslugu upravljanja podacima koji se koriste za potpisivanje u ime potpisnika smije stvoriti kopiju ovih podataka u svrhu pohrane sigurnosne kopije samo ako su ispunjeni sljedeći uvjeti:

   a. sigurnost dupliciranih podataka mora biti iste razine kao i izvornih,

   b. broj dupliciranih podataka ne smije prijeći minimalni broj koji je nužan da bi se osiguralo neometano funkcioniranje usluge (certifikacijskog autoriteta).

QSCD uređaji počeli su se koristiti krajem 20. i početkom 21. stoljeća upotrebom pametnih kartica[244] te je njihova upotreba i danas uobičajena. Na primjer, na slici 12 prikazan je USB token kakav fizičkim osobama izdaje Financijska agencija Republike Hrvatske.

Slika 12. Thales QSCD uređaj izdan fizičkoj osobi od Financijske agencije

[244] Leitold, H., & Konrad, D. (2019). Qualified Remote Signatures–Solutions, its Certification, and Use. *Proceedings of 29th SmartCard Workshop*, (str. 219-231). Preuzeto 29. 4. 2020. s https://pdfs.semanticscholar.org/3908/de8d4adb1dcbd88785e64aae37295147c58e.pdf

Iako je tehnički riječ o USB uređaju, u ovom QSCD uređaju jasno je vidljiva pametna SIM (engl. *subscriber identification module*) kartica. Da bi izbjegli logističke (i financijske) probleme vezane uz upotrebu ovakvog uređaja danas se sve više istražuju rješenja bazirana na mobilnim uređajima (pametnim telefonima) koji zadovoljavaju QSCD zahtjeve eIDASa. Primjer takvog uređaja (QSCD pametnog telefona) dali su Theuermann, Tauber i Lenz 2019.[245] Rješenje kakvo predlažu ovi autori pruža značajne prednosti u odnosu na ona temeljena na pametnim karticama jer umanjuje logističke probleme vezane uz upotrebu kvalificiranog elektroničkog potpisa da način da omogućuje da se umjesto fizičkog uređaja koji je potrebno koristiti u kombinaciji s računalom, kao što je slučaj s rješenjem Financijske agencije, koristi pametni telefon na koji je dovoljno instalirati posebnu aplikaciju.

Kvalificirani elektronički potpis koji se vodi prema navedenim smjernicama i koristi za potpisivanje digitalnog dokumenta, to jest zapisa izvodi se kroz ETSI standarde koji su tema ovog poglavlja. Oni su dio AdES[246] obitelji standarda – PAdES,[247] XAdES[248] i CAdES.[249] Sva tri standarda rezultat su rada Europskog instituta za telekomunikacijske standarde (engl. *European Telecommunications Standards Institute*, skraćeno ETSI). Iako ova neprofitna organizacija nosi ime europskog kontinenta u svom nazivu njezini standardi priznati su i izvan Europske unije i s preko 800 članova iz 66 različitih država[250] smatra se najvažnijom ustanovom za standardizaciju u području informacijsko komunikacijske tehnologije te će se ovo istraživanje, kada je to moguće, fokusirati na njihove, otvorene, standarde za razliku od zatvorenih ISO (engl. *International Organization for Standardization*) standarda.

---

[245] Theuermann, K., Tauber, A., & Lenz, T. (2019). Mobile-only solution for server-based qualified electronic signatures. *CC 2019-2019 IEEE International Conference on Communications (ICC)* (str. 1-7). IEEE. Preuzeto 5. 1. 2022. s https://www.researchgate.net/profile/Kevin-Theuermann-3/publication/334485288_Mobile-Only_Solution_for_Server-Based_Qualified_Electronic_Signatures/links/5d95c72c92851c2f70e62ea4/Mobile-Only-Solution-for-Server-Based-Qualified-Electronic-Signatures.pdf

[246] European Telecommunications Standards Institute. (2021). *ETSI EN 319 102-1 Electronic Signatures and Infrastructures (ESI) – Procedures for Creation and Validation of AdES Digital Signatures – Part 1: Creation and Validation.* Preuzeto 20. 11. 2022. s https://www.etsi.org/deliver/etsi_en/319100_319199/31910201/01.03.01_60/en_31910201v010301p.pdf

[247] European Telecommunications Standards Institute. (2016). *ETSI EN 319 142-1 Electronic Signatures and Infrastructures (ESI);PAdES digital signatures Part 1: Building blocks and PAdES baseline signatures.* Preuzeto 17. 6. 2019. s ETSI: https://www.etsi.org/deliver/etsi_en/319100_319199/31914201/01.01.01_60/en_31914201v010101p.pdf

[248] European Telecommunications Standards Institute. (2021). *ETSI EN 319 132-1: Electronic Signatures and Infrastructures (ESI); XAdES digital signatures; Part 1: Building blocks and XAdES baseline signatures.* Preuzeto 13. 2. 2020. s ETSI: https://www.etsi.org/deliver/etsi_en/319100_319199/31913201/01.02.00_20/en_31913201v010200a.pdf

[249] European Telecommunications Standards Institute. (2013). *ETSI TS 101 733 – Electronic Signatures and Infrastructures (ESI); CMS Advanced Electronic Signatures (CAdES).* Preuzeto s ETSI: http://www.etsi.org/deliver/etsi_ts/101700_101799/101733/02.02.01_60/ts_101733v020201p.pdf

[250] Prema ETSI Membership information: https://portal.etsi.org/Portal_IntegrateAppli/QueryResult.asp?Alone=1&SortBy=&SortDirection=&Param=

3.2.1. PAdES

Razvijen 2009. od stane ETSI-a, PAdES (engl. *PDF Advanced Electronic Signatures*) standard definira napredni elektronički potpis kojim se potpisuje PDF (engl. *Portable Document Format*) datoteka. Standard se bazira na ISO 32000-1[251] standardu koji opisuje PDF datoteku te je objavljen pod nazivom EN319 142-1.[252]

S obzirom na učestalost upotrebe PDF datoteka kao medija za digitalne dokumente PAdES standard je od posebne važnosti. Iako model koji je razvijeni tijekom ovog istraživanja nije razrađen do razine programskog koda koji rješava ekstrakciju potpisa iz PDF datoteke ovaj postupak je nužno razumjeti.

Izvorno je PAdES standard podijeljen na više dijelova koji definiraju različite načine digitalnog potpisivanja PDF dokumenta. Recentni PAdES standard je spojio neke od ovih inačica, a druge ukinuo ili preselio u druge standarde. Svejedno, reference na ove zastarjele standarde se još uvijek mogu naći u literaturi, a u kontekstu arhivistike uvijek je moguće naći zapis koji je napravljen prema starijim standardima. Također, druga dva ETSI standarda, XAdES i CAdES, još uvijek sadrže sličnu podjelu pa je ona ovdje navedena. Izvorni PAdES[253, 254] bio je podijeljen na:

- Osnovni PAdES definiran je u dokumentu TS 102 778 [2] te opisuje osnovni potpis, bez pohrane certifikata i bez provjere integriteta podatka. Ovakav potpis uopće ne zadovoljava zahtjeve arhivistike.
- Napredni PAdES – BES i EPES definirani su u dokumentu TS 101 733 [3] te opisuju potpis koji nužno uključuje potpisni certifikat te, opcionalno i digitalni vremenski žig.[255]

[251] International Organization for Standardization. (2008). *ISO 32000-1:2008 Document management — Portable document format — Part 1: PDF 1.7.* Preuzeto 1. 5. 2020. s ISO – International Organization for Standardization: https://www.iso.org/standard/51502.html

[252] European Telecommunications Standards Institute. (2016). *ETSI EN 319 142-1,* n. dj.

[253] European Telecommunications Standards Institute. (2013). *ETSI TS 103 172 V2 Electronic Signatures and Infrastructures (ESI); PAdES Baseline Profile.* Preuzeto 6. 1. 2022. s https://www.etsi.org/deliver/etsi_ts/103100_103199/103172/02.02.02_60/ts_103172v020202p.pdf

[254] European Telecommunications Standards Institute. (2009). *ETSI TS 102 778-1 V1.1 Electronic Signatures and Infrastructures (ESI); PDF Advanced Electronic Signature Profiles; Part 1: PAdES Overview – a framework document for PAdES.* Preuzeto 6. 1. 2022. s ETSI: https://www.etsi.org/deliver/etsi_ts/102700_102799/10277801/01.01.01_60/ts_10277801v010101p.pdf

[255] Digitalni vremenski žigovi opisani su u poglavlju 5

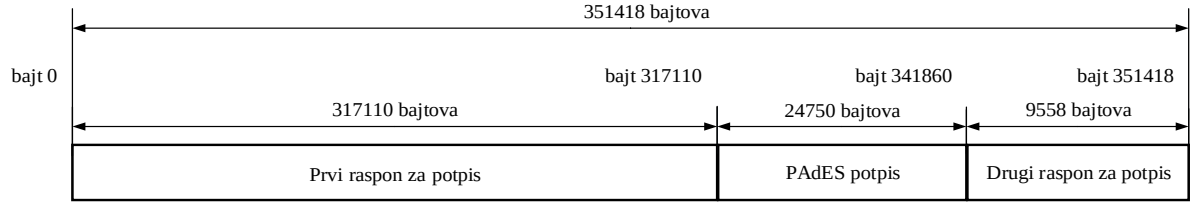
- Dugoročni PAdES – LTV definiran je u dokumentu TS 101 733 [4]. opisuje dodatak koji omogućuje dugoročnu validaciju potpisa, čak i nakon isteka potpisnog certifikata.
- Još četiri PAdES standarda koji opisuju potpisivanje XML dokumenata (ovi standardi su u međuvremenu prebačeni u zasebni XAdES standard, koji je specijaliziran za potpisivanje XML zapisa.

PDF datoteka sastoji se od niza objekata omeđenih posebnim oznakama, engl. *tag.*[256] PAdES potpis sadržan je u objektu označenim s oznakom `ByteRange` (raspon bajtova). `ByteRange` oznaka osim svog naziva sadrži i dva para cijelih brojeva. Ovi brojevi označavaju dio PDF datoteke na koji se potpis odnosi. Digitalni potpis se u pravilu odnosi na cijelu datoteku osim samog potpisa, to jest `ByteRange` objekta. Primjer kako ova oznaka može konkretno izgledati prikazan je programskim kodom 9.

Programski kod 9. PAdES ByteRange oznaka digitalnog potpisa

```
<</ByteRange[ 0 317110 341860 9558] … >>
```

Parovi cijelih brojeva u uglatim zagradama definiraju na koji dio datoteke se potpis odnosi. Prva dva broja se odnose na početak i dužinu prvog raspona (početak 0, dužina 317110), a idući par na isti način opisuje drugi raspon, sam potpis se smjestio u `ByteRange` objekt između ova dva raspona. Položaj potpisa i značaj ovih raspona prikazani su na slici 13.



Slika 13. Položaj zapisa PAdES potpisa u PDF datoteci

`ByteRange` sadrži cjelokupni potpis, to jest sve elemente PAdES potpisa (u bilo kojoj njegovoj verziji). Programski kod 10 prikazuje primjer sadržaja takvog potpisa.

[256] International Organization for Standardization. (2008). *ISO 32000-1:2008*, n. dj.

Programski kod 10. PAdES digitalni potpis

```
123 0 obj
<</ByteRange[ 0 317110 341860 9558]
      /Contents<30822ce106092a864886f70d010702a…>/
      Filter/Adobe.PPKLite
      /M(D:20180510094444+02'00')
      /Name(þÿ H R V O J E   S T A N I)
      /Prop_Build<</App<<
      /Name/Exchange-Pro
      /OS[/Win]
      /R 720905
      /REx(11.0.9)
      /TrustedMode true>>
      /Filter<</Date(Sep 12 2014 09:44:22)
            /Name/Adobe.PPKLite
            /R 131104/V 2>>
      /PubSec<</Date(Sep 12 2014 09:44:22)
            /NonEFontNoWarn true
            /R 131105>>>>
      /Reference[<</Data 193 0
            R/DigestLocation[ 342263 34]
            /DigestMethod/MD5/DigestValue<4e0bbe246f4e422633b55a05b2
            0c2d02>
            /TransformMethod/DocMDP
            /TransformParams<</P 2
                  /Type/TransformParams
                  /V/1.2>>
            /Type/SigRef>>
                  <</Data 193 0 R/DigestLocation[ 342467 34]
                  /DigestMethod/MD5
                  /DigestValue<f35528d1b7a0555fdfbbd520db9eb300>
                  /TransformMethod/FieldMDP
                  /TransformParams 181 0 R
                  /Type/SigRef>>]
/SubFilter/adbe.pkcs7.detached
/Type/Sig>>
endobj
```

Sam potpis nalazi se u `Contents` oznaci čiji je sadržaj predug da bi ovdje bio prikazan. Sadržaj te oznake zahtjeva dekodiranje na način koji je opisan u ranijem poglavlju, zapis je kodiran ASN.1 notacijom i DER formatom. Prikazane tehnike omogućuju provjeru i izolaciju digitalnog potpisa iz PDF dokumenta što je preduvjet za funkcioniranje TrustChain modela.

3.2.2. XAdES

XAdES (skraćenica za engl. XML Advanced Electronic Signatures) standard je ETSI proširenje XML-Dsig specifikacije, to jest XML sintakse za digitalni potpis World Wide Web Consortiuma.[257] Funkcionalno XML-Dsig vrlo je sličan PKCS#7[258] specifikaciji digitalnog potpisa ali je usredotočen na potpisivanje XML dokumenata. XML (engl. *Extensible Markup Language*) je engl. *markup* jezik koji omogućava organiziranu pohranu podataka u tekstualne datoteke koji je definirao World Wide Web Consortium, skraćeno W3C[259]. XML-Dsig, pa onda i XAdES potpisi prvenstveno nalaze svoju primjenu u internetskim tehnologijama, poput SOAP[260] protokola i SAML[261] jezika, gdje se koriste za autorizaciju i potvrdu integriteta podataka koji putuju internetom. Kao takvi oni su usredotočeni na sigurnosni aspekt ovakvih tehnologija ali XML-Dsig, i njegov nasljednik, XAdES su definirani dovoljno općenito da se mogu, a donekle su za to i predviđeni, koristiti za upotrebu XML zapisa u bilo kakvoj okolini.

U svojem osnovnom obliku XML-Dsig je jednostavni set podataka zapisan XML jezikom (programski kod 11).

Programski kod 11. XML-Dsig digitalni potpis

```
<Signature>
  <SignedInfo>
    <CanonicalizationMethod />
    <SignatureMethod />
    <Reference>
       <Transforms />
       <DigestMethod />
       <DigestValue />
    </Reference>
    <Reference />
  </SignedInfo>
  <SignatureValue />
  <KeyInfo />
  <Object />
</Signature>
```

[257] World Wide Web Consortium. (2013). *XML Signature Syntax and Processing Version 1.1.* Preuzeto 2. 5. 2020. s World Wide Web Consortium (W3C): https://www.w3.org/TR/xmldsig-core/
[258] Kaliski, B. (1998). *RFC2315 – PKCS #7 – Cryptographic Message Syntax.* Preuzeto 2. 5. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc2315
[259] World Wide Web Consortium. (2013). *XML Signature*, n. dj.
[260] Engl. *Simple Object Access Protocol* – protokol za razmjenu poruka između world wide web servisa
[261] Engl. *Security Assertion Markup Language* – otvoreni standard za razmjenu autentifikacijskih podatka, na primjer, za autorizaciju korisnika world wide web servisa

Ovakav potpis, iako funkcionalan, ne ispunjava pretpostavke X.509 standarda niti Uredbe eIDAS. Kritični element koji nedostaje je identitet autora. Ovakav digitalni potpis garantira integritet (nepromjenjivost podataka) ali nije dostatan za dokazivanje autentičnosti u arhivskom kontekstu.

Izvorni ETSI XAdES standard proširuje XML-Dsig specifikaciju na način da ga usklađuje s Uredbom eIDAS i definira šest profila digitalnog potpisa XML zapisa koji ispunjavaju uvjete naprednog elektroničkog potpisa. Profili definirani u izvornom XAdES standardu su 2016. povučeni te je definiran jedinstveni XAdES profil s više opcionalnih dodataka, ali suvremeni standardi i dalje održavaju kompatibilnost s njima. Ovi profili su:

- XAdES-BES (engl. Basic electronic signature), osnovna varijanta XAdES potpisa, definira najjednostavniji XML potpis koji zadovoljava eIDAS uvjete,
- XAdES-T, definira načine dodavanja vremenskog žiga digitalnom potpisu XML datoteke,
- XAdES-C, definira načine dodavanja referenci potpisnih podataka, certifikata i revokacijskih listi (bez dodavanja samih podataka),
- XAdES-X, prošireni potpis dodaje više vremenskih žigova koji pružaju zaštitu podacima definiranim u XAdES-C od kompromitiranja certifikacijskih lanaca na način da se osigura integritet podataka nakon isteka certifikata,
- XAdES-X-L definira načine pohrane podatka iz XAdES-C uz sam digitalni potpis,
- XAdES-A, arhivska varijanta XAdES potpisa, dodaje mogućnost naknadnog periodičkog dodavanja vremenskih žigova, slično XAdES-X modelu.

Struktura XAdES-BES potpisa, minimalne varijante XML potpisa koji zadovoljava eIDAS uvjete je prikazana u XML zapisu u priloženom programskom kodu 12.[262]

[262] European Telecommunications Standards Institute. (2002). *ETSI TS 101 903 V1.1 XML Advanced Electronic Signatures (XAdES).* Preuzeto 6. 1. 2022. s ETSI: https://uri.etsi.org/01903/v1.1.1/ts_101903v010101p.pdf

Programski kod 12. XAdES digitalni potpis

```
<ds:Signature ID?>
<ds:SignedInfo>
  <ds:CanonicalizationMethod/>
  <ds:SignatureMethod/>
    (<ds:Reference URI? >
      (<ds:Transforms>)
      <ds:DigestMethod>
      <ds:DigestValue>
    </ds:Reference>)
  </ds:SignedInfo>|
<ds:SignatureValue>
 (<ds:KeyInfo>)
<ds:Object>
    <QualifyingProperties>
      <SignedProperties>
        <SignedSignatureProperties>
          (SigningTime)
          (SigningCertificate)
          (SignaturePolicyIdentifier)
          (SignatureProductionPlace)
          (SignerRole)
        </SignedSignatureProperties>
        <SignedDataObjectProperties>
          (DataObjectFormat)
          (CommitmentTypeIndication)
          (AllDataObjectsTimeStamp)
          (IndividualDataObjectsTimeStamp)
        </SignedDataObjectProperties>
      </SignedProperties>
      <UnsignedProperties>
        <UnsignedSignatureProperties>
          (CounterSignature)
        </UnsignedSignatureProperties>
      </UnsignedProperties>
    </QualifyingProperties>
</ds:Object>
</ds:Signature>
```

Prikazana struktura je značajno kompliciranija od osnovne XML-Dsig strukture i treba znati da se značajan dio ovih oznaka može dalje raščlaniti. Ipak za potrebe ovog istraživanja dovoljno je napomenuti `SignedProperties` oznaku koja sadržava podatke koji, s gledišta arhivistike, značajno unaprjeđuju osnovni XML potpis. Prema ETSI-u[263] u ovoj oznaci sadržani su svi elementi koji identificiraju autora potpisa. Nadalje XAdES standard, za razliku XML-

[263] European Telecommunications Standards Institute. (2021). *ETSI EN 319 132-1*, n. dj.

Dsig-a u potpis uključuje i potpisne mehanizme koji obuhvaćaju sam certifikat te na taj način onemogućavaju zamjenu certifikata.

XAdES potpisani XML dokumenti prikladni su za arhivske svrhe, što je i bio cilj XAdES standardizacije, pa su stoga i kompatibilni s novim TrustChain modelom. S obzirom na XML strukturu ovakvih zapisa sam potpis i certifikat je tehnički lakše provjeriti i izolirati od ostatka zapisa u usporedbi s PAdES standardom. Nažalost, XML zapisi su više prisutni kao dio komunikacijskih protokola ili dio komunikacije između informacijskih servisa, a manje kao način pohrane zapisa namijenjenih (izravnoj) ljudskoj upotrebi pa su, kao takvi, i manje zastupljeni u arhivskom digitalnom gradivu.

3.2.3. CAdES

CAdES standard zadnji je ETSI-ev standard koji definira način primjene digitalnog potpisa. Za razliku od PAdES i XAdES standarda, CAdES je univerzalan, to jest namijenjen je upotrebi u kombinaciji s bilo kakvim digitalnim podacima. Da bi se omogućila univerzalna upotreba CAdES potpisi se ne ugrađuju izravno u datoteku s podacima koje potpisuju već mogu biti dodatak bilo kakvim digitalnim podacima. Takav standard ima primjenu kao mehanizam potpisa elektroničkih transakcija (bez npr. računa u PDF formatu), elektroničke pošte i sličnih zapisa čiji oblik nije striktno definiran. Ovakav standard često zovemo i standard za potpisivanje poruka.

Kriptografska sintaksa za poruke (engl. *Cryptographic Message Syntax*, skraćeno CMS) je IETF-ov standard za kriptografsku zaštitu podataka. CMS se, kao i XML-Dsig temelji na PCKS#7 specifikaciji,[264] ali je opisan u vlastitom IETF dokumentu, RFC 5652.[265] CAdES, skraćeno za engl. *CMS Advanced Electronic Signatures*, je ETSI-ov standard koji CMS potpise usklađuje s Uredbom eIDAS i opisan je u dokumentu ETSI TS 101 733[266] koji uvelike preuzima tehničke aspekte iz IETF-ovog RFC 5126.[267]

[264] Kaliski, B. (1998). *RFC2315 – PKCS #7*, n. dj.
[265] Housley, R. (2009). *RFC 5626: Cryptographic Message Syntax (CMS).* Preuzeto 2. 5. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc5652
[266] European Telecommunications Standards Institute. (2013). *ETSI TS 101 733*, n. dj.
[267] Pinkas, D., Pope, N., & Ross, J. (February 2008). *RFC 5126: CMS Advanced Electronic Signatures (CAdES).* Preuzeto 2. 5. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc5126

Slično izvornim PAdES i XAdES standardima, CAdES je podijeljen na više profila, oni su:

- osnovni CAdES-BES, osnovna varijanta CAdES potpisa, ispunjava minimalne zahtjeve Uredbe eIDAS,
- CAdES-T, definira načine dodavanja vremenskog žiga potpisu,
- dugoročni CAdES-LT, definira način dodavanja podataka koji se koriste za potvrdu potpisa, digitalnog certifikata i revokacijskih listi,
- dugoročni arhivski CAdES-LTA, arhivska varijanta CAdES potpisa, omogućava naknadno, periodičko, dodavanje vremenskih žigova potpisu.

CAdES standard prati istu strukturu podatka kao i XAdES te vlastiti zapis realizira PEM DER kodiranjem koje je opisano u ranijem poglavlju. Slično XAdES potpisima, CAdES potpisi su vrlo jednostavni za obradu (s tehničkog aspekta) ali i, za sada, rjeđe prisutni u arhivskom gradivu.

3.2.4. ASiC

ASiC standard (skraćenica za pretince povezanih potpisa, engl. *Associated Signature Containers*) zadnji je ETSI standard koji je važno razmotriti. ASiC je definiran u dokumentu ETSI TS 102 918[268] i opisuje način na koji je moguće pohraniti više potpisa u jedan pretinac (engl. *container*). Riječ *potpis* u ovom kontekstu se u naravi odnosi na vremenske žigove, te iako ovaj standard omogućava širu upotrebu, njegova primarna svrha je omogućiti tehničke uvjete koji su potrebni za realizaciju arhivske varijante ostalih ETSI standarda koji definiraju upotrebu naprednog elektroničkog potpisa (XAdES-A i CAdES-LTA, dok PAdES – LTV ovo definira u sklopu vlastitog standarda). ASiC je na taj način, uz ranije navedene varijante primjene naprednog elektroničkog potpisa, ETSI-ev odgovor na problem dugoročne pohrane digitalno potpisanoga arhivskoga gradiva. Iako ovakav pristup omogućava takvu pohranu te kao takav nudi rješenje dugoročne pohrane digitalno potpisanoga gradiva, on pretpostavlja periodičko dodavanje vremenskih žigova. Istraživanje mogućnosti izbjegavanja takvih periodičkih intervencija i razvoj modela koji bi to omogućio je osnovna svrha ove disertacije.

[268] European Telecommunications Standards Institute. (2013). *ETSI TS 102 918 – Electronic Signatures and Infrastructures (ESI);Associated Signature Containers (ASiC.* Preuzeto 9. 5. 2020. s ETSI: https://www.etsi.org/deliver/etsi_ts/102900_102999/102918/01.03.01_60/ts_102918v010301p.pdf

Zbog toga standard nije detaljno poručen već je samo spomenut, on predstavlja alternativu novom, TrustChain, modelu, a ne dio njegove funkcionalnosti.

3.3. Zaključak

Poglavlje je dalo pregled tehnološke podloge digitalnog potpisa i relevantnih standarda. Raspravljeni su hash algoritmi, asimetrični algoritmi za enkripciju te načini zapisa digitalnog certifikata. Sve navedeno čini osnovu za digitalni potpis te je pokazano da se upotrebom ovih tehnologija može ispuniti ranije raspravljene zahtjeve arhivistike i diplomatike. Osim toga, navedene tehnologije tvore temelj za novi model za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva – TrustChain.

Pri razvoju novog modela treba biti jasno koje podatke on prikuplja i pohranjuje, a to su prije svega digitalni potpis i certifikat, te kako se ostvaruje garancija integriteta podataka. Integritet podataka realiziran je upotrebom ulančanih zapisa, koji su opisani u idućem poglavlju. Ova, posebna, podatkovna struktura nije moguća bez upotrebe hash funkcija, koje su i dio digitalnog potpisa, pa im je posvećena posebna pažnja te je detaljno objašnjena najvažnija obitelj istih, obitelj SHA hash algoritama. Hash algoritmi istovremeno omogućavaju postojanje učinkovitog digitalnog potpisa i podatkovne strukture koja čini temelj postojećih sustava i novog TrustChain modela.

Na osnovu pregleda postojećih tehnologija moguće je tvrditi da suvremeni digitalni potpis zaista ispunjava potrebu za provjerom autentičnosti (digitalno) potpisane arhivske građe, možda i u većoj mjeri nego klasični, vlastoručni, potpis ili žig. Također, pokazan je način na koji se ovi podaci pohranjuju te je pokazano da se oni mogu jednostavno izdvojiti iz ostatka zapisa i neovisno o njemu pohraniti. Na ovaj način postavljeni su temelji za razvoj novog sustava koji će biti sukladan ranije raspravljenim načelima arhivistike i diplomatike.

4. Ulančani zapisi

Ovo poglavlje razmatra podatkovnu strukturu, ulančani zapis (engl. *blockchain*). Struktura podataka (engl. *data structure*) je termin iz područja računarstva koji se općenito definira kao strukturirani sistem za organizaciju, upravljanje i pohranu podataka koji omogućava učinkovit pristup i izmjenu podataka.[269] S obzirom da je svrha ovog istraživanja razviti model koji omogućava trajno očuvanje autentičnosti digitalno potpisanih (digitalnih) zapisa odabir odgovarajuće strukture podataka i njezino razumijevanje su temeljne pretpostavke istraživanja.

Za potrebe ovog istraživanja u hipotezu je uključena konkretna struktura podataka, ulančani zapis, zbog svojeg svojstva nepromjenjivosti. Nepromjenjiva struktura zapisa posebno je privlačna za digitalne arhivske sustave jer takvi (arhivski) sustavi i zahtijevaju da podaci ostanu nepromijenjeni nakon ulaska u (arhivski) sustav. Ovo svojstvo u kontradikciji je s ranije navedenom općom definicijom strukture podataka (koja omogućava izmjenu podataka). Konkretnija definicija kakvu su ponudili Wegner i Reilly bolje obuhvaća specifičnu podatkovnu strukturu kojom se ovaj radi bavi, ona definira strukturu podataka kao "skup podatkovnih vrijednosti, odnosa među njima, i funkcija ili operacija koje se mogu izvršiti nad podacima".[270]

Ovo poglavlje daje uvid u ranije spomenutu podatkovnu strukturu, ulančane zapise (engl. *blockchain*), te tako omogućava njezinu upotrebu kao temelj kasnije prikazanog modela za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva. Osim ulančanih zapisa poglavlje istražuje i dvije povezane podatkovne strukture (stablo i lista hasheva), koje se izuzetno često koriste u kombinaciji s ulančanim zapisima, kao i jedan koncept iz područja računarstva (lanac hasheva) koji je lako zabunom pomiješati s podatkovnom strukturom jer je vrlo sličan inače (u računarstvu) neimenovanoj podatkovnoj strukturi (koja se spominje u idućem poglavlju).

Pojam *ulančani zapisi* danas se često koristi u kolokvijalnom govoru ali je očigledno da je razina razumijevanja pojma relativno slaba. Termin se istovremeno koristi za opis sustava za kriptovalute, poput Bitcoin sustava,[271] i ostalih sustava koji koriste ulančane zapise, na primjer

[269] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms,* n. dj.
[270] Wegner, P., & Reilly, E. D. (2003). Data structures. U A. Ralston, E. D. Reilly, & D. Hemmendinger (Ur.). John Wiley and Sons Ltd. Preuzeto 7. 1. 2022. s https://dl.acm.org/doi/pdf/10.5555/1074100.1074312
[271] Bitcoin.org URL: https://bitcoin.org/en/

suvremenih sustava za vremenske žigove. Termin se najčešće koristi kao sinonim za kriptovalute, najčešće Bitcoin. Kao što će biti pokazano sam ulančani zapis uopće nije karakterističan za Bitcoin. Srž Bitcoina, i prava novost koju je uveo, je sustav za konsenzus koji je temeljen na dokazu rada.[272]

Ulančani zapis je pojam iz područja računarstva koji opisuje specifičnu podatkovnu strukturu. Ova podatkovna struktura sastoji se od više linearno povezanih blokova podataka (zapisa) među kojima svaki blok sadrži podatke s kojima je moguće dokazati podatkovni integritet prethodnih blokova. Ovaj dokaz integriteta prethodnih podataka postiže se upotrebom ranije opisanih kriptografskih tehnika, prije svega hash funkcija. Iz ovoga proizlazi svojstvo nepromjenjivosti zapisanih podatka i razlog odabira ove podatkovne strukture kao osnove modela za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva.

Ovakav ulančani zapis prvi puta se pojavljuje 1990. u radu Habera i Stornette "How to time stamp a digital document".[273] Iako navedeni autori u svom radu predviđaju kompleksniju strukturu od ovdje prikazane, upotrebom Merkleovog stabla,[274] njezina osnovna ideja prvi put je objavljena u tom radu. S obzirom na to možemo reći da izvorna namjena strukture podataka poznate kao ulančani zapis upravo (digitalno) arhivistička, a ne kako se danas često smatra financijska (povezana s kripto valutama).

Kriptovalute i suvremeno, kolokvijalno, shvaćanje ulančanih zapisa potječu iz dokumenta iz 2008., dakle 17 godina nakon prve pojave ulančanih zapisa. Te je godine osoba (ili osobe, jer pravi identitet autora do danas nije poznat) koja se predstavila kao Satoshi Nakamoto predstavila prvu funkcionalnu kriptovalutu – Bitcoin.[275] Nakamoto se u svome radu "Bitcoin: A peer-to-peer electronic cash system", poziva na Habera i Stornettu te njihov koncept proširuje na način koji omogućava potpuno decentraliziranu kontrolu nad ulančanim zapisom pomoću mehanizma koji je nazvao "dokaz rada" (engl. *proof of work*) i tako omogućio nastajanje prve općeprihvaćene kriptovalute, suvremenog novca koji banke ili države ne reguliraju.

Osim tog izvora (Nakamotovog rada) potrebno je spomenuti da, iako se autorstvo nad Bitcoin sustavom priznaje (neidentificiranom) Satoshiu Nakamotu, ideja decentralizirane

[272] Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system.* Preuzeto 7. 6. 2019. s https://bitcoin.org/bitcoin.pdf
[273] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.
[274] Merkle, R. C. (1979). *Secrecy, authentication, and public key systems*, n. dj.
[275] Nakamoto, S. (2008). *Bitcoin*, n. dj.

valute je 10 godina starija i prvi puta ju je predstavio Dai u svom prijedlogu iz 1998.[276] Od tada do danas, kriptovalute su se pokazale kao nepresušan izvor nesporazuma, od nerazumijevanja osnovnih pojmova povezanih uz njih, do nerazumijevanja njihovog tehničkog i financijskog funkcioniranja.

Mehanizmi dokaza rada su prava inovacija u Nakamotovom radu, dok je ulančani zapis je već bio poznat. S obzirom da je ovakva, slobodna valuta širokoj populaciji, pa čak i nekim stručnim krugovima, značajno interesantnija od arhivskog sustava za digitalne dokumente termin ulančani zapis (engl. *blockchain*) zaživio je kao sinonim za kriptovalute (umjesto općenite strukture podataka koja nalazi primjenu u arhivistici).

Ovo istraživanje će u idućem poglavlju razmotriti utjecaj, to jest mogućnost upotrebe sustava baziranih na dokazu rada (i nekim drugim decentraliziranim modelima) ali kada se govori o ulančanim zapisima podrazumijeva se izvorno značenje, povezana struktura podataka. S obzirom na to da je kasnije izrađeni model namijenjen upotrebi u arhivskim ustanovama, potreba za decentralizacijom postoji, ali je značajno manja nego u slučaju kripto valuta, te je kasnije predstavljeni model nadogradnja sustava koji su osmislili Haber i Storentta.[277] To se razlikuje od sličnih modela koji predstavljaju nastavak razvoja suvremenog, potpuno decentraliziranog, pravca razvoja ulančanih zapisa kakav koriste kriptovalute, na primjer, sustava baziranih na pametnim ugovorima (Ethereum ulančanim zapisima), ili bazama podataka koje koriste ulančane zapise kao dodatnu kontrolu integriteta podataka.[278] Ovakvi u potpunosti decentralizirani sustavi su od upitne koristi za arhivske svrhe, s obzirom na to da podrazumijevaju da postoji značajna količina korisnika kojima je u interesu koristiti sustav, a osnovna načela arhivske struke podrazumijevaju očuvanje zapisa radi dokazivanja odgovornosti za odluke donesene tijekom poslovanja ili obavljanja osnovnih funkcija institucije. Iz ovih razloga, smatram da se pri razvoju arhivskih modela treba osloniti na sudjelovanje arhivskih institucija, a ne široke populacije (kako to rade suvremene kriptovalute).

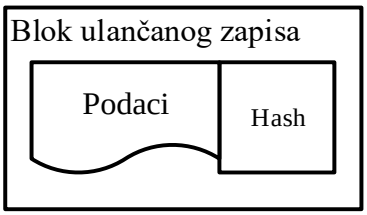
Hash funkcije čine osnovu ulančanih zapisa i oni bez njih nisu mogući. S obzirom da su hash funkcije detaljno opisane u trećem poglavlju u ovom poglavlju ih se neće dodatno razmatrati. Umjesto toga poglavlje će se usredotočiti na sam ulančani zapis i neke druge

[276] Dai, W. (1998). B-money proposal. Preuzeto 18. 5. 2020. s https://web.archive.org/web/20180328204908/http://www.weidai.com/bmoney.txt
[277] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.
[278] McConaghy, T., Marques, R., Müller, A., Jonghe, D. D., McConaghy, T., McMullen, G., . . . Granzotto, A. (2016). *Bigchaindb: a scalable blockchain database.* Preuzeto 19.11.2022. s BigChainDB: https://gamma.bigchaindb.com/whitepaper/bigchaindb-whitepaper.pdf

povezane računarske tehnike, na primjer, na Merkleova stabla[279] koja omogućuju učinkovit zapis velike količine hasheva i metode implementacije ulančanih zapisa. Prije svega poglavlje će razmotriti osnovni oblik ulančanog zapisa i arhivski model koji su predložili Haber i Stornetta.[280]

4.1. Struktura ulančanog zapisa

Struktura ulančanih zapisa podijeljena je, kao što je već spomenuto, na linearni niz blokova podataka. Ovi blokovi podataka su osnovna jedinica ovakve strukture podatka te njihov niz čini ulančani zapis. Ova, osnovna strukturna jedinica ulančanog zapisa, koji obično zovemo blok, prikazana je na slici 14.



Slika 14. Struktura bloka ulančanog zapisa

Sadržaj svakog bloka podijeljen je na barem dvije jedinstvene cjeline:

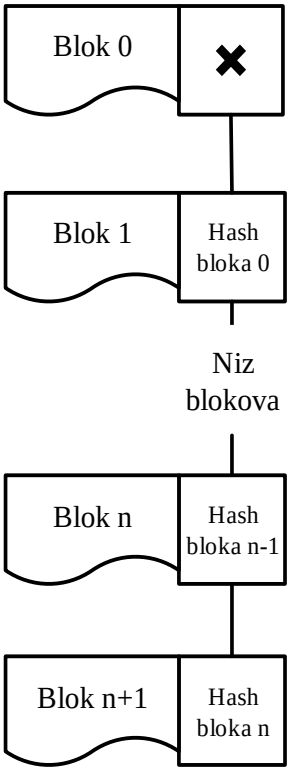
1) Podatke. Pod podacima ovdje podrazumijevamo podatke čija je pohrana svrha ulančanog zapisa. Na primjer, podatke o vremenskim žigovima ili podatke o transakcijama kriptovalute. Iako takvo ograničenje ne postoji, iz praktičnih razloga, većina sustava ulančanih zapisa u ovaj blok ne pohranjuje same kritične podatke, već njihove hasheve. Ovo se radi da bi se zaustavio pretjerani rast podatkovne strukture. Sustavi čiji kritični podaci nisu veliki mogu pohraniti i kompletne podatke unutar bloka. Dobar primjer ovakvog sustava je Bitcoin, koji su svoj ulančani zapis pohranjuje podatke o financijskim transakcijama.

2) Hash podataka. Ovaj dio bloka sadrži hash vrijednosti koje jamče integritet podataka. Termin hash vrijednosti (engl. *hash value* ili engl. *message digest*) odnosi se na rezultat hash funkcije koja je kao ulazne podatke primila cijeli prethodni blok, a u većini rješenja i podatke iz stavke broj 1 trenutnog bloka. Ova hash vrijednost može biti izračunata upotrebom bilo kojeg hash algoritma,

[279] Merkle, R. C. (1979). *Secrecy, authentication, and public key systems*, n. dj.
[280] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.

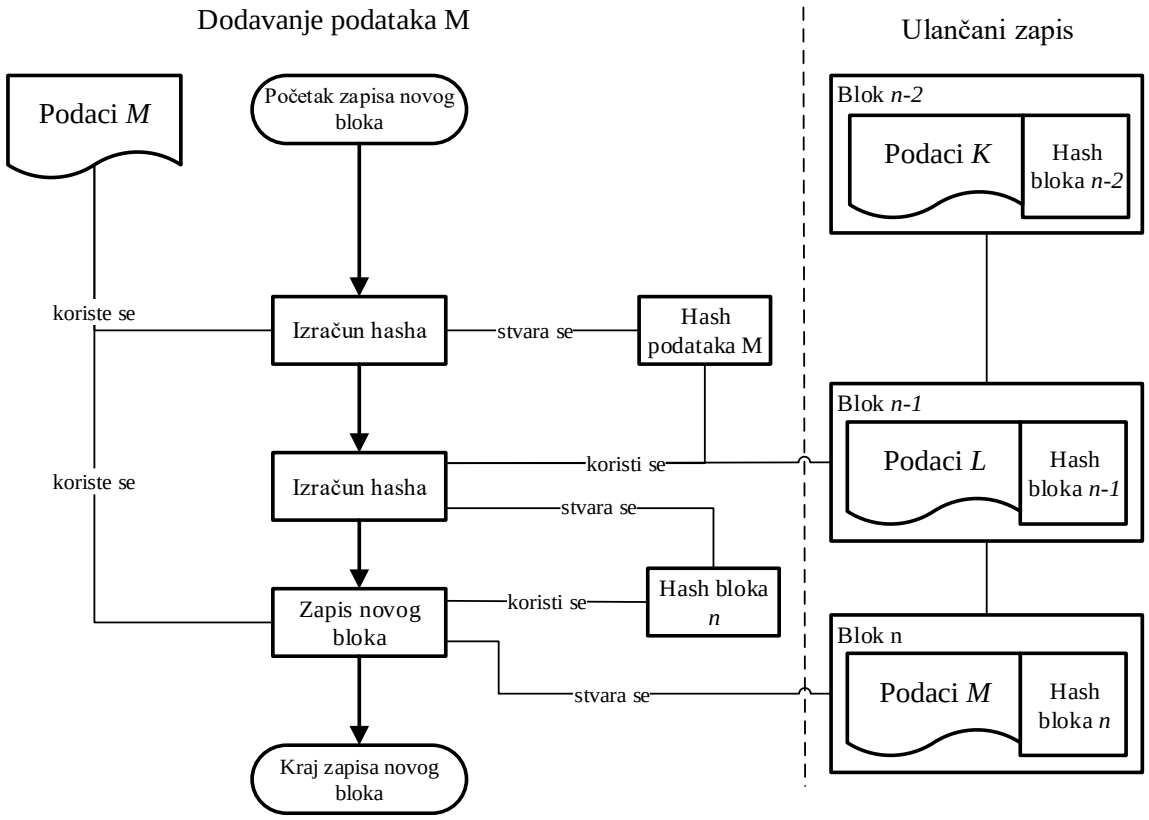
uključujući sve koji su razmatrani u ranijim poglavljima. U suvremenim sustavima izračun je najčešće realiziran upotrebom SHA-256 algoritma.

Zajedno ove dvije cjeline tvore strukturu podataka poznatu kao blok ulančanih zapisa. Povezivanjem ovakvih blokova dobiva se lanac ulančanih zapisa ili lanac blokova (engl. *blockchain*). Slika 15 prikazuje niz ovako ulančanih blokova (od bloka 0 do bloka n).



Slika 15. Ulančani zapis

Ako pretpostavimo gore opisanu, minimalnu, strukturu bloka ulančanog zapisa te zanemarimo potrebu za dodatnom obradom (meta) podatka onda je i proces stvaranja bloka relativno jednostavan. Konkretni sustavi, uključujući kasnije prikazani TrustChain model zahtijevaju kompliciranije procese. Proces stvaranja novog bloka bez ikakve dodatne obrade podataka opisan je na slici 16.



Slika 16. Proces stvaranja novog bloka ulančanog zapisa

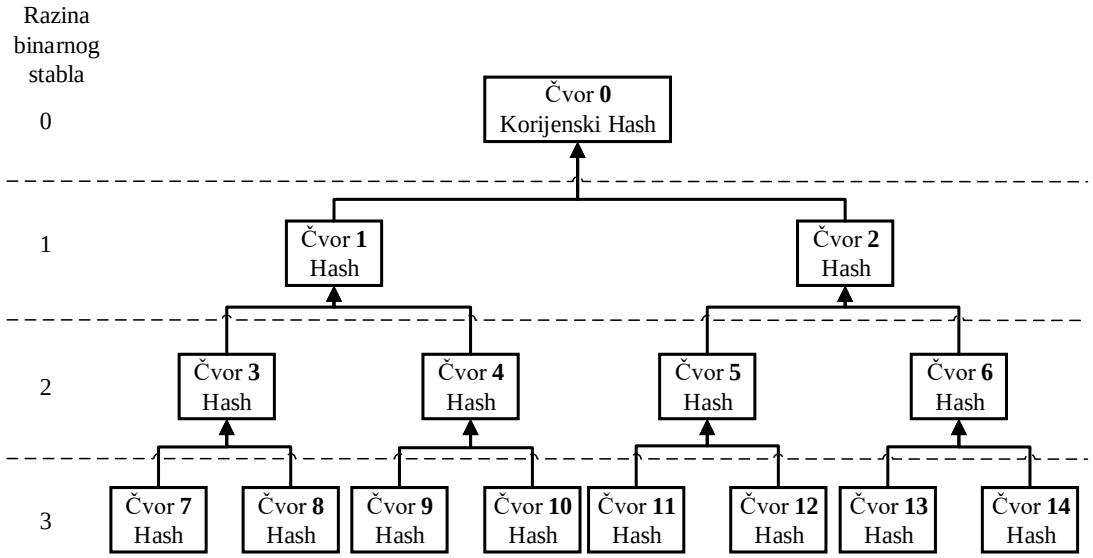
Slika 16 opisuje minimalni postupak stvaranja novog bloka. Dijagram između ostalog pokriva obje mogućnosti stvaranja bloka, stvaranje bloka s potpunim podacima i stvaranje bloka koji pohranjuje samo hash podatka. U gornjem dijagramu dodan je korak koji je moguće izbjeći. *Hash bloka n* mogao se izračunati izravno iz podataka *M* i hasha bloka *n-1*. Gornji dijagram dodaje još jedan izračun hasha koji omogućava da se umjesto samih podataka u ulančani zapis pohrani njihov hash – *Hash podataka M*.

S obzirom na ranije objašnjena svojstva hash funkcija (prije svega SHA algoritma) jasno je da je ovako organizirana podatkovna struktura nepromjenjiva. Bilo koja naknadna izmjena u prethodno uvrštenom bloku zahtjeva (zbog svojstva hash funkcija) izmjene u svim naknadnim blokovima. U slučaju da su ovi blokovi (ili njihovi hashevi) na neki način javno objavljeni te je njihovom postojanju u određenom obliku svjedočio veliki broj osoba ili ustanova takva promjena jednostavno nije moguća. Ovo je srž ideje ulančanih zapisa kakvu su predstavili Haber i Stornetta[281] i koja opisuje osnovu funkcioniranja svih sustava ulančanih blokova.

[281] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document , n. dj.

4.2. Stablo hasheva

Stablo hasheva je podatkovna struktura koja u potpunosti odgovora standardnom potpunom binarnom stablu. Pod pojmom potpunog binarnog stabla podrazumijevamo povezanu podatkovnu strukturu u kojoj svaki čvor ima najviše dva čvora djeteta te je zadnja razina stabla u potpunosti popunjena.[282] Ovo znači da će stablo koje ima n razina na zadnjoj razini imati 2^n čvorova, to jest listova stabla. Ovakvo stablo prikazano je na slici 17.



Slika 17. Merkleovo stablo (stablo hasheva, to jest kompletno binarno stablo hasheva)

Prikazana podatkovna struktura (binarno stablo) je poznata i temeljito istražena podatkovna struktura iz područja računarstva koje se bavi algoritmima i strukturama podatka, na primjer u knjizi "Uvod u algoritme".[283] Ovdje prikazano binarno stablo, koje često zovemo i hash stablo ili Merkleovo stablo specifično je po podacima koji su pohranjeni u čvorove, a to su hashevi. Hashevi se u ovakvom stablu računaju na način da se kao ulaz svake hash funkcije uzima konkatenacija (engl. *concatenation*) ili spajanje hasheva čvorova djece. Na primjer, hash čvora 3 izračunat je na osnovu spajanja hasheva koje sadrže čvorovi 7 i 8.

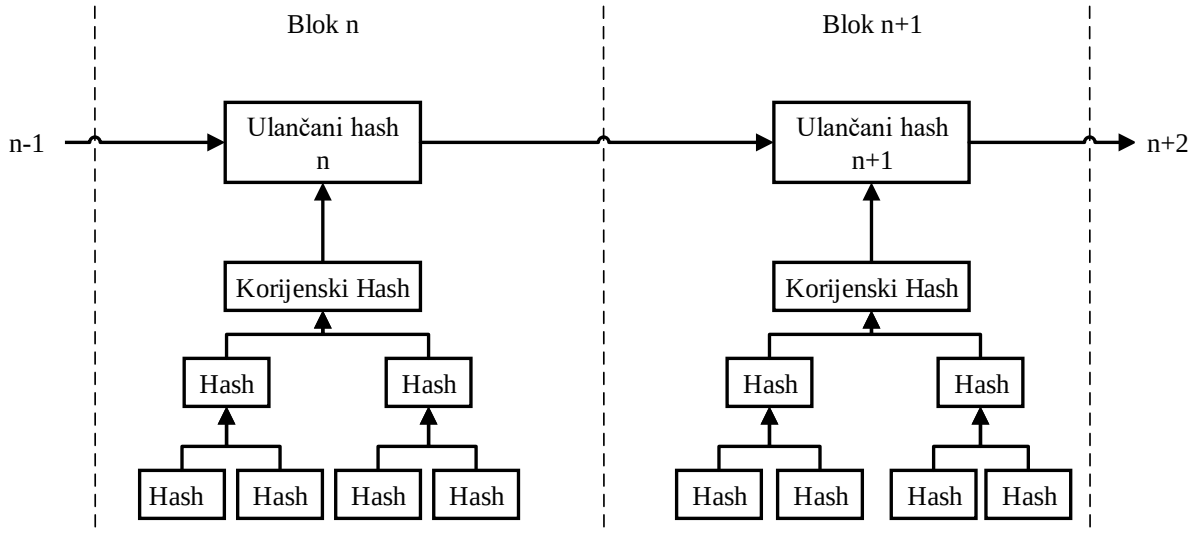
Ralph C. Merkle predložio je upotrebu binarnog stabla na ovaj način u svom radu iz 1980.[284] te ju je dvije godine kasnije i patentirao.[285]

---

[282] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms*, n. dj.
[283] Ibid.
[284] Merkle, R. C. (1980). Protocols for public key cryptosystems. *IEEE Symposium on Security and Privacy* (str. 122-134). IEEE. Preuzeto 6. 7. 2019. s https://www.researchgate.net/profile/Ralph_Merkle/publication/220713913_Protocols_for_Public_Key_Cryptosystems/links/00b495384ecda07784000000/Protocols-for-Public-Key-Cryptosystems.pdf
[285] Merkle, R. C. (1982). *Washington, DC Br. patenta U.S. Patent No. 4,309,569.* Preuzeto 8. 12. 2021. s https://worldwide.espacenet.com/patent/search/family/022107098/publication/US4309569A?q=pn%3DUS4309569

Svrha Merkleovog izuma nije izravno povezna s pohranom podatka. Ona je metoda koja olakšava, to jest eliminira potrebu za distribucijom velikog broja javnih ključeva pri potpisivanju velikog broja zapisa upotrebom jednokratnih metoda za potpisivanje zapisa, poput Lamportovog potpisa.[286]

U kombinaciji s jednokratnom potpisnom metodologijom, poput Lamportove, Merkleovo stablo omogućava da se, nakon što je generirano *n* parova privatnih i javnih ključeva, gdje je *n* broj čvorova na najnižoj razini i ukupan broj zapisa koji će se potpisati, za kasniju potvrdu ispravnosti potpisa koristi korijenski hash (kao javni ključ svih zapisa). Na ovaj način postoji potreba za distribucijom samo jednog javnog ključa, umjesto *n* javnih ključeva. Ova osobina karakteristična za postupke potpisivanja jednokratnim ključevima i nije primjenjiva na zapise potpisane (dugo)trajnim digitalnim certifikatima poput onih temeljenih na DSA ili RSA kriptografskim sustavima. Usprkos ovome, stabla hasheva su našla široku primjenu u suvremenim sustavima za vremenske žigove. U slučaju ovakve upotrebe stabla hasheva se obično ulančavaju na način koji su predložili Haber i Stornetta.[287] Ovakva podatkovna struktura prikazana je na slici 18.



Slika 18. Ulančana hash stabla

Imajući sve navedeno na umu ne predlažem upotrebu punog hash stabla u TrustChain modelu. Osnovni razlog ovome je veličina hash stabla. Ako na najnižoj razini kompletnog binarnog stabla imamo 1024 zapisa kompletno binarno stablo zahtjeva 10 razina jer je broj

[286] Lamport, L. (1981). Password authentication with insecure communication. *Communications of the ACM, 24*(11), 770-772. Preuzeto 7. 12. 2021. s http://merlot.usc.edu/cs530-s07/papers/Lamport81a.pdf
[287] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.

čvorova na najnižoj razini jednak 2^h, ako je *h* visina stabla. Ovo znači da iznad razine s početnim hashevima treba pohraniti još skoro toliko hasheva. Ako je *t* ukupni broj čvorova, a *h* visina stabla onda u binarnom stablu vrijedi.[288]

$$t = 2^{h+1} + 1$$

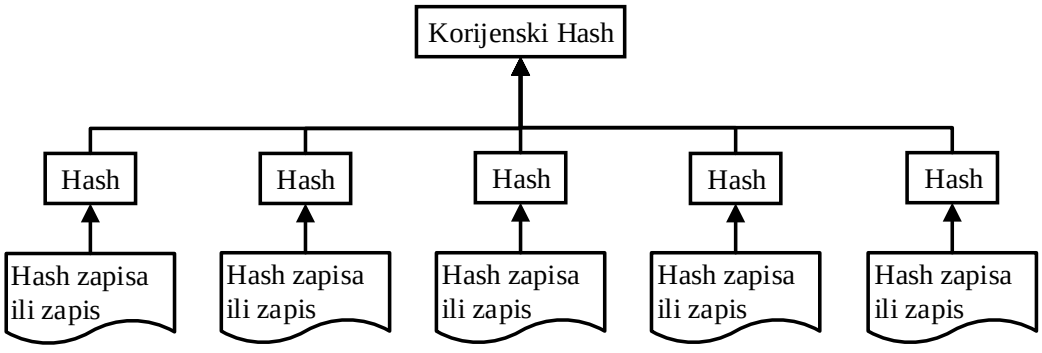
Dakle, potrebno je dodati 1023 čvora koji sadrže dodatne hasheve. Ukupan broj hasheva u takvom stablu je 2047. S obzirom na nedostatak ranije opisane osobine kombiniranja javnih ključeva i na to da je jedina funkcija hasheva u TrustChain modelu zaštita integriteta podataka, za TrustChain model, odabrana je jednostavnija strukturu, lista hasheva.

4.3. Lista hasheva

Lista hasheva je podvrsta Merkleovog stabla[289]. U svojoj osnovi riječ je o jednostavnijem stablu hasheva. Listu hasheva karakteriziraju dvije razlike u odnosu na binarno Mrekleovo stablo objašnjeno u prethodnom poglavlju. Ove razlike su:

1) stablo ima samo dvije razine,
2) stablo nije binarno.

Kada uvažimo ove dvije razlike dobivamo strukturu koja je sličnija listi nego stablu. U ovom smislu ona je slična hash tablicama za direktno adresiranje[290], kakve su kratko predstavljene u ranijem poglavlju. Ipak, s obzirom na to da ova lista hasheva sadrži korijenski hash, koji se računa iz konkatenacije svih hasheva u listi (razina stabla 1), može se reći da je riječ o nebinarnom stablu s dvije razine. Ovakva podatkovna struktura prikazana je na slici 19.



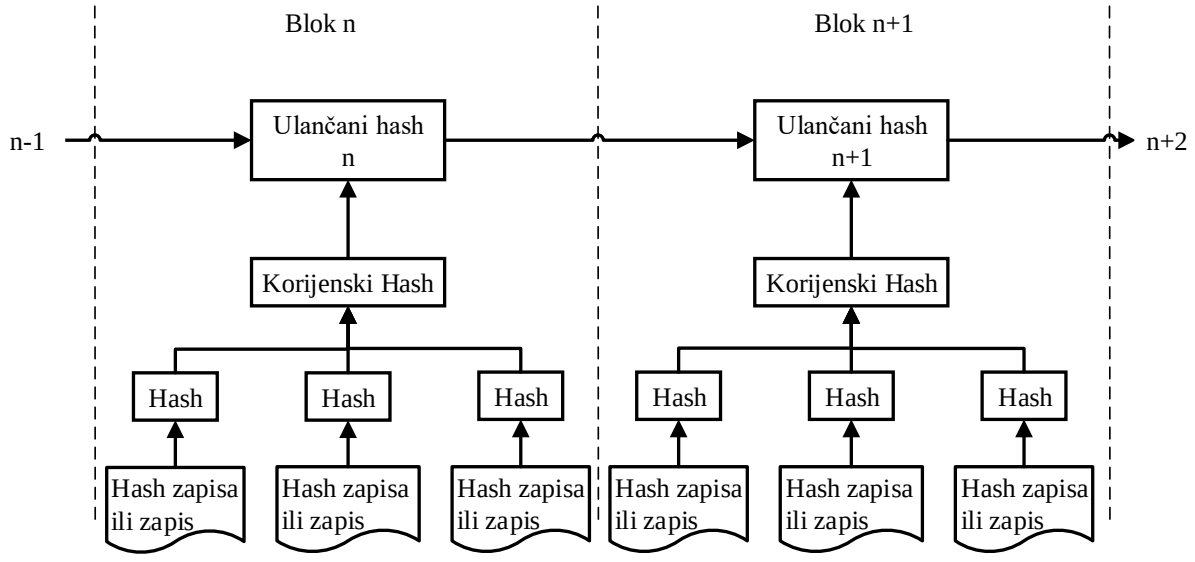
Slika 19. Lista hasheva

[288] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms*, n. dj.
[289] Merkle, R. C. (1980). Protocols for public key cryptosystems, n. dj.
[290] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms*, n. dj.

Ova podatkovna struktura izuzetno je važna za provedeno istraživanje jer čini temelj podatkovne strukture ulančanih blokova koja se koristi u TrustChain modelu koji je razrađen u kasnijem poglavlju.

Smatram da upotreba punog Merkleovog stabla nije prikladna podatkovna struktura za TrustChain lanac blokova. Ovo mišljenje temelji se na činjenici da TrustChain lanac blokova ne sadrži samo hasheve već i neke metapodatke vezane uz svaki zapis. Merkleovo stablo je učinkovit način za provjeru integriteta velikog skupa hash vrijednosti. Ovakva stabla dizajnirana su za upotrebu u sustavima koji potpisuju veliki broj podataka, kao rješenje za problem distribucije velikog broja javnih ključeva.[291,292] Iako je više sustava za vremenske žigove, na primjer GuardTime,[293] implementiralo sličnu, binarnu podatkovnu strukturu, s obzirom na to na da TrustChain lanac blokova sadrži složenu podatkovnu strukturu (binarni JSON zapis) dodavanje punog binarnog stabla iznad tih podatka stvara nepotreban teret. Integritet podataka je osiguran i upotrebom liste hasheva s korijenskim hashem.

Liste hasheva u TrustChain modelu se povezuju u ulančane blokove sukladno principu opisanom u radu Haber i Stornette.[294] Par ulančanih listi hasheva prikazan je na slici 20.



Slika 20. Ulančane liste hasheva

[291] Merkle, R. C. (1980). Protocols for public key cryptosystems, n. dj.
[292] Becker, G. (2008). *Merkle signature schemes, merkle trees and their cryptanalysis.* Bocuhm: Ruhr-Universit¨at Bochum. Preuzeto 7. 12. 2021. s https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.392.7879&rep=rep1&type=pdf
[293] GuardTime AS. (2021). *GuardTime*. Preuzeto 3. 12. 2021. s KSI blockchain timestamping: https://guardtime.com/timestamping
[294] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n.dj.

Hashevi zapisa (ili zapisi) prikazani na dnu slike 20 se u slučaju TrustChain modela odnose na binarne JSON zapise koji su detaljno prikazani u poglavlju koje predstavlja novi model. Ovdje prikazana podatkovna strukturama ulančanih listi hasheva predstavlja teoretsku osnovu za nepromjenjivu podatkovnu strukturu koju novi model koristi.

4.4. Lanac hasheva

Lanac hasheva (engl. *hash chain*) po svojoj izvornoj definiciji nije podatkovna struktura već proces kojim se pokušava povećati sigurnost autentifikacije u nesigurnim komunikacijama. U ovom smislu poglavlje se možda ne čini važnim ali je ono ipak uključeno ovdje iz dva razloga:

1) terminologija je vrlo slična podatkovnim strukturama (engl. *hash chain* i engl. *block chain*, lanac hasheva i lanac blokova) o kojima se ranije govorilo,
2) koncept na kojem se temelji termin lanaca hasheva vrlo je sličan kasnije opisanoj tehnici naknadnog potpisivanja zapisa (najčešće vremenskim žigovima).

Zbog ovih razloga je termin lanac hasheva dobro (ukratko) razjasniti da bi se izbjegli nesporazumi, a on u svojoj osnovi i jest ekvivalentan naknadnom potpisivanju zapisa. Iako mu to nije bila svrha koncept je identičan dodavanju vremenskih žigova i može služiti kao opis, to jest uvod u taj koncept. Osim toga, riječ je o specifičnoj vrsti ulančanih zapisa koji nisu ekvivalentni podatkovnoj strukturi koja je centralna ovom radu, lancu blokova (engl. *blockchain*).

Ideja lanca hasheva prvi puta se spominje 1981. u kratkom radu Lamporta.[295] Lamport iznosi ideju koja pomaže zaštititi lozinke tijekom prijenosa nezaštićenim računalnim mrežama, kao što je to Internet. Lamport je predložio da se lozinke modificiraju na način da se pri svakoj autentifikaciji koristi druga lozinka. S obzirom da bi to značilo da obje strane moraju održavati velike liste lozinki za svakog korisnika Lamport predlaže da se pohranjuje samo jedna lozinka za svakog korisnika ali se pri *n-tom* pokušaju autentifikacije koristi lozinka koja je *n* puta hashirana unaprijed dogovorenom funkcijom. To jest ako je lozinka *x*, dogovorena funkcija *F*, a hash lozinke *y* onda vrijedi:

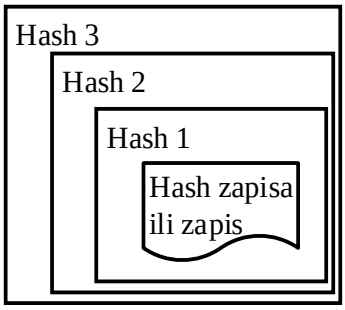
$$y_n = F_n(\ldots.(F_2(F_1(x))))$$

[295] Lamport, L. (1981). Password authentication, n. dj.

Na ovaj način, svaki put kada se korisnik autentificira preko nezaštićene mreže će putovati drukčiji hash lozinke (*y*).

Iako ovdje nije riječ o striktno podatkovnoj strukturi, ovaj koncept (uzastopnog hashiranja) gotovo je identičan postupku uzastopnog potpisivanja nekog zapisa vremenskim žigovima ili certifikatima, koji rezultira podatkovnom strukturom koja prema saznanjima autora disertacije nema poseban naziv (ali je vrlo slična ovdje objašnjenom lancu hasheva). Iz ovog razloga, kao i zbog razjašnjavanja terminologije, ovaj koncept je odabran kao temelj za objašnjenje te neimenovane, ali često korištene, podatkovne strukture.

Ovakav lanac vremenskih žigova ili potpisa rezultira dodavanjem omotnica oko izvornog zapisa i uobičajen je postupak u digitalnim arhivima koji ima za cilj dugotrajno očuvanje integriteta zapisa. Ovako strukturirani hashevi (ili vremenski žigovi ili digitalni potpisi) prikazani su na slici 21 i tema su idućeg poglavlja.



Slika 21. Lanac (engl. chain) hasheva

Ova podatkovna struktura nije kandidat za upotrebu u novom modelu. TrustChain model nema ni potrebu, niti mogućnost, za naknadno potpisivanje vlastitih zapisa. Ipak, na ovaj način omotani podaci predstavljaju podatkovnu strukturu koja je relevantna za istraživanje jer postojeća rješenja rezultiraju podacima koji su organizirani na ovaj način te ona predstavlja alternativu upotrebi TrustChain modela.

4.5. Zaključak

Poglavlje je objasnilo temeljene podatkovne strukture koje koristi novi TrustChain model i koje koriste gotova, trenutno dostupna rješenja, prije svega sustavi za vremenske žigove. Između raspravljenih struktura ulančani zapisi i lista hasheva su za predstavljeni model daleko najvažnije. To su podatkovne strukture koje predstavljaju srž modela.

Iz izloženog u ovom poglavlju i detaljnog razmatranja hash funkcija u prethodnom jasno je da podatkovna struktura ulančanih zapisa omogućuje dokazivost integriteta podataka od trenutka stvaranja ulančanog zapisa. S obzirom na to sigurno je da ova podatkovna struktura ispunjava uvjete za temeljnu podatkovnu strukturu novog modela za dugotrajnu pohranu podatka.

Osim ulančanih zapisa u poglavlju su raspravljene i druge vrste podatkovnih struktura koje se vežu uz pohranu hasheva te su relevantne za istraživanje. Ove podatkovne strukture su kandidati za osnovnu strukturu podataka unutar ulančanog bloka. Za primjenu u novom modelu odabrana je lista hasheva umjesto stabla hasheva. Pokazno je da hash stablo, u ovom slučaju primjene, ne pruža jasne prednosti, a povećava memorijske zahtjeve ulančanih blokova pa je kao temelj strukture organizacije podataka unutar ulančanih blokova odabrana jednostavnija podatkovna struktura - lista hasheva.

Ulančani zapisi, u kontekstu upotrebe povezanih hasheva odavno su poznata tehnika osiguravanja zapisa, što je i bio jedan od razloga za njihov odabir kao temeljne podatkovne strukture novog modela. Iako je istraživanje u svojim prvim koracima počelo od pretpostavke da su za ovu svrhu pogodni ulančani zapisi koji se inače koriste u druge svrhe poput Bitcoin i Ethereum lanca blokova, oni su se brzo pokazali kao sustavi koji ili ne ispunjavaju zahtjeve u potpunosti, ili stvaraju nepotrebne troškove sustavu (a ponekad i oboje). Ovo je donekle u skladu i s drugim istraživanjima u ovom području. Lemieux je 2016. proučila prijedlog implementacije Bitcoin lanca blokova te na njemu izgrađenog Factom[296] podatkovnog sloja[297] kao temelja za sustav za pohranu katastarskih zapisa Hondurasa.[298] Lemieux u zaključku predlaže oprez, ali i naglašava prednosti upotrebe javnog lanca blokova kao validacijskog mehanizma. U kasnijem radu Lemieux razrađuje topologiju tada dostupnih rješenja za pohranu zapisa temeljenih na ulančanim zapisima.[299] Razvijena topologija sastoji se tri vrste sustava koji se razlikuju po razini integracije ulančanih blokova:

---

[296] Factom URL: https://www.linkedin.com/company/factom-protocol/

[297] Snow, P., Deery, B., Kirby, P., & Johnston, D. (2015). *Factom ledger by consensus.* Preuzeto 9. 12 2021 s https://cryptochainuni.com/wp-content/uploads/Factom-Ledger-by-Consensus.pdf

[298] Lemieux, V. L. (2016). Trusting records: is Blockchain technology the answer? *Records Management Journal, Vol. 26 Issue: 2*, 110-139. Preuzeto 7. 1. 2022. s https://www.emerald.com/insight/content/doi/10.1108/RMJ-12-2015-0042/full/html

[299] Lemieux, V. L. (2017). A typology of blockchain recordkeeping solutions and some reflections on their implications for the future of archival preservation. *2017 IEEE International Conference on Big Data (Big Data)* (str. 2271-2278). IEEE. Preuzeto 3. 12. 2021. s https://www.researchgate.net/profile/Victoria-Lemieux/publication/322511343_A_typology_of_blockchain_recordkeeping_solutions_and_some_reflections_on_their_implications_for_the_future_of_archival_preservation/links/5a9626ed45851535bcdcc1f4/A-typology-of-bl

1) Zrcalni tip (engl. *mirror type*). Ova vrsta sustava koristi ulančane blokove samo za zapis hasheva arhivskih zapisa ili dokumenata. Prema Lemiux primjer ovakvog sustava je Guardtime[300] sustav.

2) Tip (punog) digitalnog zapisa (engl. *digital record type*). Sustavi ove vrste u ulančane blokove pohranjuju cijeli zapis. Lemiux kao primjer ovakvog sustava navodi Švedski sustav za prijenos vlasništva nad nekretninama koji se temelji na ChromaWayevim Esplix[301] i Postchain[302] tehnologijama.

3) Znakovni tip (engl. *token type*). Znakovni sustavi u potpunosti su integrirani s ulančanim zapisima te osim pohrane zapisa u ulančanim blokovima počinju koristiti i znakove (engl. token) za monetizaciju sustava ili pohranu dodatnih podataka. Primjer ovakvog sustava je Brazilski katastar temeljen na Ubitquityevom[303] sustavu koji koristi obojene kovanice (engl. *coloured coins*) za označavanje vlasništva.

TrustChain model se prema ovoj topologiji jasno svrstava u zrcalni tip sustava temeljen na vlastitom lancu blokova. Iako TrustChain model ne koristi ulančane zapise isključivo za pohranu hasheva, već i za pohranu djelomičnih metapodataka, sami (arhivski) zapisi se tamo ne pohranjuju.

Spomenuta prethodna istraživanja, kao i postojeći sustavi, posebno oni za vremenske žigove te istraživanja vezna uz njih, jasno ukazuju na pogodnost prikazanih podatkovnih struktura za primjenu u sustavima kojima je cilj dugotrajno očuvanje autentičnosti i integriteta digitalnih zapisa. Istraživanja u ovom području se nastavljaju te na temelju posebnog izdanja časopisa Computers koji je posvećen ulančanim blokovima i očuvanju zapisa Lemiux, urednica tog izdanja, zaključuje da je razina istraženosti upotrebe ulančanih zapisa za dugotrajno očuvanje arhivskih zapisa i dalje relativno niska te da treba nastaviti s istraživanjima.[304] Novi TrustChain model predstavlja nastavak istraživanja upravo u tom pravcu.

---

[300] GuardTime AS. (2021). *GuardTime*, n. dj.
[301] ChromaWay. (2021). *ChromaWay Technology*. Preuzeto 9. 12. 2021. s ChromaWay : https://chromaway.com/technology
[302] ChromaWay. (2017). *Postchain*. Preuzeto 9. 12. 2021. s Postchain: https://postchain-docs.readthedocs.io/en/latest/
[303] Ubitquity URL: https://www.ubitquity.io/
[304] Lemieux, V. L. (2021). Blockchain and Recordkeeping: Editorial. *Computers, 10*, 135. doi: https://doi.org/10.3390/computers10110135

5. Postojeća rješenja za dugotrajno dokazivanje integriteta podataka

Ovo poglavlje daj pregled postojećih standarda i već gotovih rješenja problema dugotrajnog očuvanja digitalno potpisanih zapisa. Uobičajeni odgovor na pitanje: "Kako dugotrajno zajamčiti očuvanje autentičnosti i integriteta digitalnog zapis?" je digitalni vremenski žig (engl. *time–stamp*). Digitalni vremenski žigovi prisutni su u informacijskim tehnologijama i informacijskim znanostima još od 1999., kada su objašnjeni prvi učinkoviti digitalni vremenski žigovi,[305] te od tada nalaze primjenu u raznim sustavima za digitalne vremenske žigove poput GuardTime[306], AbsolutProof,[307] time:beat[308] i drugih. Poglavlje nije zamišljeno kao iscrpni pregled ovih sustava već daje prikaz nekolicine sustava koje smatram karakterističnim za potvrdu postavljenih teza, relevantnim zbog primjera tehnologija koje koriste ili pogodnim kandidatima za pružanje usluge vanjskog vremenskog žiga za novi model.

Digitalni vremenski žig je oznaka vremena koja je trajno povezana s nekim digitalnim zapisom. Adams i drugi autori u dokumentu RFC3161[309] daju definiciju servisa za izdavanje vremenskih žigova:

> *"Servis za dodavanje vremenskih žigova podržava dokaz da je podatak postojao prije konkretnog vremena."* (Adams, Cain, Pinkas, & Zuccherato, 2001)[310]

Definicija to eksplicitno ne navodi ali postojanje podatka implicira i konkretni sadržaj podataka. Vremenski žigovi omogućuju dokazivanje integriteta podataka u danom trenutku. Stoga, definiciju ima smisla proširiti na način:

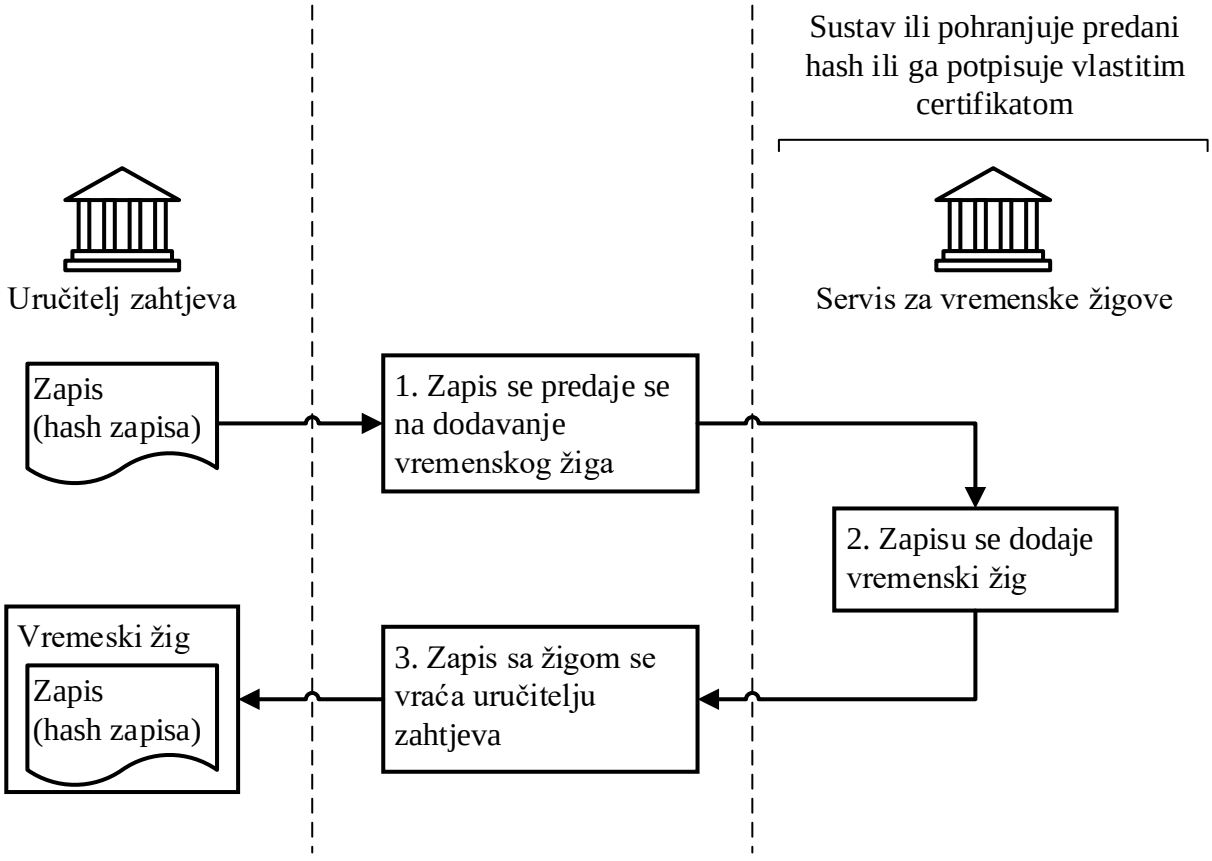
Servis za dodavanje vremenskih žigova je informacijski sustav koji podržava dokaz da je podatak ili skupina podataka postojala prije konkretnog vremena u točno određenom obliku.

[305] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.
[306] GuardTime URL: https://guardtime.com/timestamping
[307] AbosluteProof URL: http://www.surety.com/solutions/intellectual-property-protection/sign-seal
[308] time:beat URL: https://enigio.com/timebeat/
[309] Adams, C., Cain, P., Pinkas, D., & Zuccherato, R. (2001). *RFC 3161, n. dj.*
[310] Ibid.

Te je iz nje moguće izvesti definiciju samog vremenskog žiga:

Vremenski žig je skup podatka (često u obliku certifikata) koji pruža dokaz da je neki podatak postojao prije konkretnog vremena u točno određenom obliku.

Da bi ovo postigao digitalni vremenski žig mora na neki način dokazati da je zapis kojem se dodaje zaista i postojao u trenutku nastajanja žiga te da od tog trenutka nije mijenjan. Za postizanje ovog cilja većina sustava poseže za kriptografskim rješenjima koja se temelje na izvornom radu o vremenskim žigovima u kojem je osnovni element ovakvog sustava hash funkcija[311]. U svojoj osnovi, kada govorimo o dodavanju vremenskih žigova podacima izvan specifičnih sustava (poput raznih dnevnika događaja u informacijskim sustavima) možemo reći da on, u najjednostavnijoj varijanti, prati korake opisane na slici 22.



Slika 22. Jednostavni postupak dodavanja vremenskog žiga na zapis

Ovaj princip opisali su i Haber i Stornetta te ga nazvali naivnim, i iskoristili kao argumentaciju i motivaciju za razvoj novih modela za dodavanje vremenskih žigova. Najveći problem s ovakvim sustavom je da je on izuzetno ranjiv na jednoj točki – servisu za vremenske

[311] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.

žigove. U slučaju da je ovaj servis pristran, zlonamjeran, ili kompromitiran svi žigovi koje je izdao postaju nevaljani, i u tom slučaju naknadno dodavanje vremenskog žiga ne postiže potreban efekt jer je dokaz o vremenu kada je vremenski žig stvarno bio potreban izgubljeno. Iz ovog razloga svi sustavi koji se bave dugotrajnim očuvanjem integriteta i autentičnosti zapisa moraju na neki način izbjeći oslanjanje na jedan autoritet.

Kao što je slučaj i s digitalnim potpisom, digitalni vremenski žig izvorno je namijenjen digitalnim zapisima koji još nisu bili pohranjivani u digitalne arhive ili je istek digitalnih certifikata bio zanemarivan jer su bili namijenjeni računalnim protokolima u kojima je ovaj istek poželjan (na primjer SSL i TSL protokoli). Iz tih razloga većina izvornih sustava nije u potpunosti adekvatna za upotrebu u arhivske svrhe. Osim što ne zadovoljavaju specifične arhivske zahtjeve ovi sustavi često zanemaruju digitalne certifikate i probleme koji proizlaze iz njihovog isteka pri čuvanju zapisa na (izuzetno) dug ili čak neograničen vremenski period. Ipak, razvijen je značajan broj industrijskih rješenja koja prate vrlo detaljne standarde i imaju za cilj arhiviranje podatka, iako uglavnom pokušavaju biti univerzalni i primjenjivi na svu pohranu podatka, pa ih ovaj rad mora razmotriti.

Osim "klasičnih" vremenskih žigova ovo poglavlje će raspraviti i noviji koncept temeljen na prolaznim ključevima koji je tek djelomično zaživio i najjednostavnije je rješenje problema za naknadno periodičko potpisivanje zapisa nakon što su oni ušli u arhiv.

Sustavi za vremenske žigove i prateći standardi su od izuzetne važnosti za ovaj rad iz još jednog razloga. Značajni dio ovih sustava bazira se na ulančanim blokovima (engl. *blockchain*) podataka što je temelj arhitekture i modela koji su rezultat ovog istraživanja. Dapače, vremenski žigovi su jedno od prvih područja primjene ulančanih blokova uopće. Vremenski žig koji koristi ulančane blokove koji su 1990. opisali Haber i Stornetta još uvijek je osnova ne samo suvremenih sustava za vremenske žigove već i ostalih sustava koji koriste ulančane blokove.[312] Najpoznatiji takvi sustavi sigurno su kriptovalute i izvorni rad Nakamota, koji je razvio ideju Bitcoin sustava, koji se poziva na ovaj sustav vremenskih žigova.[313]

[312] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.
[313] Nakamoto, S. (2008). *Bitcoin*, n. dj.

Upotreba ulančanih zapisa u svrhe sustava za vremenske žigove je danas standardizirana u ANSI X9.95[314] i ISO 18014[315] standardima. Više dostupnih komercijalnih rješenja usklađeno je s ovim standardima te će neka od njih biti razmotrena u idućim poglavljima.

Zbog toga, oslanjanja na ulančane blokove, postojeći sustavi za vremenske žigove doprinose potvrdi prve hipoteze: "Upotrebom ulančanih zapisa produžuje se dokazivost autentičnosti digitalno potpisanoga arhivskoga gradiva". Iako ovi sustavi izvorno nisu namijenjeni provjeri autentičnosti (upotrebom digitalnih certifikata) i trude se očuvati samo integritet podataka oni su i dalje (djelomična) potvrda prve hipoteze.

5.1. Vremenski žigovi temeljeni na ulančanim zapisima i PKI sustavima

Vremenski žigovi opisani u radu Habera i Stornette, koji su osnova svih suvremenih sustava za vremenske žigove, kao i modela opisanog u ovoj disertaciji, postoje u dvije varijante. [316] Obje varijante se oslanjaju na nepromjenjivi lanac zapisa koji je osnova suvremenih ulančanih zapisa. Haber i Stornetta ove varijante nazivaju "povezana shema " (engl. *linking scheme*) i "shema distribuiranog povjerenja" (engl. *distributed trust scheme*). Obje sheme pretpostavljaju da postoji niz zahtjeva za izdavanjem vremenskih žigova koji se predaju nekom servisu, odnosno TSS-u (engl. *Time Stamping Service*). Ove zahtjeve uručuju razne pravne ili fizičke osobne, a točan redoslijed zahtjeva niti njihov sadržaj nije unaprijed poznat. Ova pretpostavka je srž sustava koji su predložili Haber i Storentta i ona je očigledno istinita te vrijedi za sve sustave koji opslužuju veliki broj klijenata. Iako se obje sheme oslanjanju na jednosmjerne hash funkcije one se razlikuju po načinu na koji koriste ovu činjenicu da bi ispunile ranije navedenu definiciju vremenskog žiga. S obzirom da tvore temelj svih suvremenih sustava za vremenske žigove obje sheme će biti razmotrene:

[314] ANSI. (2016). *ANSI X9.95-2016 Financial Services – Trusted Time Stamp Management And Security.* Preuzeto 19. 11. 2022. s American National Standards Institute : https://infostore.saiglobal.com/en-gb/Standards/ANSI-X9-95-2016-1894464/

[315] International Organization for Standardization. (2009). Preuzeto 19. 11. 2022. s ISO/IEC 18014-3:2009 Information technology -- Security techniques -- Time-stamping services -- Part 3: Mechanisms producing linked tokens: https://www.iso.org/standard/50457.html

[316] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.

1) Povezana shema vremenskih žigova.

   U ovoj shemi autori[317] su predložili povezivanje zapisa na način da ako je uručitelj zahtjeva predao podatke *y* TSS odgovara s C_n koji se formira na način:

$$C_n = (n, t_n, ID_n, y_n, L_n)$$

   Gdje je *n* redni broj, t_n vrijeme zahtjeva, ID_n identifikator uručitelja zahtjeva, y_n predani podaci, a L_n hash podataka o prethodnom zapisu to jest:

$$L_n = (t_{n-1}, ID_{n-1}, y_{n-1}, H(L_{n-1}))$$

   Ovo je upravo definicija ulančanog bloka na način na koji je opisan u prethodnom poglavlju i koja je aktualna i danas, 30 godina nakon njenog prvog opisa. Ona pretpostavlja centralni servis, TSS, koji opslužuje sve ove zahtjeve i čuva kopiju svih vremenskih žigova (koje bi trebali čuvati i uručitelji zahtjeva). Prema ovom opisu sustava svaki zapis je ekvivalentan jednom ulančanom bloku. Suvremeni sustavi, koji koriste ulančane blokove ovo izbjegavaju upotrebom stabla hasheva.[318] I autori ove sheme u svojem kasnijem radu s Byerom preporučuju upotrebu stabla hasheva.[319] Ovu shemu nije potrebno dodatno komentirati jer je već detaljno objašnjena u ranijem poglavlju. TrustChain model, koji je prikazan u kasnijem poglavlju, se uvelike oslanja na ovaj princip (ulančanih zapisa).

2) Shema distribuiranog povjerenja za vremenske žigove.

   Prema autorima[320] ova shema pretpostavlja upotrebu pseudonasumičnog generatora koji na osnovu podatka *y* generira n-torku:

$$G(y) = (ID_1, ID_2, ID_3, \dots, ID_n)$$

   ID_n vrijednosti u ovoj n-torci odgovaraju identifikatorima drugih pravnih ili fizičkih osoba koji sudjeluju u distribuiranom (decentraliziranom) sustavu za izdavanje vremenskih žigova. Uručitelj zahtjeva za vremenskim žigom šalje

[317] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.
[318] Merkle, R. C. (1987). A digital signature based on a conventional encryption function. *Conference on the theory and application of cryptographic techniques* (str. 369-378). Berlin, Heidelberg: Springer. Preuzeto 1. 12. 2021. s https://link.springer.com/content/pdf/10.1007/3-540-48184-2_32.pdf
[319] Bayer, D., Haber, S., & Stornetta, W. (1993). Improving the efficiency and reliability of digital time-stamping. *Sequences II* (str. 329-334). New York, NY: Springer. Preuzeto 1. 12. 2021. s http://www.math.columbia.edu/~bayer/papers/Timestamp_BHS93.pdf
[320] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.

svim ovim osobama (ili ustanovama) svoje podatke (najčešće hash podataka) i zauzvrat od svakog dobiva potpisanu poruku:

$$s_n = (t, ID, y)$$

Kombiniranjem svih ovih poruka uručitelj zahtjeva dobiva vremenski žig u obliku n-torke koja se sastoji od:

$$[(y, ID), (s_1, \dots, s_n)]$$

Pri provjeri ispravnosti žiga potrebno je kontaktirati sve (ili barem većinu) ustanova ili osoba koje su sudjelovale u stvaranju istog. Sigurnost ovakvog sustava temelji se na pseudonasumičnom generatoru koji za svaki konkretni hash podataka *y* generira specifičan set identifikatora osoba ili ustanova koje sudjeluju u sustavu. Na ovaj način nije moguće predvidjeti tko će izdavati podatke za koji vremenski žig. Haber i Stornetta predlažu generatore temeljene na rješenjima autora Blum i Micali[321] ili Yao.[322] Ovakvo rešenje, prema spoznajama autora disertacije, nije izravno implementirano u otvorenim sustavima za vremenske žigove. Rješenje ima ogromnu prednost izbjegavanja oslanjanja na centralni TSS, ali zahtjeva da u sustavu (trajno) sudjeluje veliki broj osoba ili ustanova koje će koristiti specijalizirani softver koji generira ranije spomenute n-torke i koji će čuvati sve podatke o svim vremenskim žigovima u kojima su sudjelovali. Model predstavljen u kasnijim poglavljima (TrustChain) ipak primjenjuje ovu shemu kao načelo distribuiranog povjerenja kroz model glasanja o ispravnosti zapisa. Na sličan način načelo su prihvatile i suvremene kriptovalute, poput Bitcoina, te drugi sustavi koji su inače temeljeni na prvom principu (ulančanih blokova), poput Ethereuma, kroz inzistiranje na distribuiranim sustavima umjesto polaganja povjerenja u centralnu instituciju. U kasnijem radu istih autora[323] naziv ove sheme je promijenjen u shemu nasumičnih svjedoka (engl. *random witness*), naziv pod kojim je i danas poznata.

[321] Blum, M., & Micali, S. (1984). How to generate cryptographically strong sequences of pseudorandom bits. *SIAM journal on Computing, 13*(4), 850-864. Preuzeto 1. 12 2021 s https://dl.acm.org/doi/abs/10.1145/3335741.3335751

[322] Yao, A. C. (1982). Theory and application of trapdoor functions. *3rd Annual Symposium on Foundations of Computer Science (SFCS 1982)* (str. 80-91). IEEE. doi: https://doi.org/10.1109/SFCS.1982.45

[323] Bayer, D., Haber, S., & Stornetta, W. (1993). Improving the efficiency, n. dj.

Obje sheme su u kasnijem radu istih autora proširene dodavanjem koncepta stabla hasheva te autori smatraju da sve tri tehnike mogu biti korištene istovremeno da bi omogućile najvišu razinu sigurnosti sustava za vremenske žigove. I zaista, kada govorimo o vremenskim žigovima najčešće je riječ o sustavima temeljenim upravo na gore navedenim shemama. Primjer ovakvog sustava, koji je temeljen upravo na radu Habera i Stornette, je AbsoluteProof, sustav za vremenske žigove koji je tvrtka Surety[324] razvila 1995. Sustav je temeljito analiziran u tehničkom izvještaju Thimblina i suradnika iz 2005.[325]

Drugi sustavi su osnovni koncept proširivali na razne načine. time:beat, sustav tvrtke Enigio Time [326] koristi nov (patentiran) koncept nasumičnih događaja [327] kojim formira negativni vremenski žig (engl. *negative time stamp*). Prema ranijoj definiciji vremenski žig garantira da je zapis postojao u određenom trenutku u vremenu, to jest da je stvoren najkasnije tada. Sustav time:beat, upotrebom negativnog vremenskog žiga, pokušava dokazati da vremenski žig nije postojao prije tog trenutka. Dokaz da sam zapis nije postojao prije trenutka dodavanja vremenskog žiga nije moguć, ali na ovaj način ipak se izbjegava određena mogućnost manipulacije podacima od strane zlonamjernog TSS-a. Sustav ovo postiže na način da u svaki vremenski žig uključuje "signal" koji je rezultat kriptografske hash funkcije koji je izračunat na osnovu javno dostupnih podatka koji su nasumično odabrani te ih je teško ili nemoguće (precizno) predvidjeti. Ti podaci mogu biti rezultati sportskih događaja, vremenske prognoze, rezultati igara na sreću, kretanja cijena na burzi, broja munja u svijetu u nekom trenutku i sl., tj. bilo koji javno objavljeni podatak koji nije moguće u potpunosti predvidjeti. Na primjer, moguće je s relativnom razinom pouzdanosti predvidjeti pobjednika neke nogometne utakmice, jer se jedna momčad smatra značajno kvalitetnijom. Dobar analitičar možda ima i ideju koji je vjerojatan broj pogodaka za obje momčadi. Ono što nije moguće predvidjeti, a javno je objavljen podatak, je točno vrijeme tih pogodaka. Kada je u vremenski žig uključen ovakav podatak on se može iskoristiti kao snažan dokaz da vremenski žig nije postojao prije određenog vremena. Sustav time:beat dodatno osnažuje dokaz na način da ne koristi samo jedan već više podataka iz nasumično odabranih izvora za svaki vremenski žig.

[324] Surety, I. (2021). *AbsoluteProof*. Preuzeto 2. 12. 2021. s Surety: http://www.surety.com/
[325] Thimblin, M., Kamisetty, N., Raman, P., & Paila, A. (2005). *Implementation of an Evidentiary Record Validation Utility and Security Analysis for Surety's AbsoluteProof.* George Mason University. Preuzeto 3. 12. 2005. s http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.136.7029&rep=rep1&type=pdf
[326] Enigio Time AB. (2021). *Enigio*. Preuzeto 2. 12. 2021. s time:beat: https://enigio.com/timebeat/
[327] Almgren, H. (2015). *Br. patenta WO 2015/020599 Al.* Preuzeto 1. 12. 2021. s https://patentimages.storage.googleapis.com/7a/25/e6/d6ba657b93aa99/WO2015020599A1.pdf

Takav koncept Enigio naziva konceptom događaja kojima su svjedočili mnogi (engl. *widely-witnessed event*).

Enigiov sustav, koji koristi vlastiti princip nasumičnih događaja, sigurno je inspiriran ranije opisanom shemom nasumičnih svjedoka (distribuiranog povjerenja) ali ju je dodatno razradio. Također, time:beat, kao što je predloženo u izvornim modelima, povremeno objavljuje hash trenutnog bloka nepromjenjivog lanca blokova (sukladno shemi povezanih vremenskih žigova) na medijima nad kojima nema kontrolu i tako pruža dodatno osiguranje da Enigio, kao tvrtka koja upravlja sustavom, nije u stanju samovoljno mijenjati zapise. I ovo je koncept predstavljen u radu Habera i Stornette.[328] Ovi, nezavisni, mediji su često novine koje se izdaju u fizičkom, a ne samo digitalnom obliku ili neki drugi medij koji ostavlja fizički trag.

5.2. Vremenski žigovi temeljeni na prolaznim ključevima

Osim ranije opisanih "klasičnih" sustava za vremenske žigove Doyle[329] je 1998. godine razvio poseban sustav koji se ne oslanja na PKI sustave. Ovakav sustav pokušao je realizirati Proofspace Inc. koji je na osnovu Doyleovog patenta razvio i objavio detaljni opis sustava za vremenske žigove Proofmark.[330]

Doylovea ideja polazi od pretpostavke da su certifikati koji se koriste u većini sustava za vremenske žigove najslabija karika takvog sustava. U slučaju gubitka certifikata potpisani dokumenti postaju kompromitirani, to jest nevjerodostojni te se stvara situacija u kojoj se takvi potpisani zapisi moraju povući. S obzirom na to da je stvaranje žiga u konkretnom trenutku ujedno i svrha vremenskog žiga, u slučaju kada dolazi do njegovog povlačenja stvara se ogromna (možda i nepopravljiva) šteta. O ovoj problematici ranije je pisao autor ove disertacije (s koautorima).[331] Autori su u radu naglasili probleme koji proizlaze iz povlačenja certifikata što je preporuka relevantnih standarda u slučaju kompromitiranja certifikata.[332] Konkretno, relevantni RFC navodi:

[328] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.
[329] Doyle, M. D. (1998). *Sjedninje Američke Države Br. patenta US6381696B1.* Preuzeto 3. 12. 2021. s https://patents.google.com/patent/US6381696B1/en
[330] Proofspace. (2007). *Proofmark System Technical Overview.* ProofSpace. Preuzeto 3. 12. 2021. s http://fios.com/proofmarksystemtech.pdf
[331] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.
[332] Adams, C., Cain, P., Pinkas, D., & Zuccherato, R. (2001). *RFC 3161*, n. dj.

> *"Kada je privatni ključ autoriteta za vremenske žigove (TSA[333]) kompromitiran svi odgovarajući certifikati (oni koji su njime potpisani) moraju biti povučeni)".* (Adams, Cain, Pinkas, & Zuccherato, 2001)[334]

Nakon ovakvog ishoda jedino što preostaje je ponovno potpisati (dodati vremenski žig) sve dokumente koji su njime potpisani. U slučaju nekih zapisa ovo može biti značajna razlika jer će novi žig možda doći i godinu dana nakon izvornog.

Kao što je rečeno, certifikat postaje kompromitiran kada njegov privatni ključ bude ukraden ili izgubljen. S obzirom na to da većina ključeva (i certifikata) traje godinama i da se svi ne čuvaju na adekvatan način, i uz najbolje mjere predostrožnosti postoji šansa da navedeni privatni ključ bude kompromitiran. Ovaj problem Doyle rješava na način da izbjegne izdavanje ključeva korisnicima (pa čak i servisu za dodavanje vremenskih žigova) te umjesto da ključ (i certifikat) budu vezani za osobu ili ustanovu (koja ga onda mora čuvati) predlaže da oni budu vezani za kratak vremenski period. Ovakav ključ nazvan je "prolazni ključ" (engl. *transient key*).

Ako je ključ vezan za kratki vremenski period onda on identificira upravo taj trenutak, umjesto neke osobe i ustanove. U ovom slučaju, problem kompromitiranja ključa, koji identificira osobu ili instituciju, te je izazvan nemarom ili neznanjem nestaje.

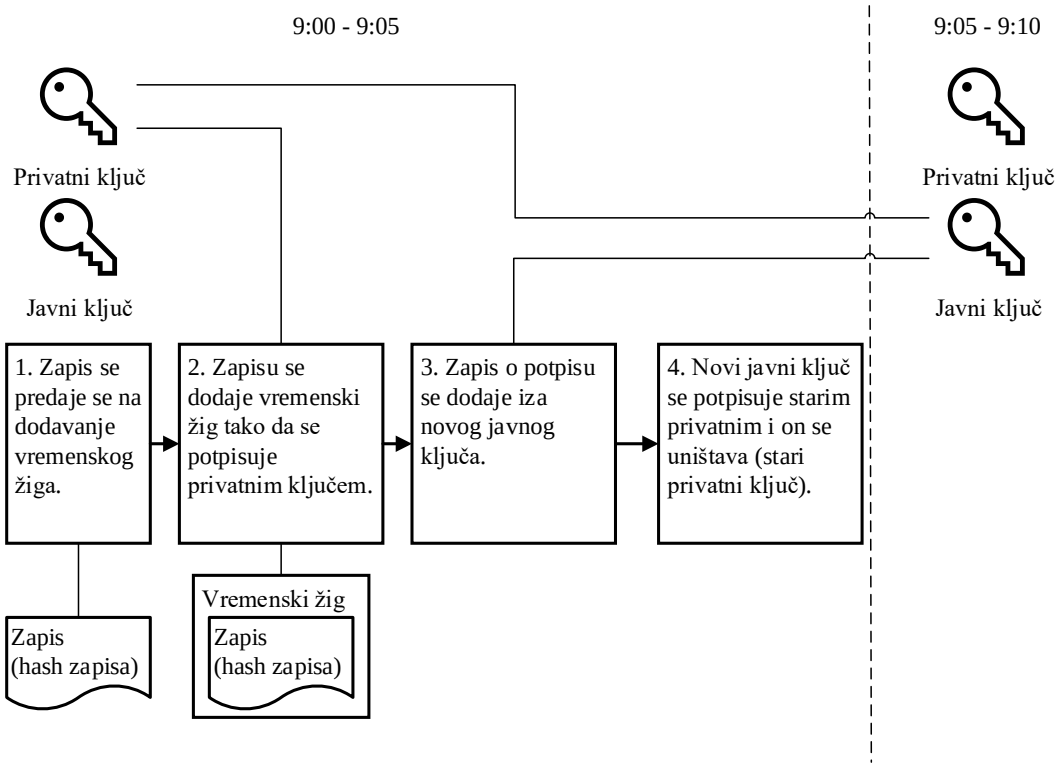
Sustav Proofmark funkcionira na način da, kao i prethodni modeli, zapise veže u nepromjenjivi lanac blokova u kojem je svaki blok potpisan prolaznim ključem koji se nakon upotrebe uništava. Samim zapisima vremenski žigovi dodaju se privatnim ključem koji je vezan uz određeni vremenski period. Ovaj ključ čuva se na poslužitelju servisa za vremenske žigove koji ga po njegovom isteku uništava te se za kasniju validaciju vremenskih žigova koristi javni ključ vezan uz konkretni vremenski period. Jedinstvenost privatnog ključa osigurava se upotrebom *nonce* (engl. *number used only once*) generatora. Generatori brojeva koji se koriste samo jednom su dobro razrađeno područje kriptografije. Pregled često korištenih generatora dao je Zenner.[335] U ovakvom sustavu ne postoji bojazan od krađe privatnog ključa jer:

[333] Engl. *Time Stamp Authority*
[334] Adams, C., Cain, P., Pinkas, D., & Zuccherato, R. (2001). *RFC 3161*, n.dj.
[335] Zenner, E. (2009). Nonce generators and the nonce reset problem. *International Conference on Information Security* (str. 411-426). Berlin, Heidelberg: Springer. Preuzeto 3. 12. 2021. s https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1061.3009&rep=rep1&type=pdf

1) Prolazni privatni ključ se koristi samo u kratkom vremenskom periodu. Krađom ključa moglo bi se stvoriti vremenski žig vezan uz taj konkretni period, te ključ ne identificira potpisnika već vrijeme potpisa.

2) Prolazni privatni ključ se ne može naknadno iskoristiti jer su hashevi svih legitimno potpisanih dokumenata dodani idućem javnom ključu te na taj način uključeni u nepromjenjivu podatkovnu strukturu (ulančane zapise).

Rad ovakvog sustava prikazan je na slici 23.



Slika 23. Rad sustava za vremenske žigove koji koristi prolazne ključeve

Proofspace je dodatno osigurao sustav uvođenjem distribuirane arhitekture poslužitelja koji su svi morali sudjelovati u stvaranju novog javnog ključa dodajući vlastiti popis zapisa kojima je dodan vremenski žig u relevantnom periodu. Razlike u popisu dobivenom od nekih drugih poslužitelja rezultirale bi neispravnim blokovima koji bi bili odbačeni.

Nažalost, sustav vremenskih žigova koji je zamislio Proofspace nije zaživio te danas ne postoji funkcionalna implementacija sustava ali su u međuvremenu istekla autorska prava na patent te se otvara mogućnost razvoja novih sustava koji koriste ovu tehnologiju. Takav sustav bio bi idealan kandidat za vanjsko dodavanje vremenskog žiga podacima sadržanim u TrustChain blokovima, modelu koji je razrađen u kasnijim poglavljima. Koncept prolaznih vremenskih ključeva danas je standardiziran u Sjedinjenim Američkim Državama u ANSI

standardu X9.95[336] koji obuhvaća sve pouzdane sheme za dugotrajno dokazivanje integriteta podataka upotrebom vremenskih žigova.

5.3. Vremenski žigovi temeljeni na ulančanim zapisima kripto valuta

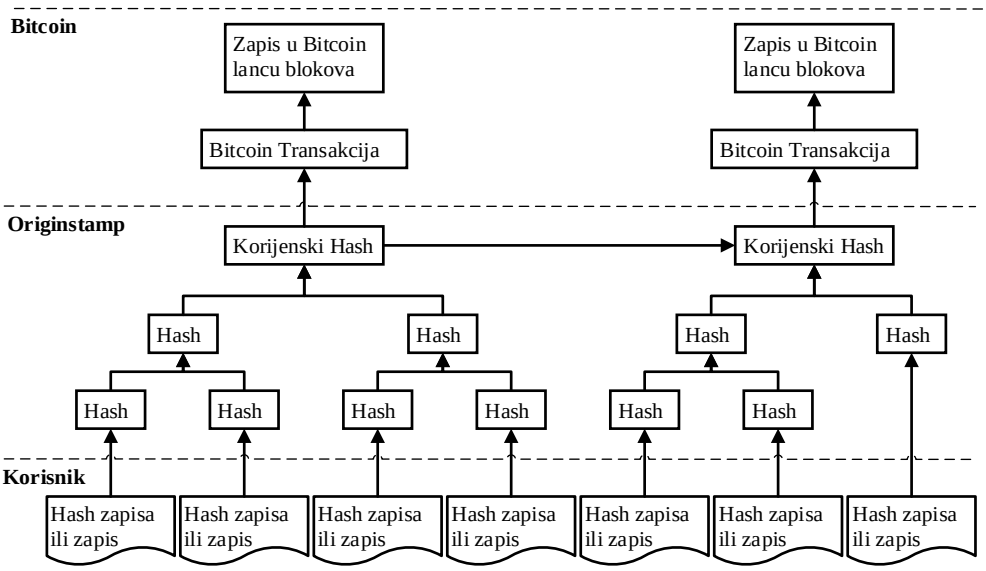
Eksponencijalni rast upotrebe kriptovaluta uzrokovao je pokušaje da se ulančani zapisi, kolokvijalno rečeno i s tuđicom *blockchain* (engl.), primjene u svim vrstama informacijskih sustava. Iako se, kao što je pokazano, ulančani zapisi u arhivske svrhe koriste duže nego kao mehanizam sigurnosti kriptovaluta nakon razvoja kriptovaluta počela su (nova) istraživanja kako iskoristiti ove postojeće lance povezanih zapisa za osiguranje integriteta podataka. Umjesto da se koriste specijalizirani ulančani zapisi predviđeni za arhivski sustav fokus je stavljen na primjenjivost postojećih struktura poput Bitcoin i Ethereum ulančanih zapisa. Dobar rani primjer ovakvih ideja dala je Lemiux.[337] Rana istraživanja pretpostavila su korištenje dodatnih polja metapodataka Bitcoin transakcija (poput OP_RETURN) kao medija za pohranu podataka o arhivskoj vezi.[338] Ovo polje bi se moglo iskoristiti i za druge svrhe, poput pohrane hasheva zapisa. Autor ove disertacije u ranim fazama vlastitog istraživanja primjene uglačanih zapisa u arhivske svrhe razmatrao je upotrebu pametnih Ethereum ugovora. Većina ranih ideja nije implementirana jer nisu pružala zaokruženo rješenje.

OriginStamp[339,340] primjer je funkcionalnog sustava za vremenske žigove koji svoje hasheve zapisuje u Bitcoin lanac zapisa. Sustav koji funkcionira na način da formira stablo hasheva (Merkleovo stablo) od primljenih zahtjeva za vremenske žigove te korijenski hash publicira kroz Bitcoin transakciju dokumentiran je 2018. godine.[341] S obzirom na to da sustav stvara vlastito stablo hasheva te većinu transakcijskih podataka čuva na vlastitim poslužiteljima

[336] ANSI. (2016). *ANSI X9.95-2016 Financial Services*, n. dj.
[337] Lemieux, V. L. (2017). Blockchain and distributed ledgers as trusted recordkeeping systems. *Future Technologies Conference (FTC).* Preuzeto 4. 12. 202.1 s https://www.researchgate.net/profile/Victoria-Lemieux/publication/317433591_Blockchain_and_Distributed_Ledgers_as_Trusted_Recordkeeping_Systems_An_Archival_Theoretic_Evaluation_Framework/links/593aa6450f7e9b3317f4d860/Blockchain-and-Distributed-Ledgers-as
[338] Lemieux, V. L., & Sporny, M. (2017). Preserving the Archival Bond, n. dj.
[339] OriginStamp URL: https://originstamp.com/
[340] OriginStamp AG. (2021). *Blockchain Timestamps for Businesses*. Preuzeto 4. 12. 2021. s OriginStamp: https://originstamp.com
[341] Hepp, T., Wortner, P., Schönhals, A., & Gipp, B. (2018). Securing physical assets on the blockchain: Linking a novel object identification concept with distributed ledgers. *Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems* (str. 60-65). New York, USA: ACM Pres. doi: https://doi.org/10.1145/3211933.3211944

ili, opcionalno, u IFPS (engl. *Inter Planetary File System)* sustavu[342] te da je jedina interakcija s Bitcoin lancem podataka zapis jedne transakcije autori smatraju da su minimalizirali cijenu objave korijenskog hasha. U 2015. prema njihovoj procjeni ova cijena iznosila je oko 0.0001 BTC ili 3 američka centa (što se kasnijim rastom vrijednosti BTC valute značajno promijenilo) te je planirana godišnja potrošnja sustava oko 10 američkih dolara godišnje.[343] Ovo je sigurno prihvatljiva cijena ali ona obuhvaća relativno mali broj godišnjih objava – samo jednu dnevno.

Osim visoke cijene zapisa, često navedeni nedostatak upotrebe Bitcoin lanaca blokova za pohranu ovakvih podataka je i vrijeme potrebno za objavu, to jest dodavanje podataka. U kasnijem radu autori Originstamp sustava navode prosječno vrijeme od 34 minute za objavu svog korijenskog hasha, s najgorim izmjerenim vremenom od 55 minuta. Ovo je u skladu s predviđenom učestalošću objave jednog hasha svakih sat vremena.[344] Ovo povećanje učestalosti objava, u odnosu na prethodna istraživanja ovih autora, bi 24 puta povećalo cijenu ali usprkos tome njena apsoluta vrijednost i dalje ostaje relativno mala, to jest oko 240 dolara godišnje (u skladu s vrijednostima BTC-a iz 2015.). Pojednostavljena shema Originstamp stabla hasheva i njegove interakcije s Bitcoin sustavom za kriptovalute prikazana je na slici 24.



Slika 24. Shema Originstamp stabla hasheva i njene interakcije s Bitcoin sustavom

[342] Benet, J. (2014). *IPFS – Content Addressed, Versioned, P2P File System.* Preuzeto 19. 11. 2022. s arxiv.org: https://arxiv.org/pdf/1407.3561.pdf

[343] Gipp, B., Meuschke, N., & Gernandt, A. (2015). Decentralized Trusted Timestamping using the Crypto Currency. *Proceedings of the iConference 2015.* Preuzeto 3. 12. 2021. s https://www.researchgate.net/profile/Thomas-Hepp-2/publication/329249467_OriginStamp_A_blockchain-backed_system_for_decentralized_trusted_timestamping/links/5c7e3313299bf1268d395112/OriginStamp-A-blockchain-backed-system-for-decentralized-trusted-timestam

[344] Gipp, B., Meuschke, N., & Gernandt, A. (2015). Decentralized Trusted Timestamping, n. dj.

Autor disertacije je u svojim istraživanjima odustao od ideje korištenja ovakvih, javnih sustava, djelomično zbog navedenih razloga (cijene i vremena objave) ali, važnije, i zbog činjenice da TrustChain model zahtjeva provjeru ispravnosti digitalnih certifikata, što je predviđeno da bi trebale odraditi uključene arhivske ustanove. Skup arhivskih ustanova, koje imaju veliki interes za sigurnu pohranu digitalno potpisanih zapisa, zasigurno ima pristup dovoljnoj vlastitoj infrastrukturi i ne moraju, a možda ni ne žele, biti u potpunosti ovisne o vanjskim servisima, što jest slučaj u slučaju primjene arhitekture kakvu koristi Originstamp. Osim toga, iako je Originstamp pokazao da cijena objave može biti relativno mala te da objave mogu biti relativno frekventne (jednom u 30 minuta) one dolaze sa značajnim ograničenjima – objavljuje se samo jedan kratki hash. Autor ove disertacije zbog ovog je razloga razmatrao upotrebu Ethereum sustava koji bi omogućio pohranu složenije podatkovne strukture. Uz to, pod pretpostavkom da Bitcoin ostane u upotrebi, i dalje ne postoji nikakva garancija da će cijene i vrijeme objave ostati takve kakve su danas.

Zbog tih razloga pristupilo se izradi modela koji se ne oslanja na vanjske sustave (ulančanih zapisa) za pohranu ključnih podataka već u skladu s ranijim modelima i standardima koristi vlastiti lanac zapisa.

5.4. Vremenski žigovi temeljeni na potpisima bez ključeva

Zadnja vrsta sustava za vremenske žigove koja će biti razmotrena je sustav tvrtke GuardTime. GuardTime [345] održava više proizvoda koji se oslanjaju na ulančane zapise uključujući i lanac ulančanih zapisa temeljen na potpisima bez ključeva, KSI sustavima (engl. *Keyless Signatures Infrastructure*). Ovaj sustav opisan je 2013. u radu Buldasa i suradnika[346] te u izvještaju tvrtke kćeri GuardTimea (GuardTime Federal LLC Proprietary),[347] a u upotrebi je i prije toga. Između ostaloga, infrastruktura potpisa bez ključa implementirana je u većinu

[345] GuardTime AS. (2021). *GuardTime*, n. dj.
[346] Buldas, A., Kroonmaa, A., & Laanoja, R. (2013). How to build global distributed hash-trees. *Nordic Conference on Secure IT Systems* (str. 313-320). Springer. Preuzeto 3. 12. 2021. s http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.434.790&rep=rep1&type=pdf
[347] GuardTime Federal LLC Proprietary. (n.d.). *Keyless Signature Infrastructure® (KSI™) Technology.* Preuzeto 3. 12. 2021. s GuardTime Library – Whitepapers: http://blockchain.machetemag.com/wp-content/uploads/2017/11/Guardtime_WhitePaper_KSI.pdf

vladinih ustanova Estonije od 2008.[348]te je prvi takav sustav koji je usklađen s Uredbom eIDAS.[349]

Centralan termin GuardTimeova sustava je KSI koncept. Prema objavljenoj literaturi ovo ne podrazumijeva puno više od stabla hasheva povezanih u ulančanu shemu (koja se u kontekstu ovog sustava zove kalendar hasheva dok sama stabla autori nazivaju agregacije hasheva. Na ovaj način garantira se nepromjenjivost podataka što je već više puta dokazivano, između ostaloga u izvornom radu o vremenskim žigovima,[350] u radovima autora stabla hasheva,[351] pa i u ovoj disertaciji. Činjenica da, prema dostupnim podacima, sustav ne uvodi nove koncepte nikako ne umanjuje njegovu vrijednost. GuardTimeova tehnologija je jedan od temelja prve digitalne državne uprave (Estonije). Na GuradTime sustavu se temelji više dijelova državne uprave, od jednostavnih upravno-administrativnih zadataka do elektroničkog sustava za glasanje[352] i medicinskih digitalnih zapisa.[353] S obzirom na svoj uspjeh GuardTime je jedan od najjačih dokaza da je tehnologija ulančanih zapisa prikladna za upotrebu u svrhe dugotrajne pohrane podatka, dakle u arhivske svrhe.

Nažalost, GuardTimeov sustav, iako snažno podupire teze ove disertacije, sam po sebi ne nudi rješenje. Nedostatci sustava su, u kontekstu dokazivanja tezi disertacije i pomoći u istraživanju, su:

1) Nedovoljno je javno dokumentiran. Mnogi dijelovi sustava uopće nisu prokomentirani u, prema istraživanju autora disertacije, javno dostupnim dokumentima. Na primjer, autori navode da se zbog sigurnosnih razloga agregacije hasheva (stabla hasheva) stvaraju u distribuiranoj okolini na, takozvanim, agregacijskim poslužiteljima ali nije objašnjeno u kakvoj su interakciji ovi poslužitelji, to jest na koji način postižu konsenzus. Ovo je kritični dio sustava koji ostaje nepoznat. Ovo je karakteristično za komercijalne sustave

[348] Vatsa, V. R., & Chhaparwal, P. (2021, July). Estonia's e-governance and digital public service delivery solutions. In *2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT)* (pp. 135-138). IEEE. Preuzeto 4. 12 2021 s https://ieeexplore.ieee.org/abstract/document/9515004
[349] European Commision. (2021). *EU Trust Services Dashboard*. Preuzeto 4. 12. 2021. iz eSignature page of the European Commision: https://esignature.ec.europa.eu/efda/tl-browser/#/screen/tl/EE
[350] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.
[351] Merkle, R. C. (1979). *Secrecy, authentication, and public key systems*, n. dj.
[352] Maaten, E. (2004). Towards remote e-voting: Estonian case. *Electronic voting in Europe-Technology, law, politics and society, workshop of the ESF TED programme together with GI and OCG*. Gesellschaft für Informatik eV. Preuzeto 4. 12. 2021. s https://dl.gi.de/bitstream/handle/20.500.12116/29132/Proceeding.GI.47-9.pdf?sequence=1
[353] Metsallik, J., Ross, P., Draheim, D., & Piho, G. (2018). Ten years of the e-health system in Estonia. *CEUR Workshop Proceedings*, (str. 6-15). Preuzeto 4. 12. 2021. s http://ceur-ws.org/Vol-2336/MMHS2018_invited.pdf

te je u sličnoj mjeri prisutno kod svih prikazanih sustava, dijelovi sustava su skriveni da bi se otežala izrada konkurentnih sustava.

2) Sustav dokazuje integritet podataka, to jest dokazuje da su podaci ostali u nepromijenjenom stanju od trenutka dodavanja vremenskog žiga. Ovo ne omogućava provjeru autentičnosti podataka kada se ta autentičnost izvorno dokazivala upotrebom u međuvremenu isteklog digitalnog certifikata.

3) Sustav ne uzima u obzir arhivsku vezu pohranjenih zapisa, to jest ona je u potpunosti prepuštena drugim sustavima koji su predviđeni za očuvanje metapodataka.

Zbog ovih razloga, kao što je slučaj i s ostalim sustavima za vremenske žigove, GuardTimeov sustav sam po sebi ne pruža adekvatne funkcionalnosti za arhivski sustav. Svejedno, on može biti upotrijebljen kao dodatak novom sustavu pružanjem vanjskih vremenskih žigova, što bi povećalo povjerenje u novi sustav. Slična upotreba GuardTimeova KSI tehnologije je već i predložena od autora Emmadi i Narumanci, oni predlažu upotrebu KSI žigova u svim tehnologijama temeljenim na ulančanim zapisima kao dodatni mehanizam osiguranja integriteta podataka.[354]

5.5. Zaključak

U poglavlju je predstavljeno više suvremenih sustava za vremenske žigove. Ovi sustavi su omogućili produživanje vjerodostojnosti digitalno potpisanih dokumenata na način da su dodali novi potpis oko prethodnog. U svojem osnovnom obliku vremenski žig nije više od posebnog digitalnog potpisa, kojeg izdaje ustanova od povjerenja i kojim se garantira postojanje podatka u trenutku potpisivanja. Vremenski žigovi su izvorno rješenje problema isteka certifikata, koje je prethodilo intenzivnom razvoju novih sustava temeljenih na ulančanim zapisima (poput ranije prikazanih sustava i modela razvijenog sljedećem poglavlju ove disertacije) te je njihova upotreba na taj način dobro dokumentirana. Prije izvođenja zaključka poglavlja predstavljen je i kratki pregled tih ranijih rješenja i prijedloga. Iako većina od njih nije suvremena, one omogućuju uvid u razvoj ove tehnologije te pružaju potvrdu principa na kojima je temeljen novi TrustChain model.

[354] Metsallik, J., Ross, P., Draheim, D., & Piho, G. (2018). Ten years of the e-health system in Estonia, n. dj.

Neki od prijedloga ne uključuju čak ni vremenske žigove već predlažu da se digitalni dokumenti ponovno potpišu novim (još ne-isteklim) certifikatom, na primjer Pharow i Blöbel.[355] Autori Hyla, Pejaś i Bielecki[356] kombiniraju metode pa navode mogućnost upotrebe oba rješenja, to jest ili novog certifikata ili vremenskog žiga.

Ipak većina prethodnih prijedloga predlaže upotrebu neke vrste vremenskog žiga za naknadno potpisivanje dokumenata. Primjeri ovakvih prijedloga dolaze od Chen i Lina[357] te od Stančića, Rajha i Brzice.[358] Glavni razlog za upotrebu vremenskog žiga je taj da su vremenski žigovi, čak i kada dolaze u obliku običnog certifikata značajno dugotrajniji (iako ni klasični vremenski žig nije vječan) te njihova upotreba zahtjeva manje postupaka naknadnog potpisivanja pa u skladu s time, posljedično, rezultira manjim rastom veličine arhiviranih dokumenata (jer se dodaje manje omotnica s novim potpisima).

Jedan od najranijih prijedloga ovakve upotrebe vremenskih žigova dali su Maniatis i Baker.[359] Ovi autori predlažu posebnu arhitekturu digitalnog arhiva koja će uključivati certifikate kojima su potpisani arhivirani dokumenti, kojima se naknadno dodaju vremenski žigovi. Iznesena ideja preteča je ideje dijela modela koji je predstavljen u kasnijim poglavljima ove disertacije (TrustChain B model).

Jedan od većih problema kod ovakvih sustava za naknadno potpisivanje leži u činjenici da je potrebno potpisati veliki broj dokumenata te se time stvara značajno opterećenje za arhivske sustave. Schwalm, kao i drugi autori, u svom prijedlogu predlaže rješenje problema tako da se potpisuju samo korijenski hashevi stabla hasheva zapisa.[360] Novi model, kao i ovaj

[355] Pharow, P., & Blöbel, B. (2005). Electronic signatures for long-lasting storage purposes in electronic archives. *International Journal of Medical Informatics, 74*(2-4), 279-287. Preuzeto 5. 12. 2021. s https://ebooks.iospress.nl/pdf/doi/10.3233/978-1-60750-939-4-316

[356] Hyla, T., W. Bielecki, W., & Pejaś, J. (2010). Non-repudiation of Electronic Health Records in distributed healthcare systems. *Pomiary Automatyka Kontrola, 56*(10), 1170-1173. Preuzeto 5. 12. 2021. s https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-article-BSW4-0086-0015/c/Hyla.pdf

[357] Chen, T., & Lin, F. (2011). Electronic medical archives: a different approach to applying re-signing mechanisms to digital signatures. *Journal of medical systems, 35*(4), 735-742. doi: https://doi.org/10.1007/s10916-009-9414-2

[358] Stančić, H., Rajh, A., & Brzica, H. (2015). Archival Cloud Services: Portability, Continuity, and Sustainability Aspects of Long-term Preservation of Electronically Signed Records. *Canadian Journal of Information and Library Science, 39*(2), 210-227. Preuzeto 5. 12. 2021. s https://muse.jhu.edu/article/590941

[359] Maniatis, P. & Baker, M.. (2002). Enabling the Archival Storage of Signed Documents. *FAST 2002 Proceedings*, (str. 31-46). Preuzeto 5. 12. 2021. s http://shiftleft.com/mirrors/www.hpl.hp.com/personal/Mary_Baker/publications/fast2002.pdf

[360] Schwalm, S. (2017). A service for the preservation of evidence and data–a key for a trustworthy & sustainable electronic business. *Open Identity Summit 2017*. Bonn: Gesellschaft für Informatik. Preuzeto 5. 12. 2021. s https://dl.gi.de/bitstream/handle/20.500.12116/3571/proceedings-10.pdf?sequence=1&isAllowed=y

prijedlog, vanjskim vremenskim žigom potpisuje samo vlastite korijenske hasheve (koje novi model naziva hashevima blokova). Upotreba ovakvog dokaza je u skladu je s Uredbom eIDAS.

Vigil i suradnici su 2011. objavili pregled i temeljitu usporedbu više postojećih rješenja za naknadno dodavanje vremenskih žigova i dugotrajnu pohranu dokumenta.[361] Usporedili su ETSI standarde, sintaksu za dokazivanje zapisa (engl. *Evidence Record Syntax*) definiranu dokumentom RFC4998,[362] ACE (engl. *Auditing Control Environment*) sustav,[363] CIS (engl. *Content Integrity Service*), sustav koji su za potrebe Hewlett Packarda razvili Haber i Kamat,[364] LOCKSS (engl. *Lots of Copies Keep Stuff Safe*)[365] i OC (engl. *Optimized Certificates*) sustav.[366] Osim ovih rješenja, za koje autori smatraju da su prihvatljiva, rad je identificirao i više sustava za koje autori smatraju da, bez dodatnih mehanizama, uopće ne omogućuju "dugotrajnu zaštitu digitalnih podataka" uključujući i ranije spomenuti model Maniatisa i Bakera. Kod rješenja za koje smatraju da su adekvatna autori su ipak upozorili da i dalje postoje "poželjne osobine (sustava) koje nisu pokrivene rješenjima za dugotrajno očuvanje". Na kraju zaključka ove opsežne analize autori ističu da "pronalazak rješenja koja zadovoljavaju dugoročne potrebe za povjerljivošću, druge zaštitne ciljeve te balansa između kriptografske zaštite i oslanjanja na ustanove kojima se vjeruje se doima kao zanimljiv i važan pravac istraživanja".[367] Balans između oslanjanja na kriptografski dokaz i povjerenje u skup ustanova upravo je temeljna

[361] Vigil, M., Cabarcas, D., Wiesmaier, A., & Buchmann, J. (2011). Authenticity, integrity and proof of existence for long-term archiving: a survey. *Cryptology EPrint Archive*. Preuzeto 4. 12. 2021. s https://d1wqtxts1xzle7.cloudfront.net/39270547/54920f5b0cf2484a3f3e092f-with-cover-page-v2.pdf?Expires=1638641185&Signature=Rkf~VTGJzZrTsQ~y3PXakhTN~TP6TwPbjzcJ4tvujltSnQQAy~EjNYoSuj1JqGwjYmdr0-z0P8o3TYeGGm7t9IL2ETrp~QRPmyF1msiTqLW5nz7Yhfd-EsDm3pjtEdf-njX

[362] Gondrom, T., Brandner, R., & Pordesch, U. (2007). *RFC4998: Evidence Record Syntax (ERS).* Preuzeto 5. 12. 2021. iz IETF Datatracker: https://datatracker.ietf.org/doc/html/rfc4998

[363] Song, S., & JaJa, J. (2009). Techniques to audit and certify the long-term integrity of digital archives. *International Journal on Digital Libraries, 10*(2-3), 121-131. Preuzeto 5. 12. 2021. s https://drum.lib.umd.edu/bitstream/handle/1903/7130/ACE-techniques-UMIACS-TR-2007-38.pdf?sequence=1&isAllowed=y

[364] Haber, S., & Kamat, P. (2006). A content integrity service for long-term digital archives. *Archiving Conference, 2006*, str. 159-164. Preuzeto 5. 12. 2021. s http://www.hpl.hp.com/techreports/2006/HPL-2006-54.pdf

[365] Maniatis, P., Roussopoulos, M., Giuli, T., Rosenthal, D., & Baker, M. (2005). The LOCKSS peer-to-peer digital preservation system. *ACM Transactions on Computer Systems (TOCS), 23*(1), 2-50. Preuzeto 5. 12. 2021. s https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.60.2234&rep=rep1&type=pdf

[366] Custódio, R., Vigil, M., Romani, J., Pereira, F., & da Silva Fraga, J. (2008.). Optimized certificates–A new proposal for efficient electronic document signature validation. *European Public Key Infrastructure Workshop* (str. 49-59). Berlin, Heidelberg: Springer. Preuzeto 5. 12. 2021. s ftp://nozdr.ru/biblio/kolxoz/Cs/CsLn/P/Public%20Key%20Infrastructure,%205%20conf...%20Theory%20and%20Practice,%20EuroPKI%202008(LNCS5057,%20Springer,%202008)(ISBN%209783540694847)(247s).pdf#page=57

[367] Vigil, M., Cabarcas, D., Wiesmaier, A., & Buchmann, J. (2011). Authenticity, integrity and proof of existence, n. dj.

pretpostavka novog modela za dugotrajno očuvanje digitalno potpisanih zapisa koji je razvijen u ovom radu.

Sva navedena istraživanja nude važan uvid u problematiku arhiviranja digitalnog potpisanih dokumenta, a ranije prikazani sustavi su konkretni primjeri sustava koji su u upotrebi i koji (donekle) ispunjavaju potrebe za dugotrajnim očuvanjem digitalno potpisanih zapisa. Problemi s postojećim sustavima već su raspravljeni. Oni ne ispunjavaju u potpunosti zahtjeve arhivistike, na primjer ne uvažavaju potrebu za evidencijom podataka o arhivskoj vezi niti pružaju dokaz autentičnosti podataka (najčešće to ni ne pokušavaju napraviti). Osim toga svi ovi sustavi nisu dovoljno (javno) dokumentirani. Dokumentacija se uglavnom svodi na patente (kada postoje), marketinške materijale koji iznose tvrdnje o učinkovitosti sustava ali ne i načinu na koji se ona postiže ili znanstvene i stručne radove koji (vjerojatno namjerno) nisu dovoljno detaljni. Čak i u slučaju sustava AbsoluteProof[368] koji tvrdi da je tehnička dokumentacija dostupna na zahtjev, autor ove disertacije nije uspio dočekati njezinu dostavu. Iako potpuno razumljivo s gledišta tvrtki koje moraju djelovati na otvorenom tržištu, ovakva ograničenja dostupnosti podataka o tehničkoj izvedbi sustava značajno umanjuju mogućnost istraživanja i znanstvenog doprinosa temeljenog na takvim sustavima. Upravo zbog toga što postojeća, komercijalna, rješenja ne omogućuju dovoljno detaljan uvid u svoj način funkcioniranja cilj ove disertacije je dokazati da je moguće upotrijebiti ulančane zapise za dugoročno očuvanje digitalno potpisanih arhivskih zapisa tako da razvije model koji je dovoljno detaljno razrađen da se nakon njegovog razumijevanja može pouzdano tvrditi da se njime točno to i ostvaruje.

[368] Surety, I. (2021). *AbsoluteProof*, n. dj.

6. Model za dugotrajnu pohranu digitalno potpisanih dokumenata

Ranije opisana problematika dugotrajne pohrane digitalno potpisanih zapisa, to jest problemi vezani uz istek certifikata i o njemu ovisnih digitalnih potpisa bili su tema istraživanja radne skupine TRUSTER Preservation Model (EU31) međunarodnog istraživanja InterPARES Trust[369] u kojima je sudjelovao i autor ove disertacije. Tijekom projekta InterPARES Trust autor disertacije je, u suradnji s Magdalenom Kuleš i Hrvojem Stančićem (koji je i mentor ove disertacije), razvio ideju o upotrebi sustava baziranih na ulančanim zapisima kao moguće rješenje ranije navedenih problema dugotrajnog očuvanja autentičnosti i integriteta digitalno potpisanih dokumenata. Nova ideja, TrustChain, bila je jedno od predloženih rješenja radne skupine za problem dugotrajne pohrane digitalno potpisanih dokumenata.

Sama ideja upotrebe ulančanih zapisa kao garancije integriteta podatka nije nova. Ovo i jest izvorna svrha ulančanih zapisa te je njihova primjena u arhivistici, primarno u digitalnim arhivima uočena još za vrijeme prošlog stoljeća. Najpoznatiji istraživači ove teme, Stuart Haber i William Scott Stornetta, 1991. godine opisali su upotrebu ulančanih zapisa, u kombinaciji s javno objavljenim fizičkim (ne-digitalnim) publikacijama kao rješenje za dugotrajno osiguranje digitalnih arhivističkih zapisa.[370] Istraživanje i model koji su razrađeni u ovoj disertaciji značajno se razlikuju od rada Habera i Stornette jer se, iako su bazirani na istim principima, fokusiraju na očuvanje autentičnosti digitalno potpisanih dokumenta. Suvremeni digitalni potpisi realizirani su upotrebom ranije opisanog digitalnog potpisa i certifikata. Digitalni certifikati nisu bili u upotrebi 1991. godine pa tadašnji (digitalni) arhivski sustavi nisu morali uzimati u obzir njihov istek. Model predstavljen u ovom poglavlju naglašava upravo rješenja za dugotrajno očuvanje digitalnih potpisa i certifikata, instrumenata za dokazivanje autentičnosti podataka koji su vremenski ograničeni. Potreba za ovakvim sustavom je i dokazana u studijama slučajeva koje su provedene tijekom InterPARES Trust istraživanja.

Izvorna ideja sustava za dugotrajno očuvanje digitalno potpisanih dokumenata s nazivom TrustChain prezentirana je na konferenciji INFuture 2017. u radu "A model for long-term preservation of digital signature validity: TrustChain".[371] Rad je opisao osnovni model

[369] https://interparestrust.org/
[370] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n.dj.
[371] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.

sustava koji je odgovoran za provjeru ispravnosti digitalnog potpisa i stvaranje novog zapisa koji je trajno pohranjen u ulančanom zapisu.

Tri godine kasnije, na osnovu saznanja o pohranjivanim dokumentima stečenim tijekom InterPARES Trust projekta i sugestijama Mats Stengårda, kolege iz blockchain inovacijske tvrtke Enigio i jednog od suradnika u InterPARES Trust istraživanju, razvijena je nova varijanta TrustChain modela koji umjesto provjere cijelog dokumenta provjerava samo potpise, to jest certifikate, i u ulančani zapis trajno pohranjuje samo digitalne certifikate (i certifikacijske lance). Ovakav sustav izbjegava probleme koji nastaju pri provjeri vjerodostojnosti potpisa na povjerljivom dokumentu. Druga verzija TrustChain modela objavljena je u časopisu Records Management Journal[372].

Zadnji javno objavljeni rezultati istraživanja TrustChain modela, prije završetka pisanja ove disertacije, pojavili su se u časopisu Computers 2021.[373] godine. U tamo objavljenom radu autor i mentor disertacije istražuju načine rješavanja problema veznih uz upotrebu nepromjenjive podatkovne strukture (ulančanih zapisa) i zahtjeva arhivistike za mogućnošću izmjene, to jest proširivanja metapodataka zapisa. Ova promjenjivost je posebno izražena kada se u obzir uzme potreba za evidencijom podataka o arhivskoj vezi. Autori su prihvatili mogućnost upotrebe standardnih bazi podataka te način odabira i njihovu primjenu opisali u spomenutom radu.[374]

U disertaciji će, u idućim poglavljima, biti prikazana oba postojeća modela kao osnova za novi model koji integrira obje ranije navede varijante modela: pohranu cijelih digitalno potpisanih dokumenata, to jest zapisa, i pohranu digitalnih certifikata, to jest lanaca digitalnih certifikata. Struktura poglavlja kronološki prati tri ranije navedena izvora te je prema njima strukturirana. Poglavlje počinje s pregledom izvornog TrustChain modela koji je ovdje značajno razrađen i prikazan kao modul većeg sustava. Osim razrade postojećeg modela u početni dio većim su dijelom uključeni i elementi, to jest procesi i podatkovne strukture konačnog sustava koji su univerzalni za izvorni model i onaj prezentiran u časopisu Records Management Journal 2020.[375]

[372] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving: digital signature certification chain preservation. *Records Management Journal, 30*(3), 345-362. Preuzeto 4. 5. 2021. s https://www.emerald.com/insight/content/doi/10.1108/RMJ-08-2019-0043/full/html
[373] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.
[374] Ibid.
[375] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving, n. dj.

U narednom poglavlju je razrada modela, i prerada u modul većeg sustava, odrađena za TrustChain model iz 2020.[376] Nakon razrade oba modela detaljnije je razmotreno kako pohraniti metapodatke i odabrati tehnologiju korištenu za pohranu metapodataka, to jest pomoćnu bazu podataka baziranu na radu iz 2021.[377]

Tijekom ovog poglavlja, zbog lakše organizacije teksta, modeli se referenciraju na sljedeći način:

- *TrustChain A* se odnosi na izvornu ideju, predstavljenu 2017. na konferenciji INFuture. Model je često referenciran kao TrustChain A modul jer je postao dio većeg modela, to jest informacijskog sustava.
- *TrustChain B* se odnosi na varijantu sustava predstavljenu 2020. u časopisu Records Management Journal. Kao što je bio slučaj i s TrustChain A, i ovaj model je transformiran u modul većeg sustava.
- Termin *TrustChain* se odnosi na novi integrirani model koji je izvorno razrađen u ovoj disertaciji i koji uzima u obzir sva ranije provedena istraživanja.

Osim ovih, novih, naziva za TrustChain modele, za potpuno razumijevanje teksta dobro je odmah raspraviti i sljedeće pojmove:

- *TrustChain lanac* blokova se odnosi na niz ulančanih zapisa (engl. *blockchain*) koji je centralni repozitorij podataka TrustChain sustava. Ovo je nepromjenjiva podatkovna struktura koja omogućuje dugotrajno dokazivanje integriteta i autentičnosti pohranjenih zapisa.
- *TrustChain pomoćna baza podataka* odnosi se na novi sustav koji je baziran na bazi metapodataka prezentiranoj 2021. u časopisu Computers ali koristi istu bazu i istu tehnologiji da bi pohranila i neke druge podatke kritične za funkcioniranje TrustChain sustava (poput reda zapisa za dodavanje, popisa čvorova i drugih kasnije uvedenih elemenata.

[376] Ibid.
[377] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.

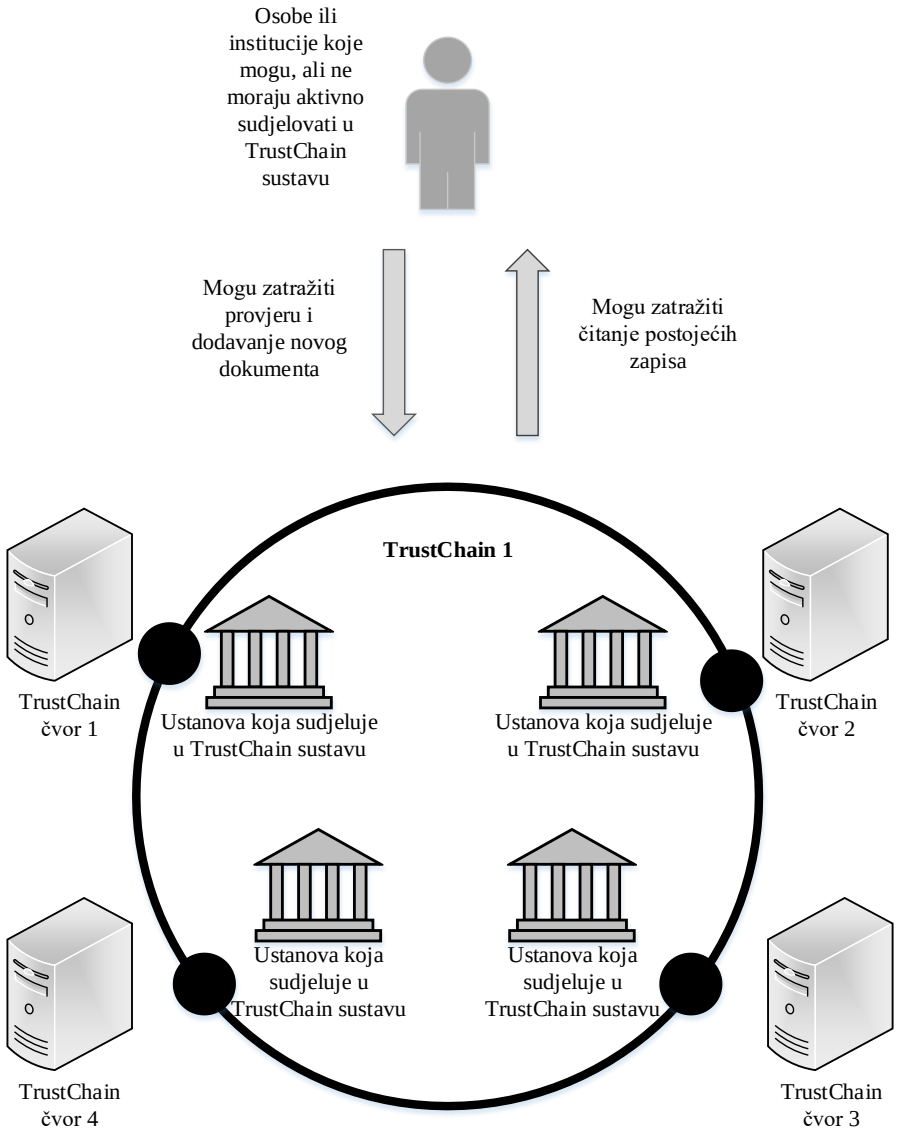
- Procesi svih modula su označeni kraticama koje počinju kraticom TC (TrustChain) te nose informaciju o modulu (A ili B) i rednom broju procesa, na primjer, TCA1. Opći procesi, koji nisu karakteristični za konkretne module ne nose oznaku modula, na primjer TC1.
- Svi potprocesi su navedeni kao postupci u procesu te nose redni broj. Na ovaj način je u raspravi ili daljnjoj razradi TrustChain modela (ili razvoju sustava) jednostavno referencirati konkretni postupak (potproces), na primjer, TCA1.1.

Iduća poglavlja će prikazati prethodne modele te detaljnije prikazati novi model, koji se na njima temelji i koji u ovoj disertaciji preuzima naziv TrustChain, te koji obuhvaća funkcionalnosti obaju sustava i u njih uključuje nova saznanja poput potrebe za promjenjivosti metapodataka i potrebe za upotrebom vanjskog vremenskog žiga.

6.1. TrustChain A

Izvorni TrustChain model,[378] koji je u daljnjem tekstu nazvan TrustChain A, opisuje informacijski sustav čiji je cilj dugotrajno očuvanje digitalno potpisanih dokumenta (karakteristika TrustChain A modela) na način da podatke o njima (metapodatke i hasheve potpisanih dokumenta) pohranjuje u nepromjenjivu podatkovnu strukturu ulančanih blokova. Naknadna istraživanja uočila su problem vezan uz ovaj model koji proizlazi iz činjenice da TrustChain A pretpostavlja da će više od jedne nezavisne ustanove potvrditi ispravnost svakog digitalno potpisanoga dokumenta. Ovaj koncept prikazan je na slici 25.

[378] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.



Slika 25. TrustChain A koncept. Izvor: Bralić, Kuleš, Stančić. A model for long-term preservation of digital signature validity: TrustChain. 2017.

Shema sa slike 25 prikazuje osnovni koncept i interakciju osobe (ili ustanove) koja koristi usluge informacijskog sustava baziranog na TrustChain A modelu sa više (arhivskih) ustanova koje sudjeluju u realizaciji sustava. Prije zapisa digitalno potpisanoga dokumenta u nepromjenjivu podatkovnu strukturu potrebno je dokazati da je u trenutku zapisa digitalni potpis (i pripadajući digitalni certifikat) ispravan. TrustChain model se u ovoj, a ni u kasnijim varijantama, nije oslanjao na koncepte koji su uobičajeni u sustavima za kripto valute temeljenim na ulančanim zapisima, poput dokaza rada (engl. *proof of work*) ili dokaza udjela (engl. *proof of stake*). Ovi koncepti garantiraju ispravnost podataka u potpuno otvorenim sustavima poput BitCoin-a i Ethereum-a. S obzirom na to da je sustav baziran na TrustChain modelu, na razini provjere i stvaranja novog zapisa, zatvoren i namijenjen arhivskim

institucijama za ovakvim dokazima nije bilo potrebe, pa se stoga može ustvrditi da za njegov rad nije potrebno mnogo struje, kao što je to potrebno za izračun novog bloka nekih kriptovaluta. Usprkos tome, da bi se osiguralo da jedna institucija ne može manipulirati podacima, TrustChain je dizajniran sa sustavom glasanja, to jest potvrđivanja ispravnosti svih ulaznih zapisa, digitalnih potpisa i pripadajućih digitalnih certifikata. Svaki zapis, prije dodavanja u podatkovnu strukturu ulančanih blokova, prolazi proces provjere ispravnosti potpisa od strane više ustanova. Svaka ustanova mora provjeriti je li digitalni certifikat ispravan i odgovara li sam potpis podacima u zapisu (u kontekstu TrustChain A modela ovo je obično digitalni dokument, na primjer PDF datoteka). Ustanove kvalificiranom većinom potvrđuju da su potpis i certifikat ispravni te se on nakon toga pohranjuje u lanac zapisa. Naknadno je uočeno da, iako ovakav pristup provjeri ispravnosti dokumenta značajno smanjuje mogućnost manipulacije podacima, pristup stvara nepremostiv problem u slučaju povjerljivih podataka. TrustChain je najpouzdaniji kada u njemu sudjeluje veliki broj potpuno nezavisnih institucija, neke od kojih mogu biti i državni arhivi iz više različitih država. U ovakvom okruženju sasvim sigurno će se pojaviti potreba za provjerom ispravnosti digitalnih dokumenta koji ne mogu napustiti nadležnost arhiva koji ih pohranjuje ili države u čijoj su nadležnosti. Osim toga, u slučaju otvaranja sustava prema javnosti, davanjem mogućnosti predaje zahtjeva za provjerom ispravnosti digitalno potpisanoga dokumenta od strane bilo koga (fizičkih osoba), problem će biti još izraženiji. Ovo je najveći nedostatak izvornog modela, koji je riješen u sljedećem modelu – TrustChain B. Usprkos tome, TrustChain A je uključen u ovu disertaciju te je njegova funkcionalnost opcionalni dio konačnog modela koji je ovdje opisan. Osim toga TrustChain A zbog svoje jednostavnosti i fleksibilnosti pruža dobru osnovu za razvoj složenijeg modela. Da bi omogućili ovakvu integraciju izvorni model je modificiran na način da funkcionira kao dio većeg sustava.

Osim navedene izmjene, TrustChain A model koji je ovdje prezentiran, za razliku od izvornog modela, uključuje mogućnost upotrebe vanjskog vremenskog žiga kako bi podigao razinu sigurnosti i povjerenja u zapise te pretpostavlja upotrebu sekundarne baze podataka koja omogućuje pohranu više metapodataka, promjenu metapodataka i jednostavnu pretragu ulančanih blokova.

Na ovaj način TrustChain A model je iz nezavisnog, potpunog modela, transformiran u modul većeg sustava koji omogućava provjeru cijelog digitalno potpisanoga zapisa kada za to nema prepreka i za time postoji potreba (što je samo dio funkcionalnosti potpunog TrustChain modela). Ovo poglavlje prikazuje TrustChain A u ovom izmijenjenom obliku.

6.1.1. TrustChain A procesi

Ovo poglavlje opisuje osnovne procese u TrustChain A modulu. U odnosu na izvorni model, predstavljen 2017., ovdje opisani procesi podrazumijevaju da se model koristi kao dio većeg sustava koji obuhvaća procese iz A i B modela. Pregled osnovnih procesa počinje s opisom proces TCA1 – dodavanje novog dokumenta. Ovaj proces prikazan je niže, na slici 26. Slika sadrži niz postupaka (i povezanih odluka) koje su, kako je ranije objašnjeno, da bi omogućili lakšu raspravu numerirane brojevima od 1 do 9. Svi postupci koji su prikazani shemama u ovom poglavlju uključuju sličnu numeraciju.

Prikazani proces (TCA1) podrazumijeva stvaranje novog individualnog TrustChain zapisa i seta podatka koji se odnose na individualni dokument (u slučaju TrustChain A modela to je najčešće digitalno potpisana datoteka) koji se u kasnijem procesu dodatno provjerava i uključuje u TrustChain lanac blokova. Proces se u odnosu na izvorni model[379] razlikuje po tri dodatka:

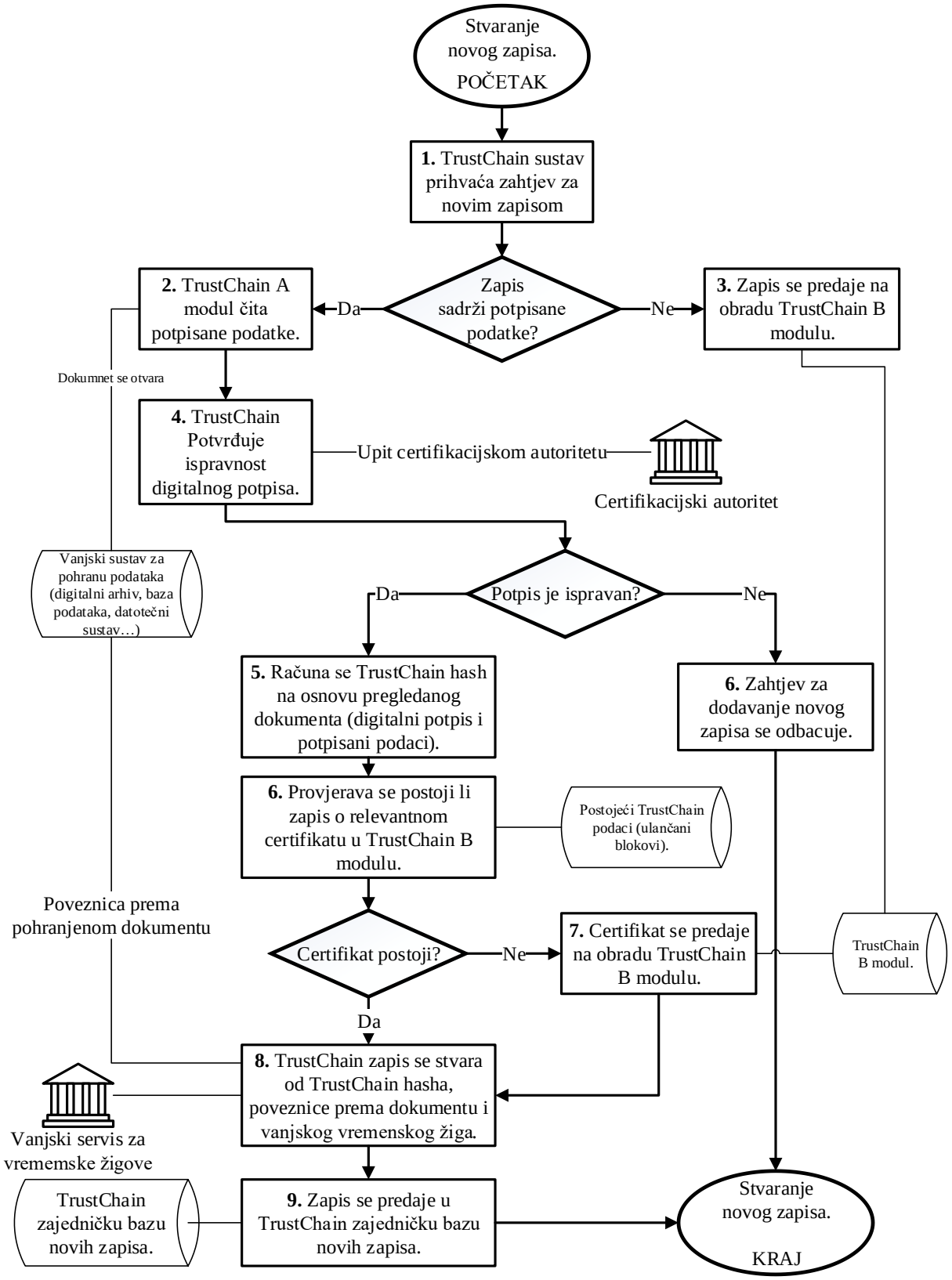
a) Dodana je mogućnost da predani zapis uopće ne sadrži potpisane podatke – u ovom slučaju zapis se predaje na obradu TrustChain B modulu koji je ovdje apstrahiran pod postupkom TCA1.3.

b) Dodan je novi postupak koji provjerava postoji li zapis o digitalnom certifikatu koji se provjerava u postojećem TrustChain lancu blokova. Ako certifikat ne postoji on se automatski predaje TrustChain B modulu koji će i njega uključiti u idući TrustChain ulančani blok. Ovo je dodatni proces koji ne isključuje pregled predanog dokumenta i pohranu njegovog hasha. Ovaj proces opisan je postupcima TCA1.6. i TCA1.7.

c) Dodana je upotreba vanjskog vremenskog žiga u postupku TCA1.8.

U odnosu na izvorni model naknadno je uočena i potreba za upotrebom vanjske baze podataka koja će pohranjivati metapodatke i omogućiti lakšu pretragu ulančanih blokova. Iako se možda čini kao da bi zapis ovih podataka trebao biti dio ovog procesa on još nije moguć jer zahtjeva podatke o rednom broju bloka u koji će zapis biti uvršten. Iz tog razloga ovaj dodatak pokriven je kasnijim procesom (TC2).

[379] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.



Slika 26. Proces TCA1 – dodavanje novog zapisa u Trust Chain A modul

Najvažniji dio procesa stvaranja novog zapisa (TCA1) sadržan je u postupcima TCA1.4., TCA1.5. i TCA1.8. pa će oni biti detaljnije opisani.

Postupak TCA1.4. provjerava ispravnost digitalnog potpisa na način na koji bi to napravio bilo koji drugi servis, računanjem hasha potpisanih podataka i dekripcijom digitalnog potpisa upotrebom pripadajućeg javnog ključa. U slučaju neispravnog ili isteklog potpisa zahtjev za zapisom se odbacuje. Kao dio procesa stvaranja novog zapisa ovu provjeru odradit će samo čvor koji stvara zapis, a u kasnijem procesu (TC2 – stvaranje bloka te TC3 i TC4 – glasanje o njegovoj ispravnosti) provjera mora biti ponovljena i ispravnost mora potvrditi većina čvorova koji sudjeluju u TrustChain sustavu.

Postupak TCA1.5 računa novi hash koji obuhvaća digitalni potpis i potpisane podatke, to jest cjeloviti skup podataka koji su predani na provjeru. Ovaj hash će biti uvršten u TrustChain lanac blokova te kasnije može biti korišten za provjeru vjerodostojnosti zapisa. Uvrštavanje cijelog dokumenta bi nepotrebno opteretilo cijeli sustav. Ranija poglavlja su detaljno objasnila zašto smatramo da je upotreba hasha dovoljan dokaz postojanja podataka u izvornom obliku. Osim ovog zapisa, u procesima opisanim u sljedećim poglavljima će i digitalni certifikat koji je korišten za stvaranje potpisa biti zapisan u TrustChain lanac blokova upotrebom B modula (ako već nije zapisan).

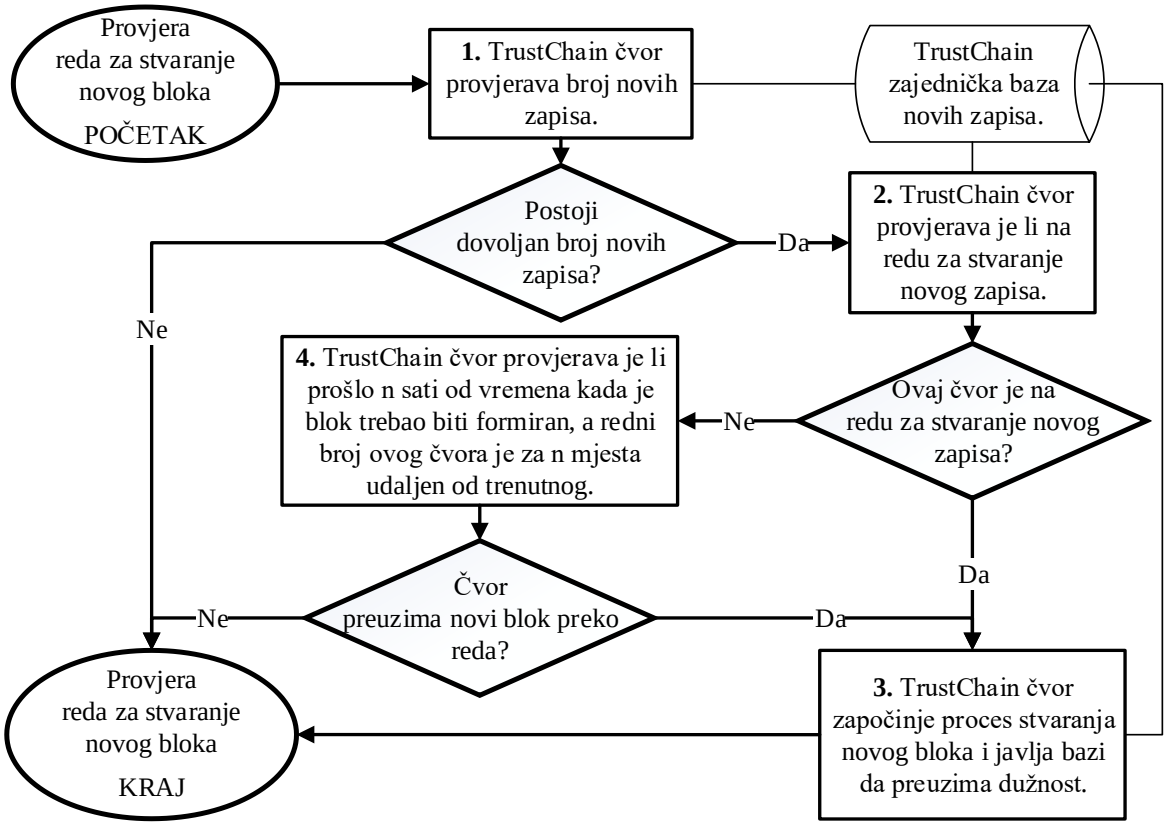
Konačni zapis stvara se u postupku TCA1.8. Rezultat ovog postupka je dvostruk. Stvaraju se dva odvojena seta podatka koji će u idućem postupku (TCA1.9.) biti korišteni za stvaranje zapisa u dva odvojena repozitorija:

1) Stvara se set podatka zapisan prema XML standardu koji se predaje na daljnju provjeru i uvrštavanje u idući TrustChain ulančani blok.

2) Stvara se set metapodataka prema XML standardu koji se u kasnijem procesu – TC3, nakon dodjele rednog broja ulančanog bloka pod kojim su zapisani podaci iz točke 1, pohranjuje u bazu podataka koja služi za pohranu metapodataka i pretragu ulančanih blokova.

Zaključivanje ova dva seta podataka označava kraj procesa dodavanja novog zapisa u TrustChain A modulu i početak procesa stvaranja novog (ulančanog) bloka podataka. Novi podaci će privremeno biti sadržani u TrustChain zajedničkoj bazi podataka, koja djeluje kao red za dodavanje zapisa u lanac blokova, te će iz njega biti preuzeti za daljnju obradu u idućem procesu. Prije nego što započne postupak formiranja i dodavanja novog bloka potrebno je utvrditi koji čvor će preuzeti odgovornost za formiranje novog bloka. Svi čvorovi koji sudjeluju u TrustChain sustavu linearnim redom preuzimaju ovu dužnost prema niže prikazanom postupku, to jest kada utvrde da je a) u zajedničkoj bazi akumulirano dovoljno zapisa za novi

blok i b) da je njihov red da formiraju novi blok. Ovaj jednostavni proces prikazan je na slici 27 te ga svi čvorovi provode svakih sat vremena (učestalost provjera može biti češća ili rjeđa ovisno o brzini kojom se stvaraju novi zapisi). Ovaj proces je univerzalan za TrustChain A i B modul pa, iako je ovdje prikazan zbog logičnog slijeda procesa, nosi oznaku TC1 (bez šifre modula).



Slika 27. Proces TC1 – odabir čvora koji stvara novi blok

U procesu sa slike 27 treba pojasniti korak 4. U ovom postupku čvor odlučuje je li vrijeme da preko reda preuzme dužnost stvaranja novog bloka. Ako čvor koji trenutno radi provjeru ima redni broj 9, a baza kaže da je na redu čvor broj 8 te je od vremena kada je blok trebao biti formiran prošlo sat vremena, ovaj čvor (redni broj 9) će preko reda preuzeti dužnost formiranja novog bloka. Ovaj korak uveden je da se izbjegne blokada sustava u slučaju ispada jednog ili više čvorova. Da bi ovaj postupak funkcionirao u bazi zajedničkih zapisa se vodi posebna stavka s brojem čvora koji je na redu i vremenom kada bi trebao preuzeti dužnost. Baza zajedničkih zapisa je baza podataka bazirana na postojećim tehnologijama koja je opisana u kasnijim poglavljima ali je istovremeno pokrenuta na istom poslužitelju kao i baza metapodataka.

Proces odabira čvora koji formira novi blok je nova razrada TrustChain A modela, takav proces nije postojao u ranijem modelu. U izvornom modelu čvorovi su čekali da sami, lokalno, zaprime dovoljan broj zahtjeva prije nego što su počinjali postupak stvaranja novog bloka od vlastitih zapisa. Na ovaj način čvorovi mogu dijeliti nove zapise te je proces optimiziran na način da je izbjegnuto predugo čekanje zapisa u redu.

Kada čvor utvrdi da je njegov red da formira novi blok započinje novi proces – TC2. Ovaj proces stvara novi blok, nadovezuje se na procese TC3 i TC 4 koji opisuju glasanje čvorova te završava uvrštavanjem bloka u TrustChain lanac blokova. Proces TC2 opisan je slikom 28.

Kao i u slučaju procesa prikazanog na slici 27, proces TC2 se bazira na izvornom radu prezentiranom na INFuture konferenciji 2017.[380] ali je ovdje značajno razrađeniji te uključuje tri važna dodatka u odnosu na izvorni model:

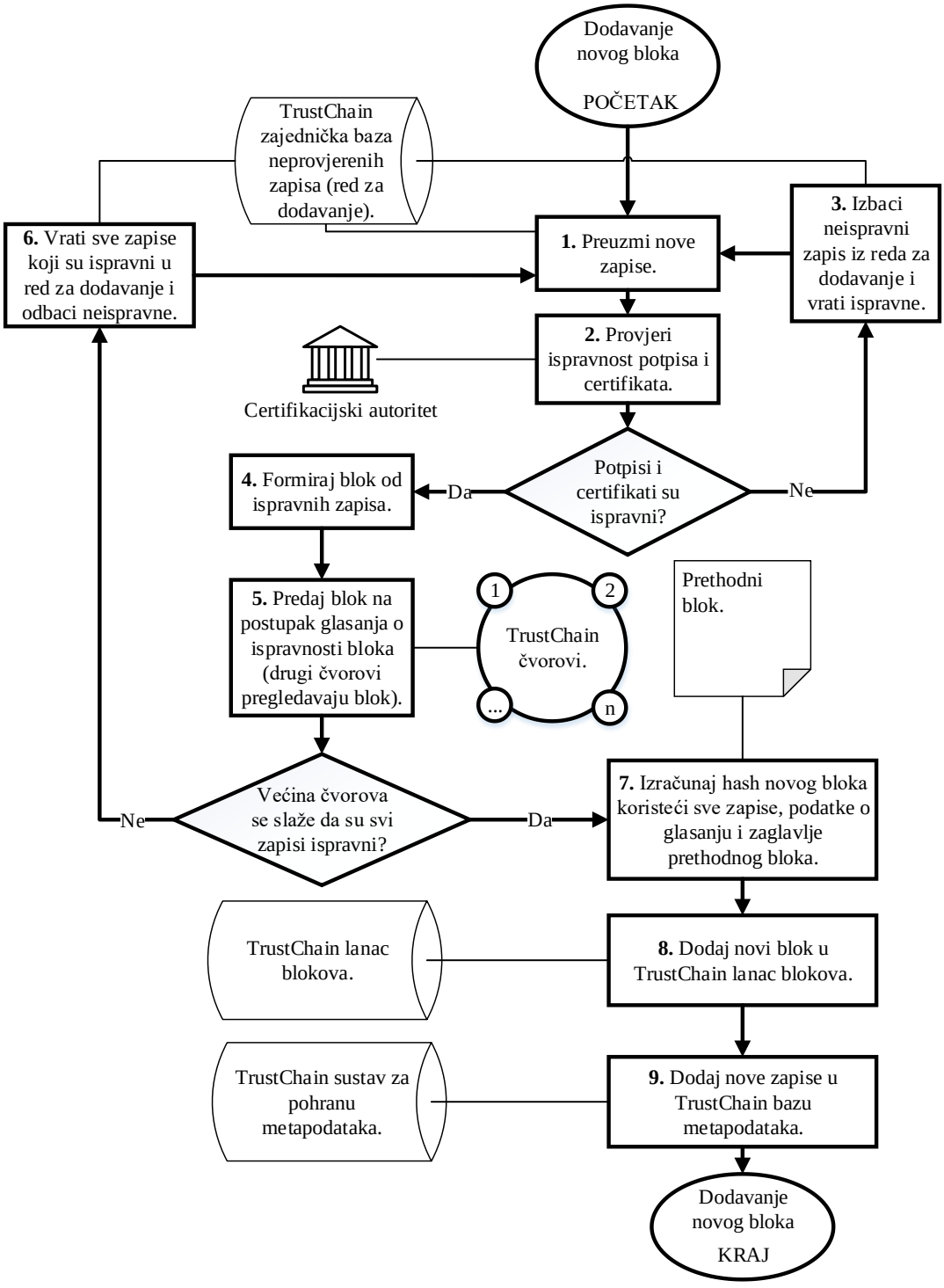
a) Dodana je funkcionalnost čitanja podataka iz zajedničke baze podatka. Izvorni model predvidio je da svaki čvor skuplja vlastite zapise dok ih se ne skupi dovoljno za cijeli blok kada čvor formira blok od svojih zapisa. Ovakav proces problematičan je u slučaju da čvor rijetko prima zapise jer je moguće da će predani zahtjevi predugo čekati na uvrštavanje u lanac blokova. Stoga je u novoj verziji uvedena zajednička baza neprovjerenih zapisa koji još čekaju na dodatne provjere ispravnosti i uvrštavanje u lanac blokova. Postupci TC2.1,, TC2.3. i TC2.6. su u interakciji s ovom bazom zapisa. Baza novih zapisa može biti, i bilo bi poželjno da je, realizirana upotrebom iste tehnologije za pohranu podataka kao i baza metapodataka. O ovoj bazi raspravlja se u kasnijem poglavlju koje se bavi strukturama podataka.

b) U postupku TC2.9. dodana je funkcionalnost zapisa metapodataka u za to predviđenu bazu podatka.

c) Postupak TC2.2 je promijenjen na način da naglasi da se provjerava i ispravnost certifikata, a ne samo potpisa kako je to slučaj u izvornom modelu. Ovaj naglasak pojašnjava raniju situaciju (certifikat je i prije trebalo provjeriti ali to nije bilo naglašeno) te omogućava da se isti postupak formiranja bloka novih zapisa koristi i za zapise koje je generirao TrustChain B modul. Iako su blokovi

[380] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.

iz modula drukčijeg sadržaja postupak njihovog dodavanja u blok je sada identičan te zapisi iz modula A i B mogu biti uvršteni u isti blok.



Slika 28. Proces TC2 – dodavanje novog bloka

Osim navedenih novosti neke postupke treba dodatno pojasniti. Proces stvaranja novog bloka počinje odlukom TrustChain čvora da je njegov red formirati novi blok prema ranije prikazanom postupku.

Postupak TC2.1. podrazumijeva preuzimanje, to jest čitanje (preuzimanje) i brisanje, zapisa iz zajedničke baze. Zapisi se odmah po preuzimanju brišu jer je, s obzirom da su već prošli jednu provjeru ispravnosti, pretpostavka da je većina njih ispravna. Kada ne bi bilo tako, bilo bi učinkovitije da se uz njih postavi zastavica koja označava da je čvor preuzeo zapise i da ih drugi čvorovi ne trebaju preuzimati. S obzirom na to će situacija u kojoj se zapisi vraćaju u ovaj red biti relativno rijetka ima smisla odmah ih izbaciti iz reda te naknadno, u rijetkom slučaju odbijanja nekog zapisa, vratiti u zajednički bazen kroz postupke TC2.3 i TC2.6.

Nakon što je čvor preuzeo nove zapise, u postupku TC2.2., ponavlja se postupak provjere ispravnosti digitalnog potpisa i certifikata. Ovaj proces istovjetan je provjeri iz procesa TCA1 i provodi se uz sudjelovanje vanjskog certifikacijskog autoriteta. Ponavljanje postupka je nužno jer ovaj čvor možda (vjerojatno) preuzima zapise koje su stvorili drugi čvorovi, a čak i u slučaju vlastitih zapisa određeno vrijeme je prošlo i to, iako je vjerojatno bilo vrlo kratko, može značiti da je istekao certifikat.

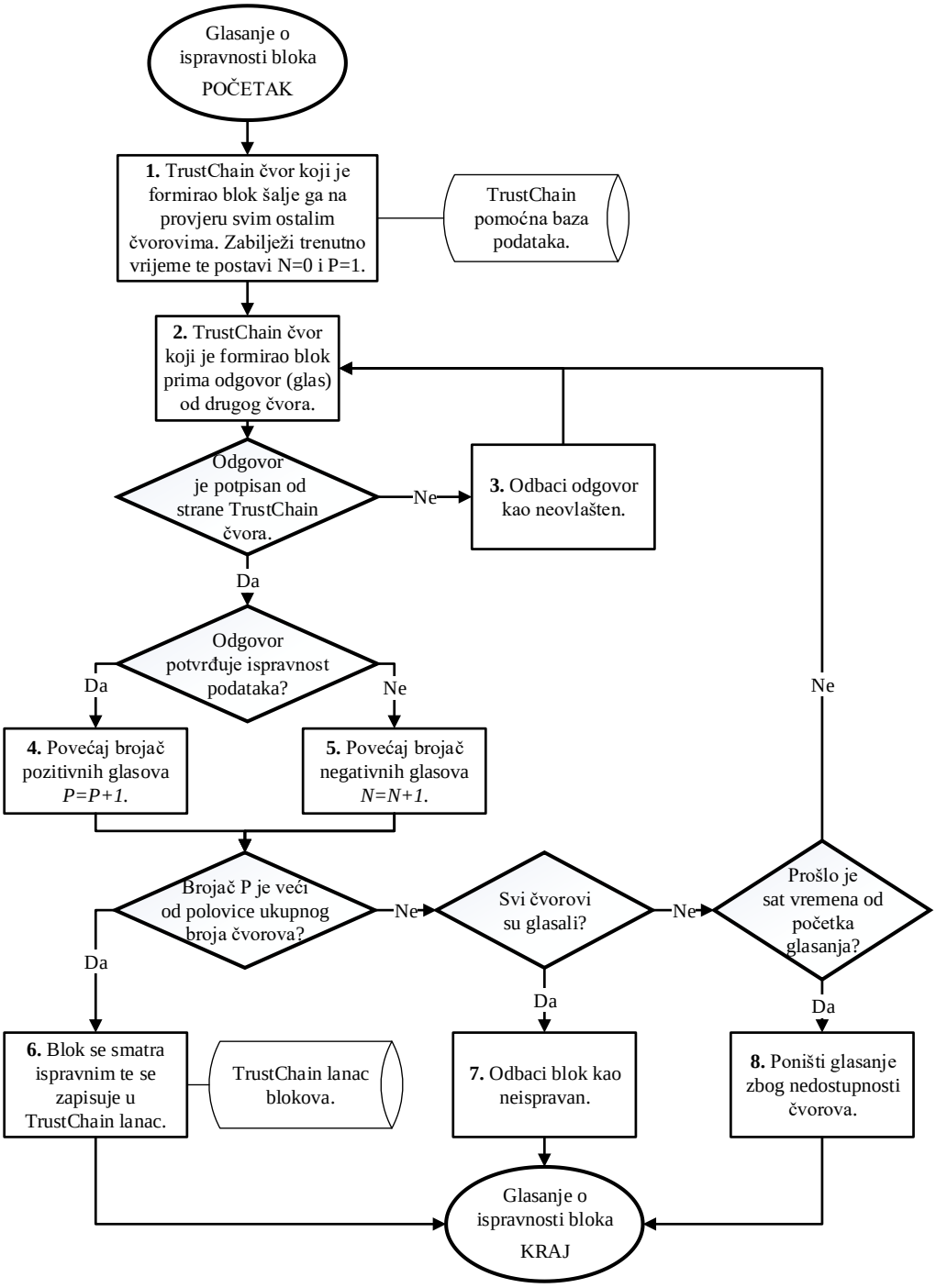
Postupak TC2.4. je najvažniji korak u ovom procesu, u njemu se stvara sam blok. Stvaranje novog bloka podrazumijeva skupljanje svih zapisa u jedinstvenu podatkovnu strukturu koju je moguće jednostavno hashirati i proslijediti na daljnju provjeru ostalim čvorovima. Struktura novog bloka je detaljnije pojašnjena u idućem poglavlju koje proučava sve podatkovne strukture u TrustChain A modulu.

Idući postupak, TC2.5., je apstrakcija kompliciranijeg postupka (TC3), koji je detaljno opisan u daljnjem tekstu. U ovom koraku traži se potvrda ostalih čvorova da je upravo formirani blok ispravan. Preostali čvorovi koji sudjeluju u TrustChain sustavu će provjeriti da je blok formiran prema zahtjevima sustava te da se sastoji samo od zapisa koji imaju ispravne potpise i/ili certifikate.

Postupak TC2.7. računa hash novog, potvrđeno ispravnog bloka koji se računa na osnovu hash-a trenutnog i prethodnog bloka. Na ovaj način garantira se trajni integritet podataka nakon što blok bude uvršten u TrustChain lanac blokova. Osim novog hasha u ovom koraku novi blok dobiva i svoj redni broj u TrustChain lancu. Ovaj redni broj je od izuzetne važnosti za zapis metapodataka u idućim koracima.

Iduća dva postupka (TC2.8. i TC2.9.) dodaju ranije pripremljene podatke u TrustChain lanac blokova (dodaje se sam blok, postupak TC2.8.) i u bazu metapodataka (zapisuju se svi dostupni metapodaci). Oba zapisa moguća su tek sada, kada je utvrđeno da je novi blok ispravan te da mu je dodijeljen redni broj koji se koristi pri zapisu metapodataka te u svojoj osnovi

indeksira zapisane podatke što omogućuje kasniju brzu pretragu lanaca blokova. Detaljni prikaz svih potpunih podatkovnih struktura koje su spomenute dan je u idućem poglavlju.



Slika 29. Proces TC3 glasanja čvorova

Detaljan opis procesa glasanja o novom bloka nije postojao u izvornom TrustChain A niti TrustChain B modelu, proces je do pisanja ove disertacije naveden kao apstrakcija uobičajenog procesa glasanja. Ovaj proces dizajniran je da bude jedinstven za oba TrustChain modula.

Nakon razrade procesa vidljivo je nekoliko točaka koje je potrebno dodatno raspraviti. Prvo, čvor koji je formirao novi blok odgovoran je i za vođenje procesa glasanja te je cijeli proces na slici 16 prikaz postupaka koje ovaj čvor provodi. Na ostalim, glasajućim čvorovima provodi se jednostavniji proces TC4 koji je prikazan, na slici 30.

Čvor koji je formirao blok, u postupku 1 kontaktira sve ostale TrustChain blokove prema podacima iz zajedničke baze podataka. Ova baza dodatak je bazi podataka koja je ranije spomenuta i čija je primarna funkcija pohrana metapodataka i pohrana privremenog reda zapisa koji još nisu dodani u lanac blokova. Osim kontaktnih podataka, najvjerojatnije domenske ili IP[381] adrese, čvor koji vodi postupak ovdje može preuzeti i relevantne certifikate, to jest javne ključeve koji će biti korišteni za potpisivanje odgovora. Osim slanja zahtjeva u ovim početnim postupcima potrebno je i postaviti dva brojača na početne vrijednosti. Brojač glasova koji nisu dali suglasnost počinje na 0, brojač suglasnih glasova počinje na 1 jer je čvor koji je formirao novi blok već provjerio njegov sadržaj. Ovi koraci opisani su postupkom TC3.1.

Ostali postupci u ovom procesu većim dijelom opisuju sami sebe pa većinu njih nema potrebe posebno objašnjavati. Postupci koje je jest potrebno dodatno raspraviti su dva zaključna postupka, postupak TC3.7. i TC3.8.

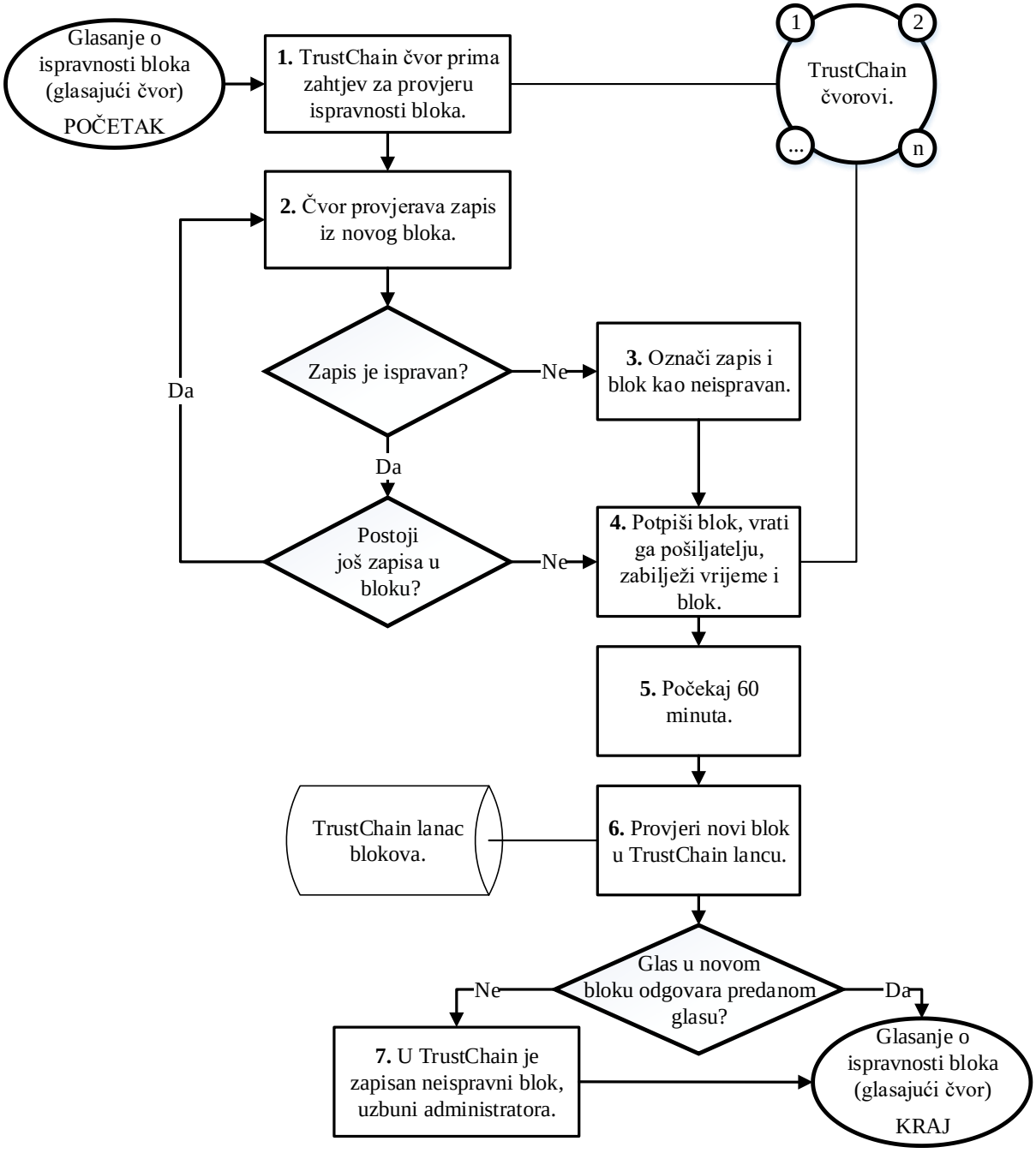
Postupak TC3.7., odbacivanje bloka kao neispravnog podrazumijeva ponovni pregled bloka i izbacivanje zapisa za koje se pokazalo da nemaju ispravan potpis ili certifikat te vraćanje ostalih u red za preuzimanje.

Postupak TC3.8. također završava proces glasanja bez dodavanja novog bloka, ali u ovom slučaju blok nije odbačen zbog neispravnih zapisa već zato što se većina čvorova uopće nije izjasnila. Ovo je indikacija značajnijeg problema nedostupnosti ili neresponzivnosti čvorova odnosno prekida komunikacije te sustav u ovom slučaju mora uzbuniti osoblje koje održava sustav. Moguće je da čvor koji vodi postupak ima problema u komunikaciji s ostalim čvorovima na mrežnoj razini ili da su ostali čvorovi nedostupni što može biti znak većeg komunikacijskog ispada (na razini Interneta) ili napada na TrustChain sustav. U svakom slučaju aktivacija postupka TC3.8. zahtjeva izvanrednu ljudsku intervenciju.

Za potpuni uvid u proces glasanja potrebno ga je sagledati i sa sastajališta glasajućih čvorova. Proces TC4, opisan na slici 30 ima dvostruku funkciju – osim provjere ispravnosti

[381] Internet Protocol – IPv4 RFC791, IPv6 RFC2460

bloka tijekom ovog procesa svaki čvor ujedno i provjerava da je njegov glas ispravno dodan u lanac blokova.



Slika 30. Proces TC4 – glasanje o ispravnosti bloka (glasajući čvor)

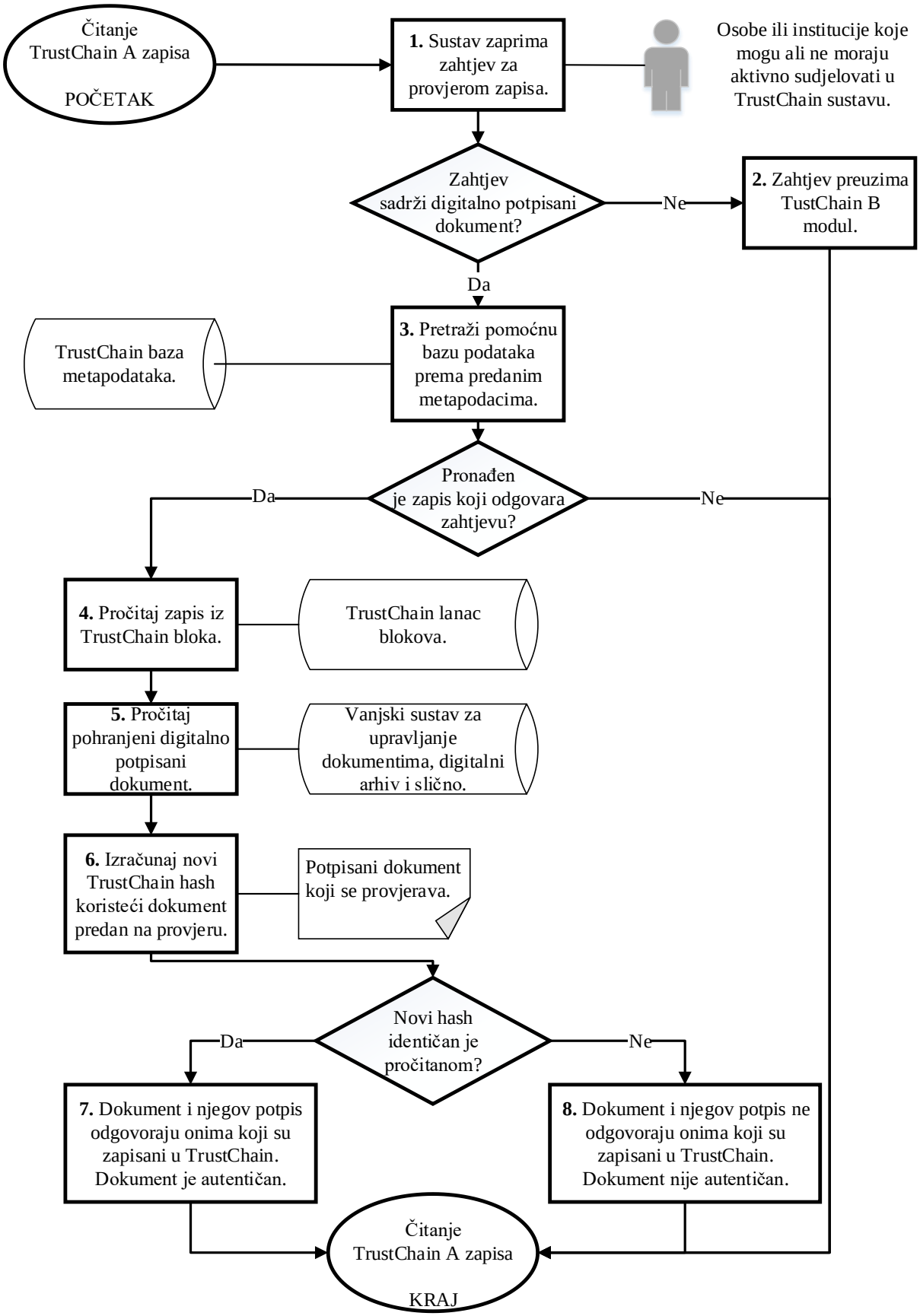
Proces TC4 je dio prethodnog proces koji se odvija na svim čvorovima osim onog koji je formirao blok. Postupak TC4.7. je jedni koji zahtjeva posebnu pažnju. Ovaj korak omogućuje konačnu provjeru ispravnosti bloka prije nego što blok postane trajni dio TrustChain lanca. U slučaju da je provjera koja prethodi postupku utvrdila da zapisani blok nije jednak onom koji je čvor predao prestaje automatizacija sustava te se uzbunjuje administrativno osoblje.

Najvjerojatniji uzrok ove neusklađenosti je napad na TrustChain sustav. Moguće je da je potpuno kompromitiran čvor koji je formirao blok, da je kompromitiran privatni ključ kojim ovaj čvor potpisuje odgovore ili da je na neki, vjerojatno dugotrajan, način komunikacija između čvorova kompromitirana. U ovom slučaju moguće je i automatski odbaciti blok te nastaviti s novim zapisom ali s obzirom na ozbiljnost omogućih uzroka (napadi na sustav) smatram da je, kao i u slučaju postupka TC3.8., riječ o iznimnoj situaciji koja zahtjeva ljudsku intervenciju. Iz tog razloga prikladnije je zaustaviti sustav dok se ne utvrdi točan uzrok problema.

Procesi glasanja o ispravnosti bloka završavaju razradu dodavanja novih zapisa u TrustChain A model. Preostalo je još razraditi proces čitanja postojećih zapisa. Ovaj proces se opet značajno razlikuje od izvornog modela iz 2017. prije svega zbog uvođenja baze metapodataka koja se prvi put pojavljuje u rad u časopisu Computers iz 2021. godine[382] i koja se detaljnije razrađuje u idućim poglavljima.

Proces čitanja zapisa ima za svrhu naknadno provjeriti ispravnost digitalno potpisanoga dokumenta te tako omogućiti potvrdu integriteta i autentičnosti digitalnih podataka. Ovakva potvrda je i osnovna svrha TrustChain sustava i razlog za stvaranje nepromjenjivog lanaca podataka. Proces čitanja i provjere ispravnosti zapisa je opisan na slici 31 te počinje pregledom baze metapodataka. Postupci pretrage baze metapodataka najveća su novost u odnosu na izvorni TrustChain A model[383]. Ova novo uvedena baza realizirana je poznatim tehnologijama, poput SQL i noSQL sustava za upravljanje bazama podatka, te omogućava pohranu i izmjenu metapodataka i brzi pronalazak bloka koji sadrži traženi zapis. Bez ovakvog indeksiranja pri svakom postupku čitanja bilo bi potrebno pretražiti cijeli TrustChain lanac, to jest trebalo bi provesti linearnu pretragu podataka. Izvorni model je ovaj postupak pretrage apstrahirao na jedan postupak "Pronalazak bloka koji sadrži traženi zapis". Osim ovih dodataka u novi proces dodana je i provjera sadrži li on digitalno potpisani dokument. Za razliku od ostalih novouvedenih procesa koji su jedinstveni za TrustChain A i TrustChain B model pri čitanju TrustChain B zapisa nužno je provesti drukčiji postupak provjere ispravnosti zapisa pa će za taj modul u idućem poglavlju biti potrebno uvesti zaseban proces provjere ispravnosti, to jest alternativu procesu prikazanom na slici 31.

[382] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.
[383] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.



Slika 31. Proces TCA2 – čitanje zapisa iz TrustChain A modela

6.1.2. TrustChain A podatkovne strukture

U ovom poglavlju se razrađuju neke od podatkovnih struktura koje su potrebne da bi se objasnili procesi opisani u prethodnom poglavlju. Poglavlje neće opisati sve povezane podatkovne strukture. Konkretno, bit će preskočena struktura TrustChain lanca blokova i podatkovne strukture u pomoćnoj bazi podataka kojima se bave kasnija poglavlja. Podatkovne strukture su na ovaj način podijeljene kroz više poglavlja da bi se naglasila razlika između TrustChain A i B modela koji čine temelje za ovdje prezentirane i dodatno razrađene modele i koncepte.

Sve podatkovne strukture u ovom poglavlju prikazane su kao JSON (engl. *JavaScript Object Notation*) datoteke[384]. XML (engl. *extensible markup language*) datoteke[385] su također razmotrene kao kandidat za prikaz podatkovnih struktura u modelu ali je prihvaćen JSON standard iz nekoliko razloga:

1) JSON je kraći i čovjeku čitljiviji.
2) JSON značajno jasnije prikazuje nizove podataka.
3) JSON se jednostavnije ugrađuje u sustave bazirane na JavaScript jeziku. Očekujem da će eventualna konkretna implementacija modela biti realizirana web aplikacijom i da će se kao takva djelomično oslanjati upravo na ovaj programski jezik.

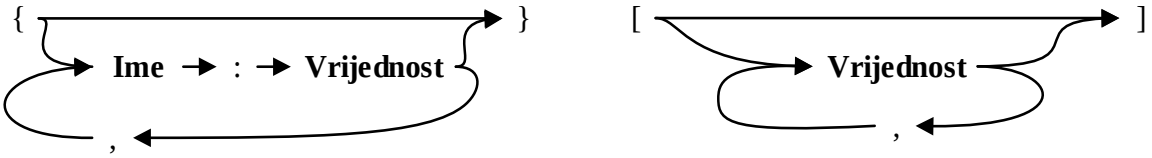
Dodatno, smatram da je za potrebe disertacije jasnoća prikaza modela prioritet.

S obzirom da je JSON standard odabran, ukratko će biti objašnjena i sintaksa ovog standarda. Osnova JSON notacije su parovi ime-vrijednost, između kojih se nalazi znak ":". Ovaj par se može opisati i kao naziv(opis)-vrijednost te u većini slučajeva sam sebe opisuje. Niz ovakvih parova odvaja se zarezima i nužno mora biti omeđen vitičastim zagradama. Ovakva notacija opisana je shemom na slici 32. Ovakav omeđen niz parova naziva se objekt. Dio para s desne strane, vrijednost, može biti i drugi objekt u kojem slučaju dobivamo ugniježđene objekte koji se obično uvlače u desnu stranu radi veće preglednosti. Osim objekata, JSON poznaje i nizove podataka (engl. *array*), koji se od objekta razlikuju po tome što su

[384] Internet Engineering Task Force. (2017). *RFC8259: The JavaScript Object Notation (JSON) Data Interchange Format*. (T. Bary, Urednik) Preuzeto 18. 11. 2021. s Internet Engineering Task Force (IETF): https://datatracker.ietf.org/doc/html/rfc8259

[385] World Wide Web Consortium. (2008). *Extensible Markup Language (XML) 1.0 (Fifth Edition).* Preuzeto 2. 5. 2020. s World Wide Web Consortium (W3C): https://www.w3.org/TR/2008/REC-xml-20081126/

omeđeni uglatim zagradama i sadrže samo vrijednost (bez imena). JSON niz je prikazan na desnoj strani slike 32. Vrijednost u ovakvom nizu podataka može biti i objekt (ili drugi niz).



Slika 32. Shema JSON notacije. Izvor: https://www.json.org/json-en.html

Bez obzira na odabrani način prikaza ovi načini zapisa su kompatibilni i lako prevodivi pa odabir jednog u jednom koraku ne isključuje upotrebu drugog u nekom kasnijem (što će i biti slučaj u nekim varijantama pomoćne baze podataka).

Programski kod 13 predlaže osnovnu podatkovnu strukturu za zapise stvorene upotrebom TrustChain A modula. Dakle, zapise koji referenciraju digitalno potpisanu datoteku koja je pohranjena u nekom vanjskom sustavu (na primjer u digitalnom arhivu ili sustavu za upravljanje dokumentima). Ova podatkovna struktura ostala je u obliku koji je gotovo istovjetan originalnom obliku u kojem je prezentirana 2017.[386]

Model se fokusira na datoteke ali je ovaj zapis prikladan za potpisane podatke u bilo kojem obliku (za rad s kojim je TrustChain sustav prilagođen).

U prikazanoj strukturi postoje dvije izmjene u odnosu na izvornik:

1) Dodano je ime `"type"`. Ovo ime označava tip zapisa. Za razliku od izvornog modela, sada je potrebno naglasiti da je riječ o zapisu stvorenom u TrustChain A modulu (a ne B). Zapisi iz B modula će se razlikovati po poljima koja sadrže u "record" zbirci.

2) Dodano je ime `"recordHash"`. Ovo ime sadrži hash zapisa, to jest zbirke `"record"`. Na ovaj način dodan je još jedan sloj sigurnosti zapisa te je postignuta struktura liste hasheva koja je objašnjena u četvrtom poglavlju.

[386] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.

Programski kod 13. TrustChain A JSON struktura zapisa

```
{
"id": "<id zapisa>",
"version": "<verzija modela>",
"type": "A",
"recordHash": "<hash zapisa>",
"record": {
            "timestamp": "<vrijeme stvaranja zapisa>",
            "certAuthName": "<naziv certifikacijskog autoriteta>",
            "certAuthID": "<id certifikacijskog autoriteta>",
            "certAuthApiLink": "<poveznica certifikacijskog autoriteta>",
            "data": {
                        "TrustChainHash": "<hash dokumenta>",
                        "docLnk": "<poveznica na dokument>" },
            "metadata": {
                        "docRefCode": "<šifra dokumenta>",
                        "docTitle": "<naziv dokumenta>",
                        "docCreator": "<naziv ili ime autora>",
                        "docCreationDate": "<datum stvaranja>" }
          }
}
```

Struktura zapisa prikazana u programskom kodu 13 sadrži `"metadata"` zbirku u kojoj su zapisani osnovni metapodaci o dokumentu na koji se zapisani hash odnosi. Izvorni TrustChain model odabrao je podatke sadržane u ovom polju prema ISAD(G) (engl. *General International Standard Archival Description*) nužnom setu metapodataka[387] Isti standard razmotren je i kasnije, 2021. godine, tijekom istraživanja vezanih uz premošćivanje problema vezanih uz nepromjenjivost ulančanih podatkovnih struktura[388] koje je rezultiralo razvojem pomoćne baze metapodataka. Zadržavanje ovih metapodataka u zapisu koji će biti zapisan u TrustChain lanac blokova, kako predlaže izvorni TrustChain A model, je redundantno ali namjerno. Iako model prezentiran u ovoj disertaciji uključuje i pomoćnu bazu podataka koja je specijalizirana za pohranu upravo ovakvih metapodataka, odlučio sam zadržati ova četiri osnovna metapodataka i u osnovnoj podatkovnoj strukturi zapisa jer:

1) Četiri dodatna i, u novom modelu, opcionalna niza znakova neće značajno opteretiti TrustChain lanac blokova.

2) Njihovo uključivanje u nepromjenjivu strukturu podataka (TrustChain lanac blokova) povećava pouzdanje u zapise u njoj. Ovaj minimalni (po ISAD(G)-u esencijalni) set metapodataka povećava sigurnost da se priloženi hash zaista

[387] Brothman, B. (1992). Isad (g): general international standard archival description. *Archivaria*, *34*, 17-32. Preuzeto 7. 1. 2022. s https://archivaria.ca/index.php/archivaria/article/view/11838/12790
[388] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.

odnosi na dokument na koji pokazuje poveznica. Zapise u pomoćnoj bazi metapodataka je moguće promijeniti te kao takvi predstavljaju ranjivost sustava.

3) Osim povećanja pouzdanja u vjerodostojnost zapisa ovi metapodaci pružaju i dodatnu sigurnost u slučaju potpunog gubitka pomoćne baze podatka. Ovakav gubitak mogao bi biti rezultat prirodne ili ljudski uzrokovane, namjerne ili nenamjerne, katastrofe.

4) Njihova prisutnost u TrustChain lancu blokova čini ovaj lanac samodostatnim. Pomoćna baza podataka je nužna za lakše funkcioniranje sustava i omogućava učinkovitu pohranu podataka vezanih uz arhivske veze ali ona nije krucijalna za osnovnu funkcionalnost TrustChain sustava. Na ovaj način, kada (ako) bi sustav bio umirovljen dovoljno je arhivirati (ili migrirati u novi sustav) samo lanac blokova, pomoćna baza može biti preskočena ako je njeno čuvanje problematično ili nemoguće (npr. zbog razloga iz točke 3).

Iz ovih razloga smatram da je zadržavanje osnovnog seta metapodataka u TrustChain lancu blokova opravdano.

Gornja podatkovna struktura rezultat je procesa TCA1 – formiranje novog TrustChain A zapisa. Skup ovakvih (i moguće TrustChain B zapisa) formira blok kandidat. Blok kandidat, podatkovna struktura prikazana programskim kodom 14 se nakon formiranja šalje ostalim TrustChain čvorovima na provjeru i glasanje (procesi TC2).

Programski kod 14. TrustChain JSON blok kandidat

```
{
"blockID": "<id bloka, redni broj>",
"nodeSig": "<potpis podataka>",
"block": {
          "previousBlockHash": "<hash prethodnog bloka>",
          "timestamp": "<vrijeme formiranja bloka>",
          "nodeID": "<id čvora koji formira blok>",
          "records": [
                    { <zapis1> },
                    { <zapis2> },...]}
}
```

Ova podatkovna struktura prvi put je opisna u ovoj disertaciji. Izvorni TrustChain A model je proces glasanja u potpunosti apstrahirao, tj. naveo primjer sličnog procesa koji se koristi u BigChainDB prijedlogu SQL (engl. *Structured Query Language*) baze podataka koja

koristi ulančane zapise kao garanciju integriteta podataka.[389] Sama struktura je pojednostavljena i modificirana verzija potpunog bloka koji je prezentiran u programskom kodu 16 te sastoji se od općih podataka o bloku i skupa svih zapisa koji ga čine (`"records"` zbirka). Novost u odnosu na izvorni model je dodatak imena `"nodeSig"` i pratećeg digitalnog potpisa čvora koji je formirao blok. Nedostatak ovog polja u izvornom modelu je rezultat potpune apstrakcije proces glasanja. Bez detaljnog razmatranja procesa glasanja ovaj nedostatak je bilo teže uočiti.

Po primanju bloka kandidata čvorovi postupaju prema uputama u procesu TC4 te proizvode podatkovnu strukturu opisanu programskim kodom 15. Ova JSON podatkovna struktura se šalje nazad čvoru koji formira blok te će, ako blok bude prihvaćen, biti uključena u njega i trajno pohranjena u TrustChain lanac blokova.

Programski kod 15. TrustChain JSON podatkovna struktura glasa o ispravnosti bloka

```
{
"nodeID": "<id TrustChain čvora>",
"nodeSig": "<potpis podataka>",
"vote": {
          "blockCandidate": "<id bloka o kojem se glasa>",
          "is_block_valid": "<true | false >",
          "timestamp": "<vrijeme glasanja>"
          "record_votes": [
                          "record_id": "<true | false >",
                          "record_id": "<true | false >",
                          "record_id": "<true | false >",...]}
}
```

Podaci o glasanju bili su prisutni u izvornom TrustChain A modelu, a podatkovna struktura je ovdje proširena na način da se dodao niz podatka `"record_votes"` koji se sastoji od identifikacijskih brojeva svih zapisa koji čine blok, to jest čiju je ispravnost provjerio čvor koji formira podatke o glasanju. Ovo je značajno povećalo veličinu podataka o glasanju ali omogućava učinkovito uočavanje zapisa za koje ovaj čvor smatra da su neispravni. Ova optimizacija je važna za kasnije korake jer je u slučaju odbijanja bloka lakše prepoznati problematične zapise, a u slučaju prihvaćanja bloka (npr. suprotno glasu ovog čvora) ne gubi se informacija o (prema ovom čvoru) problematičnim zapisima (već je ona trajno zapisana).

[389] McConaghy et al. (2016). *Bigchaindb*, n. dj.

Konačna JSON podatkovna struktura opisana u ovom poglavlju je sadržaj bloka koji više nije kandidat već sadrži potreban broj glasova i spreman je za dodavanje u TrustChain lanac blokova. Ova podatkovna struktura opisana je u programskom kodu 167.

Programski kod 16. TrustChain JSON podatkovna struktura ulančanog bloka

```
{
"blockHash": "<hash ovog bloka>",
"blockID": "<id bloka, redni broj>",
"block": {
          "previousBlockHash": "<hash prethodnog bloka>",
          "timestamp": "<vrijeme formiranja bloka>",
          "nodeID": "<id čvora koji formira blok>",
          "nodeSig": "<potpis podataka>",
          "records": [
                    { <zapis1> },
                    { <zapis2> },...],
          "votes": [
                    { <glas1> },
                    { <glas2> },...]}
}
```

Gornji blok, prikazan u JSON formatu, sastoji se od svih podataka koje treba pohraniti u novi TrustChain blok. Nizovi `"records"` i `"votes"` sastoje od podatkovnih struktura prikazanih u programskom kodu 14 i 15. Ova podatkovna struktura, u potpunom obliku, kada bi se blok sastojao od samo jednog zapisa i samo jednog glasa, prikazana je niže u programskom kodu 17. Kao i u slučaju podataka koji su se slali na glasanje dodan je potpis čvora koji formira blok (i ostali dodaci koji su preneseni ugrađivanjem u ovu podatkovnu strukturu.

Programski kod 17. TrustChain JSON blok potpuni prikaz

```
{
"blockHash": "<hash ovog bloka>",
"blockID": "<id bloka, redni broj>",
"block": {
      "previousBlockHash": "<hash prethodnog bloka>",
      "timestamp": "<vrijeme formiranja bloka>",
      "nodeID": "<id čvora koji formira blok>",
      "nodeSig": "<potpis podataka>",
      "records": [{"id": "<id zapisa>",
            "version": "<verzija modela>",
            "type": "A",
            "recordHash": "<hash zapisa>",
            "record": {
                  "timestamp": "<vrijeme stvaranja zapisa>",
                  "certAuthName": "<naziv certifikacijskog autoriteta>",
                  "certAuthID": "<id certifikacijskog autoriteta>",
```

```
                    "certAuthApiLink": "<poveznica cert. autoriteta>",
                    "data": {
                          "TrustChainHash": "<hash dokumenta>",
                          "docLnk": "<poveznica na dokument>"},
                    "metadata": {
                          "docRefCode": "<šifra dokumenta>",
                          "docTitle": "<naziv dokumenta>",
                          "docCreator": "<naziv ili ime autora>",
                          "docCreationDate": "<datum stvaranja>"}}},...],
      "votes": [
            {"nodeID": "<id TrustChain čvora>",
            "nodeSig": "<potpis podataka>",
            "vote": {
                    "blockCandidate": "<id bloka o kojem se glasa>",
                    "is_block_valid": "<true | false >",
                    "timestamp": "<vrijeme glasanja>"},
                    "record_votes":[
                          "record_id": "<true | false >",...]}},...]
}
```

Ovakav zapis može se dodati kao novi blok u TrustChain lancu ulančanih blokova. S ovim je zaključeno poglavlje o podatkovnim strukturama u TrustChain A modelu s napomenom da su sve podatkovne strukture, osim prve – JSON zapisa o individualnom dokumentu, univerzalne te će biti iskorištene i u TrustChain B modelu. Osim ovoga treba se podsjetiti da nije raspravljena struktura metapodataka za pomoćnu bazu koji se formiraju u TrustChain A modulu.

6.2. TrustChain B

TrustChain B model proizašao je iz razrade izvornog modela tijekom 2018. i 2019. godine u nastavku InterPARES Trust projekta.[390] Novi model istražen je nakon što je uočeno ograničenje izvornog, TrustChain A modela, povezano s povjerljivim arhivskim gradivom. S obzirom na to da je TrustChain sustav zamišljen kao međunarodna suradnja više arhivskih (ili drugih zainteresiranih) ustanova te da njegov sustav glasanja o ispravnosti dokumenta (i digitalnog potpisa) podrazumijeva da sve sudjelujuće ustanove (ili barem većina njih) pregledaju dokument sustav nije prikladan za pregled povjerljivih dokumenta koje mora pregledati više ustanova iz, vjerojatno, više država. Ovo ograničenje sigurno predstavlja značajan problem državnim arhivima koji moraju rukovati s povjerljivim dokumentima, ali može biti problem i privatnim osobama ili tvrtkama koje žele iskoristiti TrustChain sustav da

[390] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving, n. dj.

bi dugotrajno osigurale dokazivost autentičnosti povjerljivog dokumenta, npr. poslovnog ugovora. TrustChain A model u svojoj osnovnoj premisi pretpostavlja provjeru ispravnosti certifikata i samog potpisa konkretnog dokumenta te kao takav nema načina za zaobilazak ovog problema.

Mats Stengård, suradnik na InterPARES Trust projektu primijetio je da provjera digitalnog potpisa, to jest cijelog dokumenta možda nije potrebna. Minimalna provjera koja će osigurati autentičnost (u smislu ispravnosti certifikata, to jest ispravnosti podataka o autorstvu nekog zapisa ili dokumenta) podrazumijeva samo provjeru certifikata. Sam potpis garantira identitet potpisnika i integritet potpisanih podataka. Bez višestruke (međunarodne) provjere cijelog dokumenta TrustChain ne može garantirati da je potpis ispravan, ali može provjeriti certifikat i njega pohraniti u TrustChain lanac blokova. Na ovaj način trajno je osiguran digitalni certifikat te se on može koristiti u kombinaciji s arhiviranim dokumentima da bi se dokazalo da je (barem) certifikat bio ispravan u prošlosti. Ovo pruža manju razinu sigurnosti od TrustChain A modela, ali ipak dostatnu za slučajeve kada bolja nije moguće (što je slučaj s povjerljivim dokumentima).

Ovakav model, nazvan TrustChain 2.0, predstavljen je 2020. godine u časopisu Records Management Journal[391] te je u ovom radu iskorišten kao osnova za TrustChain B modul. Kao dio većeg sustava ovaj modul je u interakciji s TrustChain A modulom jer mu se u novom modelu predaju svi certifikati, uključujući one proizašle iz dokumenta predanih na provjeru TrustChain A modulu. Dakle, osim što omogućava, djelomično, osiguranje autentičnosti dugoročno pohranjenih povjerljivih dokumenata modul nadograđuje funkcioniranje TrustChain A modula.

6.2.1. TrustChain B procesi

Ovo poglavlje razrađuje najvažnije procese vezane uz rad TrustChain B modula. Slično prethodnom poglavlju i ovdje se nastavlja ranije započet način označavanja procesa. Procesi vezani uz TrustChain B modul počet će s oznakom TCB koja je popraćena rednim brojem procesa. Za razliku od ranije spomenutog procesa, ovdje će biti prikazani samo procesi koji su specifični za ovaj modul, a ranije prikazani univerzalni procesi neće biti ponovljeni.

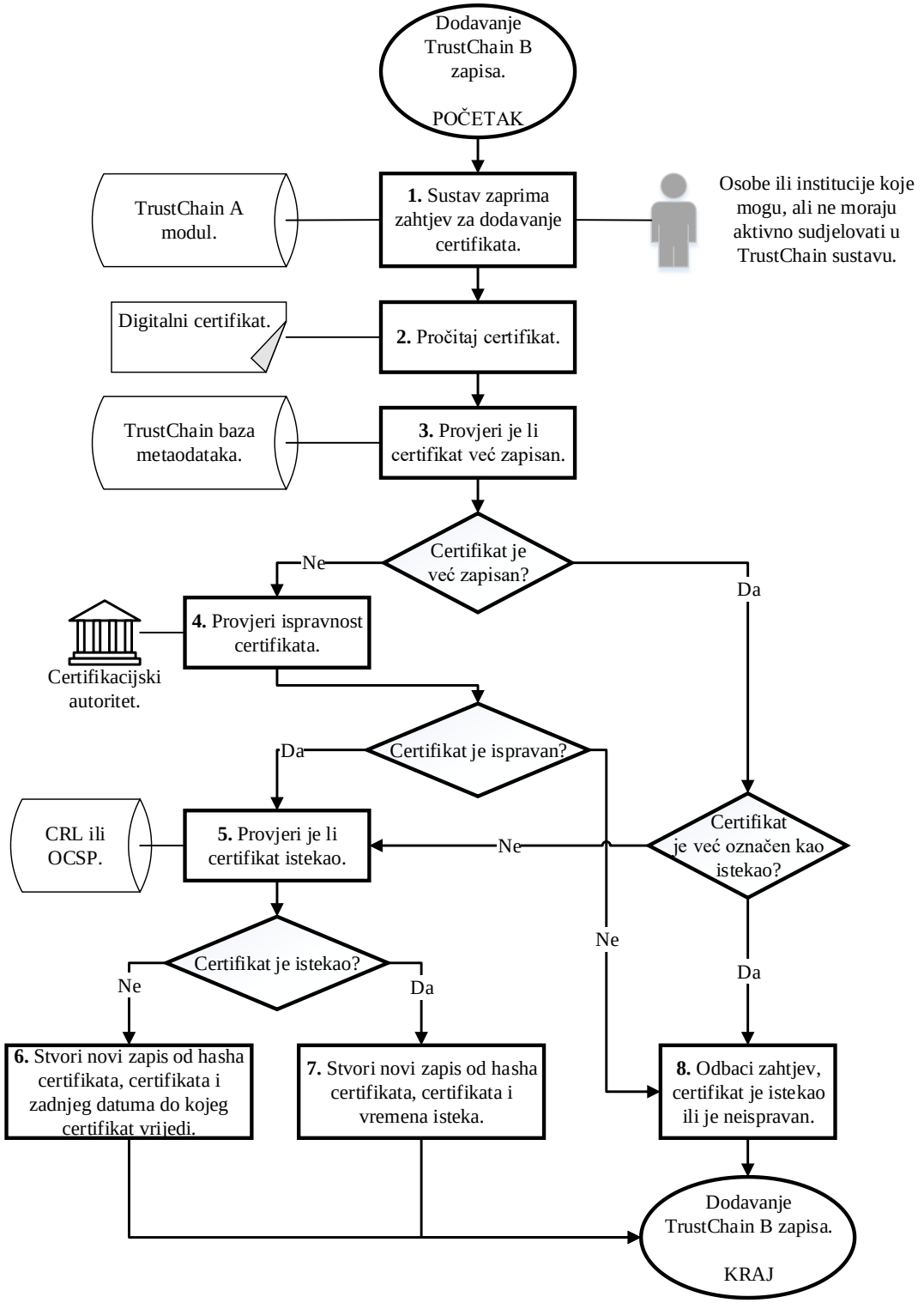
[391] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving, n. dj.

Za razumijevanje procesa iz TrustChain B modula važno je raspraviti nekoliko termina. Većina njih je detaljnije obrađena u poglavlju koje se bavi tehnološkom osnovom TrustChain modela, prije svega digitalnim certifikatima, ali svejedno smatram svrsishodnim pojasniti dvije skraćenice, CLR i OCSP koje su vezane za revokacijske liste, tj. liste opozvanih certifikata. Ove skraćenice koriste se od samog početka opisa procesa TrustChain B modula. Revokacijska lista označava popis koji izdaju certifikacijski autoriteti u kojem su navedeni certifikati koji su iz nekog razloga povučeni prije vremena isteka. Kao takve, revokacijske liste su izuzetno važan mehanizam u TrustChain B modulu jer je on odgovoran za provjeru ispravnosti i pohranu digitalnih certifikata. Dok je revokacijska lista jasan i dobro poznat koncept, mehanizmi njene distribucije su manje poznati. Revokacija certifikata provodi se na dva načina:

- Dohvaćanjem revokacijske liste, skraćeno zvane CRL (engl. *Certificate Revocation List*), od certifikacijskog autoriteta.
- Upotrebom OCSP protokola (engl. *Online Certificate Status Protocol*), protokola specijaliziranog za provjeru revokacijskog statusa certifikata (Santesson, i dr., 2013).[392]

Proces kojim počinje opis TrustChain B modula, TCB1, opisuje dodavanje novog zapisa u blok kandidat i prikazan je na slici 33.

[392] Santesson, S., Myers, M., Ankney, R., Malpani, A., Galperin, S., & C., A. (2013). *RFC 6960*, n. dj.



Slika 33. Proces TCB1 – dodavanje novog zapisa u TrustChain B modulu. Izvor: V Bralić, H Stančić, M Stengård. A blockchain approach to digital archiving: digital signature certification chain preservation. 2020.

Proces TCB1 fokusiran je provjeru ispravnosti certifikata koji je predan na provjeru TrustChain sustavu. U odnosu na proces prikazan 2020. ovdje je najveći dodatak upotreba pomoćne baze podataka, koja je navedena u postupku TCB1.3.

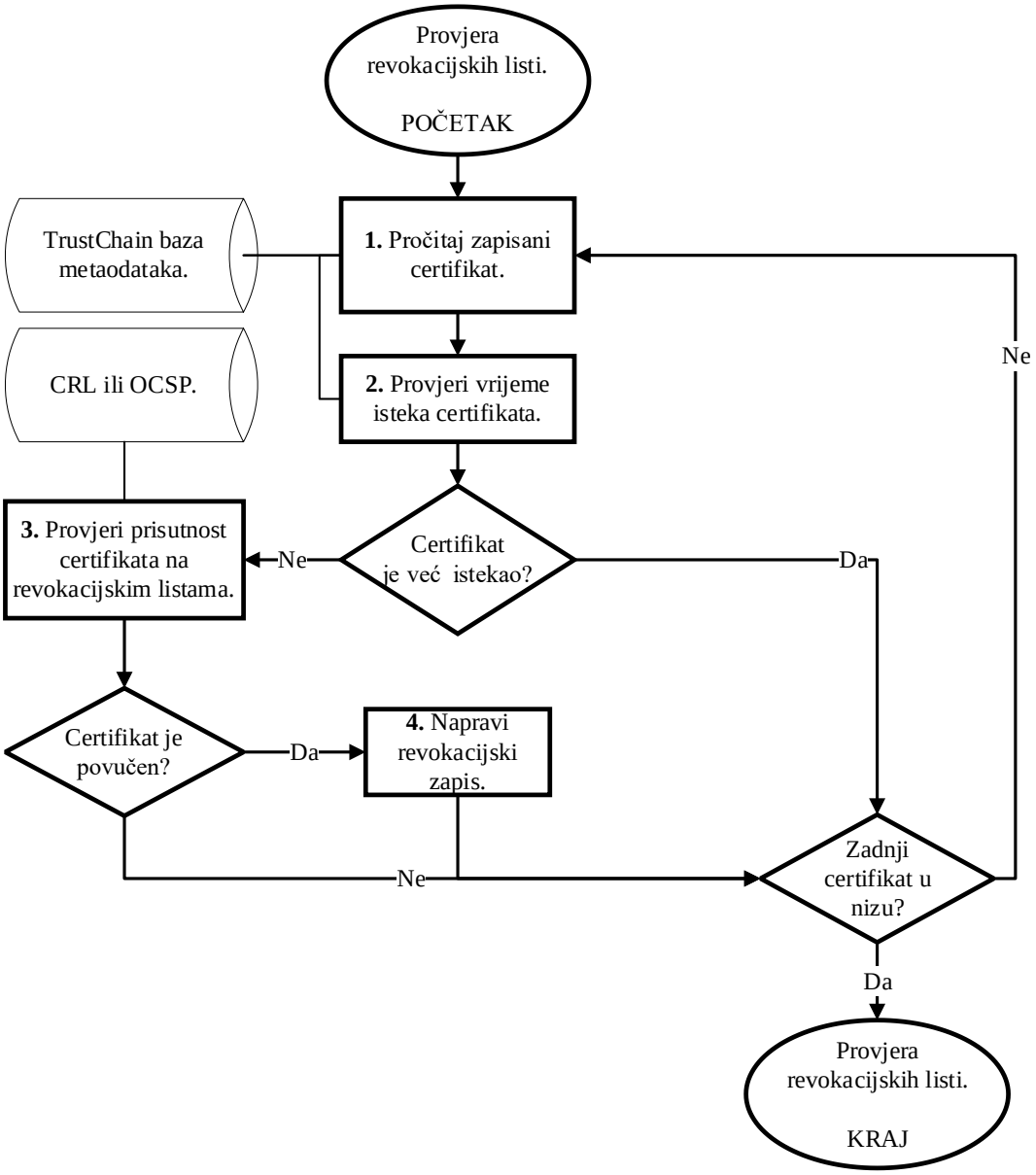
Potrebno je dodatno pojasniti postupke TCB1.6. i TCB1.7., to jest razliku između njih. Ovi postupci na prvi pogled izgledaju vrlo slično, ali se razlikuju po konačnom rezultatu, to jest po opsegu podataka koji će biti sadržani u konačnom zapisu. Proces TCB1.6. podrazumijeva zapis novog (za TrustChain sustav) certifikata koji još nije istekao, stoga će uključiti redovno vrijeme isteka certifikata. Za razliku od njega, proces TCB1.7. rukuje s (opet iz perspektive TrustChain sustava) novim certifikatom koji još nije istekao, ali se već našao na revokacijskim listama. U ovom slučaju u konačni zapis o certifikatu dodaje se i revokacijski zapis. Ovi zapisi su se mogli odraditi na identičan način, ali smatram da je činjenica da je certifikat povučen značajna informacija koju treba sačuvati. S obzirom da je revokacija, to jest ranije povlačenje certifikata posljedica izvanrednog događaja, najčešće krađe privatnog ključa vezanog uz certifikat ili pak (izvanrednog) prestanka dužnosti pojedine osobe koja u tom svojstvu ima mandat potpisivati dokumente, važno je da sustav koji trajno pohranjuje certifikate sačuva tu informaciju. Iako pregled podataka o certifikatima nije osnovna svrha TrustChain sustava, TrustChain B modul je stvorio upravo tu mogućnost, to jest pretvorio je sustav u arhiv digitalnih certifikata. Takav arhiv se u budućnosti može koristiti u više svrha te iz tog razloga sustav na ovaj način čuva podatak da je certifikat opozvan prije isteka njegovog normalnog životnog vijeka.

Osim ovih, potrebno je razraditi i postupak TCB1.8. Ovaj postupak, koji označava kraj procesa bez dodavanja novog zapisa, može biti rezultat dvaju značajno različitih uzroka. Uzrok ovome može biti neispravan certifikat ili certifikat koji je istekao, što je već naznačeno u TrustChain sustavu. U oba slučaja rezultat je isti – zahtjev za stvaranje novog zapisa se odbija bez obzira je li uzrok neispravan certifikat ili certifikat koji je već u potpunosti dokumentiran u TrustChain sustavu kao istekao. Dakle, već postoji zapis o samom certifikatu kao i o njegovom sada već prošlom datumu isteka (bez obzira je li on uzrokovan revokacijom ili ne).

Rezultat procesa TCB1 je individualni zapis, ekvivalentan rezultatu TCA1 procesa iz TrustChain A modula (slika 26). Ovaj zapis (o ispravnosti certifikata) je razrađen u idućem poglavlju.

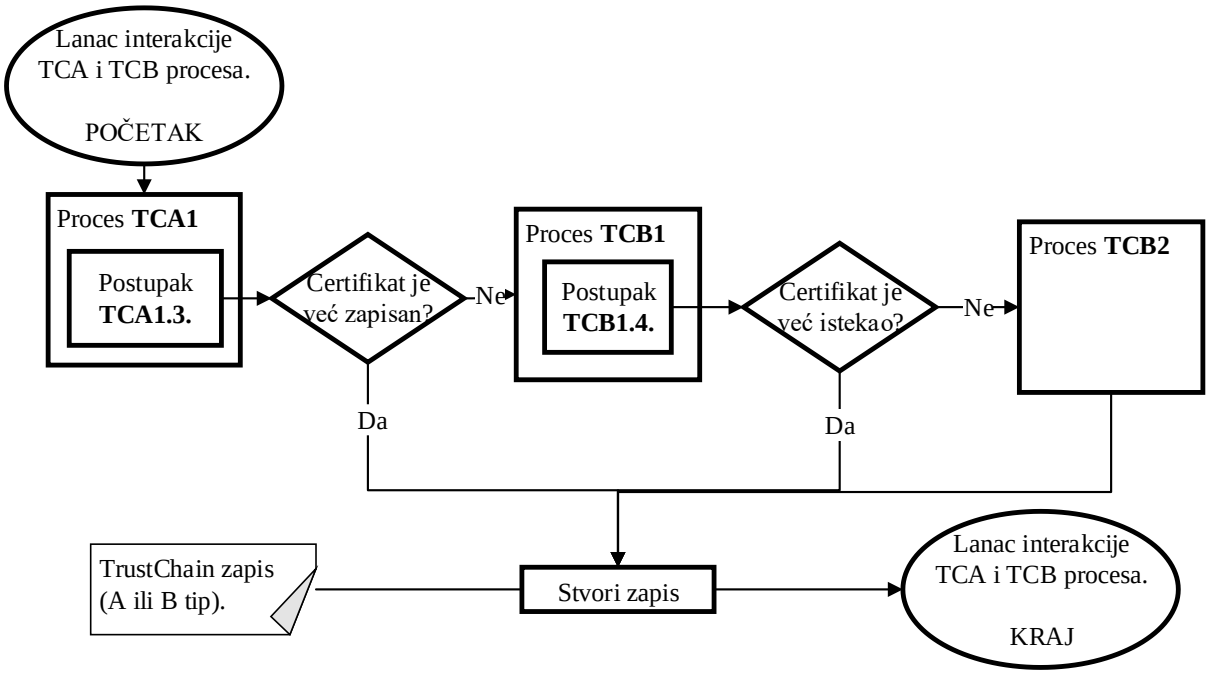
Proces TCB2, prikazan na slici 34, opisuje postupke čiji je rezultat pregled revokacijskih listi s ciljem pronalaženja isteklih certifikata koji su već zapisani u sustav ili su

upravo u postupku zapisivanja. Ovaj proces započinje kao dio procesa TCB1, konkretno njegovog postupka TCB1.5. ili, opcionalno, kao samostalni proces u kojem se provjerava prisutnost certifikata na revokacijskim listama. Iz ovog razloga proces je dizajniran na način da može provjeriti jedan ili više certifikata. Osim toga proces je na ovaj način dizajniran zbog samog načina funkcioniranja revokacijskih listi. Revokacijske liste je potrebno ili povremeno provjeravati ili u trenutku provjere ispravnosti certifikata odraditi i pregled revokacijskih listi. U slučaju poziva iz procesa TCB1 proces će svakako provjeriti samo jedan certifikat. U slučaju da se poziva samostalno, proces TCB2 može pregledati i skupinu certifikata. Pokretanje postupka na ovaj način biti će opisano u konačnom poglavlju o TrustChain modelu.



Slika 34. Proces TCB2 – provjera revokacijskih listi. Izvor: V Bralić, H Stančić, M Stengård. A blockchain approach to digital archiving: digital signature certification chain preservation. 2020.

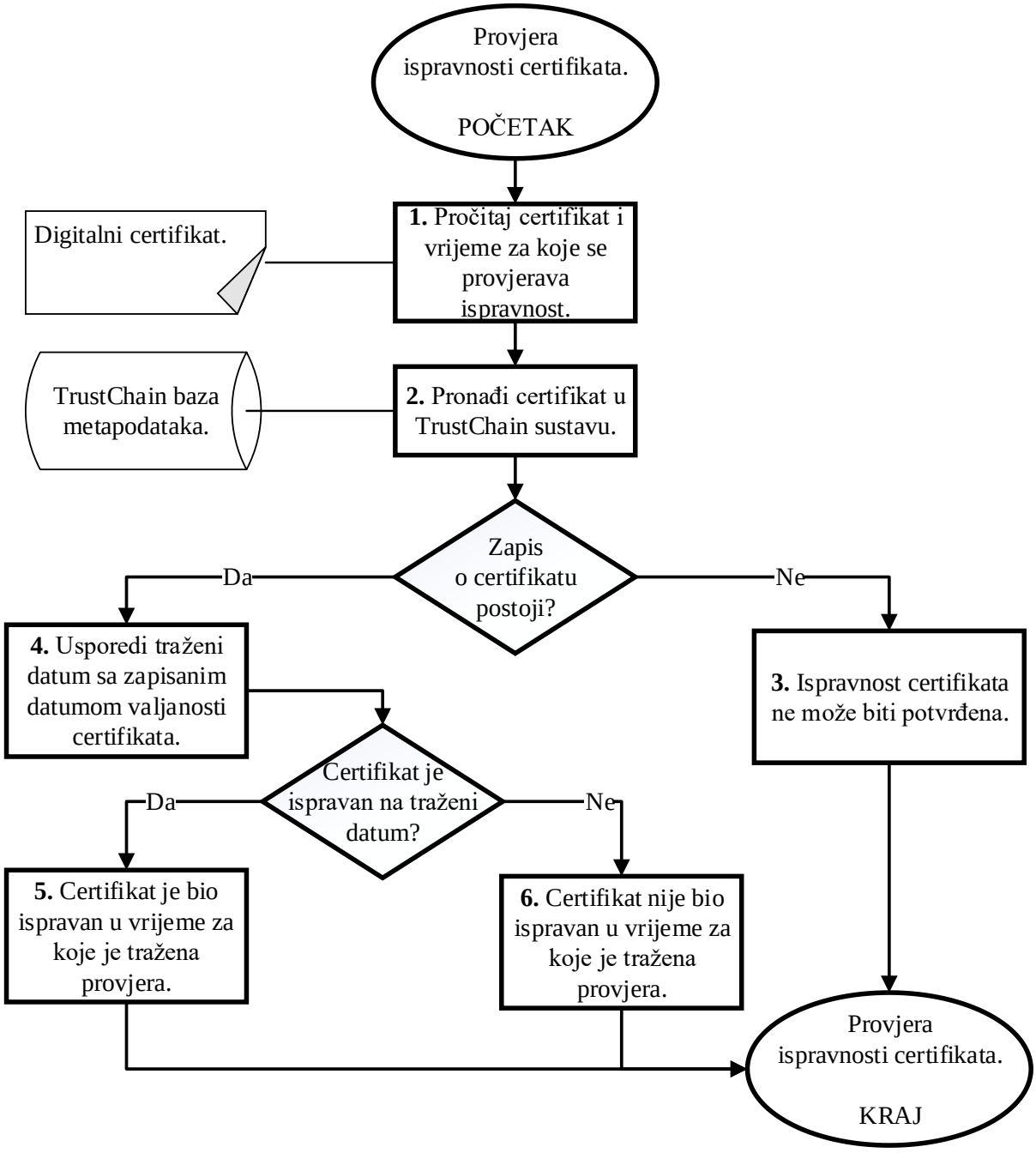
S obzirom na to da je proces TCB1 možda pokrenut u tijeku rada procesa TCA1 stvara se lanac međusobno povezanih procesa koji je prikazan na slici 35. Interakcija TrustChain A i B modula je ostavljena za posljednje poglavlje koje se bavi, novim, integriranim modelom, ali s obzirom na to da u su ovom slučaju procesi u izravnoj ovisnosti ova veza je razrađena već u ovom poglavlju.



Slika 35. Interakcija procesa iz TrustChain A i B modula

Osim procesa TCA1 i TCB1 u nekim slučajevima u ovom lancu može sudjelovati i proces TCB2, koji provjerava prisutnost certifikata u revokacijskim listama. Najduža varijanta ovog lanca, koja uključuje sva tri procesa, je slučaj u kojem je TrustChain A modulu predan zahtjev za pohranu dokumenata koji sadrži ispravan certifikat koji još nije zapisan kao istekao.

Zadnji postupak koji je karakterističan za TrustChain B modul je postupak TCB3, koji omogućuje provjeru ispravnosti zapisanog certifikata. U ranijem tekstu, kada se govorilo o provjeri ispravnosti certifikata podrazumijevalo se provjeru ispravnosti i isteka s certifikacijskim autoritetom. U slučaju procesa TCB3 govorimo o provjeri je li konkretni certifikat identičan certifikatu koji je zapisan u TrustChain lanac blokova te je li certifikat bio valjan na neki, konkretni, datum u prošlosti. Ovaj proces ekvivalentan je procesu TCA2. Oba procesa koriste podatke zapisane u TrustChain lancu blokova da bi dokazali ispravnost nekog dokumenta ili certifikata, za razliku od ostalih prikazanih procesa koji sudjeluju u stvaranju novih zapisa. Proces je prikazan na slici 36.



Slika 36. Proces TCB3 – provjera ispravnosti certifikata. Izvor: V Bralić, H Stančić, M Stengård. A blockchain approach to digital archiving: digital signature certification chain preservation. 2020.

Prikazani procesi omogućuju dodavanje i čitanje, to jest provjeru zapisa o certifikatima u TrustChain lancu blokova. Ovi procesi su srž TrustChain B modula koji proširuje funkcioniranje TrustChain sustava na način da omogućava pohranu i provjeru digitalnih

certifikata bez potrebe da se distribuiraju i provjeravaju sami potpisi te tako omogućava da sustav pomogne u dugotrajnom očuvanju povjerljivih digitalno potpisanih dokumenata.

6.2.2. TrustChain B podatkovne strukture

Podatkovne strukture povezane uz TrustChain B modul prikazane su kao i ranije, u JSON obliku. Kao što je bio slučaj i s procesima, univerzalne podatkovne strukture, prvenstveno one vezane uz glasanje ovdje nisu ponovljene već se uvode dvije nove, specifične za TrustChain B modul te se nadograđuje ranije prikazana struktura TrustChain bloka (programski kod 18).

Poglavlje počinje JSON opisom podatkovne strukture osnovnog TrustChain B zapisa. Za razliku od A modula svrha ovog zapisa, koji je prikazan u programskom kodu 18 je očuvanje podataka o certifikatima s kojima je TrustChain B modul imao doticaja. U novom TrustChain modelu, koji obuhvaća funkcionalnosti obaju prethodnih modela, ovo podrazumijeva sve certifikate s kojima je sustav imao doticaj.

Programski kod 18. TrustChain B modul JSON struktura zapisa

```
{
"id": "<id zapisa>",
"version": "<verzija modela>",
"type": "B",
"recordHash": "<hash zapisa>",
"record": {
           "timestamp": "<vrijeme stvaranja zapisa>",
           "certData": {
                     "CertificateID": "<ID certifikata>",
                     "CertificateHash": "<hash certifikata>",
                     "ValidityStart": "<vrijeme od kada vrijedi>",
                     "ValidityEnd": "<vrijeme do kada vrijedi>",
                     "CertificateOwner": "<naziv ili ime vlasnika>",
                     "CertificateIssuerID": "<CA identifikator>" },
           "RevocationData": {
                     "RevocationTime": "<vrijeme revokacije>",
                     "CRLlink": "<id revokacijske liste>" }
            }
}
```

Sami certifikati nisu pohranjeni u TrustChain lancu blokova. Kao što je TrustChain A model u lanac blokova pohranjivao samo hash provjerenih dokumenata tako i TrustChain B u lanac pohranjuje samo hash certifikata. S obzirom na to da su digitalni certifikati, zbog svoje

svrhe, nužno javni podaci (osim privatnog ključa) njih može provjeriti više TrustChain čvorova (ustanova) te mogu biti trajno pohranjeni (za razliku od dokumenata iz A modula koji su se čuvali u vanjskim repozitorijima). Iako je digitalni certifikat zapis definirane veličine koji nije memorijski zahtjevan, zbog velikog broja certifikata koji će biti provjereni (i zapisani) odlučio sam certifikat ne pohraniti u TrustChain lanac blokova. Umjesto toga, certifikati mogu biti pohranjeni u pomoćnoj bazi koja je raspravljena u idućem poglavlju, a u lancu blokova samo njihov hash s osnovnim metapodacima. Zapis prikazan u programskom kodu 18 navodi na svojem kraju zbirku `"RevocationData"`. Iako se možda čini nelogičnim uključiti podatke, koji se u većini slučajeva dodaju naknadno, u nepromjenjivu podatkovnu strukturu, TrustChain B modul je dizajniran da zabilježi i već istekle digitalne certifikate. Ova funkcionalnost omogućava širi opseg rada TrustChain B modula kao arhiva digitalnih certifikata. Iz ovog razloga u osnovnu podatkovnu strukturu je uključena ova, opcionalna, zbirka koja omogućava da se u istom zapisu označi da je riječ o isteklom certifikatu. Na ovaj način je poboljšanja učinkovitost kasnijih pretraga TrustChain lanca blokova – u slučaju certifikata koji su kasnije povučeni bit će potrebno naći blok sa zasebnim zapisom o opozivu. Ovaj zapis je uvijek rezultat procesa TCB2 (provjera isteka certifikata) koji je pozvan od strane procesa TCB1 (stvaranje novog TrustChain B zapisa).

Podatkovna struktura spomenutih zasebnih podataka o revokaciji certifikata čiji zapis već postoji u TrustChain lancu blokova prikazana je u JSON programskom kodu 19. Ovo je zapis koji je rezultat procesa TCB2 pozvanog pri naknadnoj provjeri isteka certifikata za već zapisani certifikat. Naknadna provjera podrazumijeva da izvorni zapis o certifikatu više nije moguće promijeniti te se ovaj naknadno stvoreni zapis dodaje u kasniji blok i u zapisu o certifikatu u pomoćnoj bazi podataka se stvaraju nove poveznice prema ovom zapisu. Poveznica prema bloku sadržana je i u samom zapisu o revokaciji certifikata. Ovaj postupak sličan je načinu na koji TrustChain model održava arhivsku vezu.

Programski kod 19. TrustChain B struktura zapisa podataka o opozivu certifikata

```
{
"RevocationId": "<id zapisa>",
"RevocationRecordHash": "<hash zapisa>",
"RevocationRecord": {
          "timestamp": "<vrijeme stvaranja zapisa>",
          "RevocationTime": "<vrijeme revokacije>",
          "CRLlink": "<id revokacijske liste>",
          "BlockID": "<id bloka koji sadrži zapis o certifikatu>",
          "CertID": "<id zapisa o certifikatu>"}
}
```

Ovo naknadno dodavanje revokacijskih podataka je dodatni razlog za dodavanje posebne promjenjive baze podataka uz nepromjenjivi lanac blokova. Pri razvoju TrustChain B modela[393] postalo je očito da će takva baza podataka biti nužna, ali ona nije bila detaljnije razrađena sve do objavljivanja rada u časopisu Computers[394] koji je bio motiviran potrebom za uključivanjem podataka o arhivskoj vezi, optimiziranjem načina pretrage TrustChain lanca i eventualnim korekcijama metapodataka.

Ovi zapisi (o opozvanim podacima) nisu standardni TrustChain A ili B zapis te se dodaju u skupu na kraj zadnjeg stvorenog bloka, iza skupine zapisa o glasanju. Ova proširena podatkovna struktura TrustChain bloka prikazana je (sa sažecima opisa individualnih zapisa) u programskom kodu 20.

Programski kod 20. TrustChain JSON podatkovna struktura bloka s podacima o revokaciji certifikata.

```
{
"blockHash": "<hash ovog bloka>",
"blockID": "<id bloka, redni broj>",
"block": {
     "previousBlockHash": "<hash prethodnog bloka>",
     "timestamp": "<vrijeme formiranja bloka>",
     "nodeID": "<id čvora koji formira blok>",
     "nodeSig": "<"<potpis podataka>",
     "records": [
          { <zapis1> },
          { <zapis2> },...],
     "votes": [
          { <glas1> },
          { <glas2> },...],
     "revocations":[
          { <opoziv1> },
          { <opoziv2> },...]}
}
```

Iako je, upravo uvedena zbirka `"revocations"` opcionalna, očekujem da će u slučaju učestale upotrebe TrustChain sustava većina blokova starijih od dvije do tri godine sadržavati ovakav dodatak. Nova zbirka sastoji se od niza zapisa o revokacijama kakvi su opisani ranije u programskom kodu 19. Ista podatkovna struktura ali u proširenom, detaljnom, prikazu koji se

[393] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving, n. dj.
[394] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.

sastoji od jednog TrustChain A zapisa, jednog TrustChain B zapisa, jednog zapisa o glasanju o jednom arhivskom zapisu i jednog revokacijskog zapisa prikazana je u programskom kodu 21.

Programski kod 21. TrustChain JSON podatkovna struktura bloka s podacima o revokaciji certifikata – prošireni prikaz

```
{
"blockHash": "<hash ovog bloka>",
"blockID": "<id bloka, redni broj>",
"block": {
	"previousBlockHash": "<hash prethodnog bloka>",
	"timestamp": "<vrijeme formiranja bloka>",
	"nodeID": "<id čvora koji formira blok>",
	"nodeSig": "<potpis podataka>",
	"records": [{"id": "<id zapisa>",
		"version": "<verzija modela>",
		"type": "A",
		"recordHash": "<hash zapisa>",
		"record": {
			"timestamp": "<vrijeme stvaranja zapisa>",
			"certAuthName": "<naziv CA>",
			"certAuthID": "<id CA>",
			"certAuthApiLink": "<poveznica cert. aut.>",
			"data": {
				"TrustChainHash": "<hash dokumenta>",
				"docLnk": "<poveznica na dokument>" },
			"metadata": {
				"docRefCode": "<šifra dokumenta>",
				"docTitle": "<naziv dokumenta>",
				"docCreator": "<naziv ili ime autora>",
				"docCreationDate": "<datum stvaranja>" }}},
		{"id": "<id zapisa>",
		"version": "<verzija modela>",
		"type": "B",
		"recordHash": "<hash zapisa>",
			"record": {
				"timestamp": "<vrijeme stvaranja zapisa>",
				"certData": {
					"CertificateID": "<ID certifikata>",
					"CertificateHash": "<hash certifikata>",
					"ValidityStart": "<vrijeme od kada vrijedi>",
					"ValidityEnd": "<vrijeme do kada vrijedi>",
					"CertificateOwner": "<naziv / ime vlasnika>",
					"CertificateIssuerID": "<CA identifikator>" }
				"RevocationData": {
					"RevocationTime": "<vrijeme revokacije>" ,
					"CRLlink": "<id revokacijske liste>"}}},...],
	"votes": [
		{"nodeID": "<id TrustChain čvora>",
		"nodeSig": "<potpis podataka>",
		"vote": {
			"blockCandidate": "<id bloka o kojem se glasa>",
			"is_block_valid": "<true | false >",
			"timestamp": "<vrijeme glasanja>"}
			"record_votes":[
				"record_id": "<true | false >"],...}},...],
```

```
        "revocations":[
              {"RevocationId": "<id zapisa>",
              "RevocationRecordHash": "<hash zapisa>",
              "RevocationRecord": {
                    "timestamp": "<vrijeme stvaranja zapisa>",
                    "RevocationTime": "<vrijeme revokacije>",
                    "CRLlink": "<id revokacijske liste>",
                    "BlockID": "<id bloka sa zapisom o certifikatu>",
                    "CertID": "<id zapisa o certifikatu>"}},...]
}
```

Programskim kodom 21 je prikazana gotovo potpuna podatkovna struktura TrustChain bloka koji se uvrštava u nepromjenjivi lanac. Dio koji još nedostaje je dopunjavanje neiskorištenog prostora do maksimalne veličine bloka. Jednaka dužina svih blokova omogućuje učinkovitu pretragu podataka u slučaju kada je cijeli lanac pohranjen u jednoj datoteci. Ako su svi blokovi (i zbirke u njima) jednake dužine pristup podacima je moguć algoritmom konstantne vremenske složenosti. Princip je dobro poznat u računarstvu, na primjer u slučaju pristupa podacima pohranjenim u polju. Ako je svaki blok jednake dužine, onda upotrebom početne adrese i rednog broja konkretnog bloka možemo izračunati adresu traženog bloka te mu izravno pristupiti. Konkretno, adresa (memorijska lokacija) traženog bloka, A_x jednaka je:

$$A_x = A_0 + (B * n)$$

ako je A_0 adresa na kojoj se nalazi prvi blok, B veličina bloka (konstanta) i n redni broj bloka koji se traži. Isti princip izravnog pristupa upotrebom izračuna primjenjiv je i na razini individualnog bloka u TrustChain lancu te se na ovaj način, upotrebom univerzalnih programerskih tehnika, može osigurati učinkovit pristup podacima u lancu blokova čak i u okolnostima u kojima je on sadržan u jednoj velikoj datoteci ili više sekvencijalno pohranjenih datoteka.

Dopunjavanje neiskorištenog prostora (engl. *padding*) kao i rasprava o maksimalnoj veličini bloka ostavljena je za konkretne implementacije modela. Veličina bloka izravno je ovisna o broju zapisa koji se stvaraju u jedinici vremena te je kao takva ovisna o konkretnom slučaju upotrebe sustava. Pojam dopunjavanja je ranije objašnjen i ne zahtijeva posebno objašnjenje – riječ je o nizu nuli ili jedinica.

6.3. Pomoćna baza podataka

Potreba za dodatnim sustavom pohrane podataka, koji će funkcionirati kao pomoć TrustChain lancu blokova, uočena još od tijekom razvoja prvog TrustChain modela[395] ali se nije nametnula kao kritična komponenta do trenutka kada je počelo razmatanje TrustChain modela fokusiranog na pohranu digitalnih certifikata.[396] Prema mojim saznanjima svi sustavi temeljeni na ulančanim blokovima (engl. *blockchain*) koriste neku vrstu pomoćnog informacijskog sustava za pretragu podataka u ulančanom bloku ili pretraživača ulančanih blokova (engl. *blockchain explorer*). Svrha ovakvog sustava je osigurati učinkovitu pretragu lanca ulančanih blokova s obzirom na to da su ulančani blokovi:

a) linearna podatkovna struktura,

b) organiziraju podatke kronološki prema redu zapisa, bez posebne organizacije te

c) ne sadrže indeks podatka.

Osim u slučaju kada se podatak traži prema vremenu zapisa, lanac ulančanih blokova pretražuje se linearno, to jest sekvencijalno. Ovakva pretraga podrazumijeva algoritam linearne vremenske složenosti i najmanje je učinkovit način pretrage podataka.[397] S druge strane, iako bi u slučaju pretrage prema vremenu zapisa mogli, barem djelomično, upotrijebiti binarnu pretragu, O(log n) algoritam, to je izuzetno rijedak slučaj u bilo kojem sustavu za pohranu podataka. Podaci se obično pretražuju prema autoru, naslovu, vremenu nastajanja dokumenta (ne zapisa u bazu podataka) i drugima. Iz ovih razloga svi sustavi koji koriste ulančane blokove koriste i neki pomoći sustav za njihovo pretraživanje.

Primjere ovakvih sustava u slučaju kriptovaluta (engl. *cryptocurrency*) baziranih na ulančanim blokovima možemo naći u izvještaju Yang i Li o njihovom pretraživaču ulančanih blokova[398] i u pregledu "SilkViser" sustava, vizualnog pretraživača ulančanih Bitcoin zapisa autora Zhong i ostalih.[399]

---

[395] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term, n. dj.
[396] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving, n. dj.
[397] Rahim, R., Nurarif, S., Ramadhan, M., Aisyah, S., & Purba, W. (2017). Comparison searching process of linear, binary and interpolation algorithm. *Journal of Physics: Conference Series, 930*(1), 012007. doi: https://doi.org/10.1088/1742-6596/930/1/012007
[398] Yang, R., & LI, Y. (2020). *The Final Report for A Blockchain Explorer*. Hong Kong: Hong Kong University of Science and Technology. Preuzeto 22. 11. 2021. s https://file-1252789527.cos.ap-shenzhen-fsi.myqcloud.com/MSBD6000D-Blockchain/MSBD6000D%20Group%2016%20report.pdf
[399] Zhong, Z., Wei, S., Xu, Y., Zhao, Y., Zhou, F., Luo, F., & Shi, R. (2020). SilkViser: A Visual Explorer of Blockchain-based Cryptocurrency Transaction Data. *2020 IEEE Conference on Visual Analytics Science and Technology (VAST)* (str. 95-106). IEEE. Preuzeto 22. 11. 2021. s https://arxiv.org/pdf/2009.02651.pdf

S obzirom na sve navedeno od samog početka razvoja modela bilo je jasno da će slični pomoćni sustav biti potrebno izraditi i za informacijski sustav baziran na TrustChain modelu, ali se kasnije pokazalo da je u slučaju TrustChain modela potreban značajno kompleksniji sustav nego što su to standardni pretraživači koji je i sam baza podataka za neke elemente sustava.

Kao što je ranije pokazano, TrustChain B modul, posebno u svojim interakcijama s TrustChain A modulom, pretpostavlja pohranu velikog broja digitalnih certifikata, pa njihova pohrana (zbog velikog broja) u sam lanac blokova nije racionalna. Ovo je bio prvi poticaj da se razvoju pomoćnog sustava za pretragu TrustChain lanca blokova pristupi kao punoj bazi podatka koja podržava podatke u lancu blokova.

Osim navedenog, prethodna istraživanja su pokazala da TrustChain sustav mora omogućiti naknadnu izmjenu metapodataka. Svi sustavi za pohranu podataka pate od ljudske greške pri unosu podatka. Zbog ovoga dobro je omogućiti naknadno ispravljanje uočenih grešaka, što je i zahtjev Uredbe GDPR. Zbog specifičnih zahtjeva za sustav u slučaju TrustChain sustava postoji povećana potreba za ovakvim izmjenama. Ovo je rezultat potrebe da sustav koji je specijaliziran za dugotrajnu pohranu odgovara zahtjevima arhivistike i diplomatike. Jedan od tih zahtjeva je i evidencija podataka o arhivskoj vezi[400]. Kao što je u ranijem poglavlju objašnjeno arhivska veza nije potpuno definirana (i podložna je promjena) dok dokument ne postane neaktivan (ali je njezino očuvanje i dalje potrebno)[401]. Bez klasične baze podataka naknadne izmjene ne bi bile moguće (jer bi svi podaci bili sadržani u nepromjenjivoj podatkovnoj strukturi).

Može se zaključiti da TrustChain sustav za normalno funkcioniranje zahtjeva razvoj pomoćnog sustava koji omogućava izmjenu podatka. U slučaju TrustChain sustava ovaj pomoćni sustav još je kompleksniji od onih koje koristi većina sustava temeljenih na ulančanim zapisima zbog:

a) potrebe za učinkovitom pretragom podatkovne strukture ulančanih blokova,

b) potrebe za pohranom velikog broj digitalnih certifikata,

c) povećane potrebe za naknadnom izmjenom metapodataka izazvane održavanjem podataka o arhivskoj vezi.

[400] Zhong, Z., Wei, S., Xu, Y., Zhao, Y., Zhou, F., Luo, F., & Shi, R. (2020). SilkViser: A Visual Explorer, n. dj.
[401] Duranti, L., & Macneil, H. (1996). The Protection of the Integrity of Electronic Records, n. dj.

Ovo poglavlje razrađuje razvoj takvog sustava. Iako ideja o njemu postoji od samog početka razvoja modela, TrustChain pomoćni sustav izvorno je predstavljen 2021. u časopisu Computers.[402] Prije razvoja samog sustava potrebno je odabrati tehnologiju koja će biti korištena. Tijekom istraživanja iz 2021. provedeno je pregled postojećih sustava za pohranu podataka temeljenih na ulančanim blokovima da bi se utvrdile najčešće korištene tehnologije te se iskoristile u razvoju pomoćne baze podataka TrustChain sustava. S obzirom da je ostatak ovog poglavlja uvjetovan rezultatima ovog ranije provedenog istraživanja u nastavku su parafrazirani i razrađeni najvažniji zaključci proizašli iz tog istraživanja.

Pregled postojećih sustava (Computers[403]) uzeo je u obzir tri arhivska sustava temeljena na ulančanim blokovima:

1) ARCHAIN sustav.[404] Razmatrani rad je upozorio na neke probleme vezane uz upotrebu ulančanih blokova kao podrške arhivskim sustavima. Nažalost, rad nije uopće spomenuo tehnologiju koju koristi (digitalni) arhiv i na koju se ARCHAIN nadograđuje. Svejedno, rad je pružio u uvid u problematiku i dao potvrdu ideji da se ulančani zapisi mogu koristiti na ovaj način kao garancija integriteta digitalnih zapisa pohranjenih u drugom sustavu (slično TrustChain A modulu).

2) Cilegon E-Archive sustav[405] koristi IPFS[406] (engl. *Interplanetary File System*) za pohranu podataka. U IPFS sustavima podaci (datoteke) se, često fragmentirani, pohranjuju na više računala, to jest poslužitelja koji sudjeluju u sustavu.[407] Ovaj interesantni koncept omogućava potpuno distribuiran digitalni arhivski sustav i opciju proširenja TrustChain A modela. Sličnu ideju već su razradili Naz i drugi autori 2019.[408] Prema njima upotreba IPFS sustava i

[402] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.
[403] Ibid.
[404] Galiev, A., Prokopyev, N., Ishmukhametov, S., Stolov, E., Latypov, R., & Vlasov, I. (2018). Archain: a novel blockchain based archival system. *2018 Second World Conference on Smart Trends in Systems, Security and Sustainability* (str. 84-89). IEEE. Preuzeto 7. 1. 2022. s https://arxiv.org/ftp/arxiv/papers/1901/1901.04225.pdf
[405] Permatasari, I., Essaid, M., Kim, H., & Ju, H. (2020). Blockchain Implementation to Verify Archives Integrity on Cilegon E-Archive. *Applied Sciences, 10*(7). Preuzeto 7. 1. 2022. s https://www.mdpi.com/2076-3417/10/7/2621
[406] IPFS URL: https://ipfs.io/
[407] Benet, J. (2014). *IPFS*, n. dj.
[408] Naz, M., Al-zahrani, F. A., Khalid, R., Javaid, N., Qamar, A. M., Afzal, M. K., & Shafiq, M. (2019). A secure data sharing platform using blockchain and interplanetary file system. *Sustainability, 11*(24), 7054. Preuzeto 7. 1. 2022. s https://www.mdpi.com/2071-1050/11/24/7054

ulančanih blokova može biti osnova za izuzetno pouzdan digitalni arhiv. Nisam prihvatio ovakvo rješenje u slučaju TrustChain sustava zbog sljedećih razloga:

a) S obzirom na to da IPFS radi na razini datoteke upotreba bi, zbog lakše pretrage podatka opet zahtijevala upotrebu dodatne baze podataka. Sustav bi i dalje zahtijevao posebnu bazu podatka.

b) IPFS dodatno naglašava probleme s povjerljivošću dokumenata koji proizlaze iz upotrebe TrustChain A modula. Dokumenti se sada više ne šalju samo na udaljenu provjeru već će i trajno biti pohranjeni na nepoznatim računalima.

c) TrustChain je od samog početka zamišljen kao kolaboracija više arhivskih ustanova, što je nužno za normalno funkcioniranje TrustChain sustava za glasanje. Kao takav TrustChain već je distribuirani sustav te ima manje koristi od sigurnosnog elementa upotrebe specijaliziranog sustava za distribuiranu pohranu podataka.

d) Jedna od važnih prednosti upotrebe IPFS sustava je mogućnost upotrebe postojeće (i rastuće) infrastrukture. S obzirom na razlog naveden u točki c, suradnju više arhivskih institucija, dostatna infrastruktura će, vjerojatno, već postojati ili će se moći uspostaviti. Zbog toga je ova prednost upotrebe IPFS sustava značajno umanjena.

3) Lekana sustav[409]. Kao i raniji sustav koji su razvili Bandara i drugi, ovaj sustav koristi distribuiranu bazu podatka – Apache Cassandra.[410]

[409] Bandara, E., Liang, X., Shetty, S., Ng, W. K., Foytik, P., Ranasinghe, N., . . . Larsson, D. (2020). Lekana-Blockchain Based Archive Storage for Large-Scale Cloud Systems. *International Conference on Blockchain* (str. 169-184). Cham: Springer. Preuzeto 7. 1. 2022. s https://www.researchgate.net/publication/344372675_Lekana_-_Blockchain_Based_Archive_Storage_for_Large-Scale_Cloud_Systems

[410] Apache Cassandra URL: https://cassandra.apache.org/_/index.html

Uz ove dedicirane arhivske sustave razmotrena su još četiri općenita (ne arhivska) sustava za pohranu podataka oni su:

1) BigChainDB.[411] Ovaj sustav koristi ulančane zapise da bi garantirao integritet podataka pohranjenih u MongoDB[412] distribuiranu bazu podataka.[413] U svojim ranijim verzijama BigChainDB je koristio klasične SQL baze podataka, poput Microsoft i Oracle SQL poslužitelja ali je u kasnijim iteracijama napustio ove sustave.

2) ChainSQL [414] sustav je jedinstveno rješenje jer je univerzalan. Ovaj sustav omogućuje potvrdu integriteta podatka pohranjenih u bilo koji SQL ili noSQL sustav za pohranu podataka.

3) EthernityDB[415] sustav je distribuirana baza podatka koja se oslanja na Etherum lanac blokova. Ethereum je distribuirani sustav ulančanih blokova koji upotrebom vlastitog programskog jezika omogućava pisanje pametnih ugovora.[416] Upotreba Ethereuma kao osnove za arhivski sustav razmotrena je u ranim fazama razvoja TrustChain A modela ali se od nje odustalo jer Etherum naplaćuje sve transakcije provedene kroz svoj mrežu te je procijenjeno da bi njegova upotreba kao sustava za pohranu podatka (čak i samo hasheva) bila preskupa.

4) Mystiko sustav.[417] Sustav je ranija verzija kasnijeg arhivskog sustava – Lekana. Iako, za razliku od Lekane, nije dedicirani arhivski sustav i Mystiko se oslanja na Apache Cassandra sustav.

---

[411] McConaghy et al. (2016): *Bigchaindb*, n. dj.
[412] MongoDB URL: https://www.mongodb.com/
[413] MongoDB Inc. (2020). *MongoDB Documentation*. Preuzeto 19. 11. 2022. iz MongoDB: https://docs.mongodb.com/
[414] Muzammal, M., Qu, Q., & Nasrulin, B. (2019). Renovating blockchain with distributed databases: An open source system . *Future generation computer systems*, 105-117. doi: https://doi.org/10.1016/j.future.2018.07.042
[415] Helmer, S., Roggia, M., El Ioini, N., & Pahl, C. (2018). Ethernitydb–integrating database functionality into a blockchain. *European Conference on Advances in Databases and Information Systems* (str. 37-44). Cham: Springer. Preuzeto 7. 1. 2022. s https://www.researchgate.net/publication/327309357_EthernityDB_-_Integrating_Database_Functionality_into_a_Blockchain
[416] Wood, G. (2014). *Ethereum: A secure decentralised generalised transaction ledger*. Preuzeto 23. 11. 2021. s Ethereum project yellow paper: https://files.gitter.im/ethereum/yellowpaper/VIyt/Paper.pdf
[417] Bandara, E., Ng, W. K., Zoysa, D., K., F., N., T., S., M. P., & Jayasuriya, N. (2018). Mystiko—blockchain meets big data. *2018 IEEE International Conference on Big Data*. IEEE. Preuzeto 7. 1. 2022. s https://www.researchgate.net/publication/330632586_Mystiko-Blockchain_Meets_Big_Data

Rezime provedene rasprave prikazan je u tablici 11. Tablica prikazuje najvažnije osobine razmotrenih sustava: korištenu tehnologiju, tip sustava (arhivski ili ne) i pristup pohrani podataka (centraliziran sustav ili ne).

Tablica 11. Pregled tehnologija postojećih sustava za pohranu podataka baziranih na ulančanim blokovima. Izvor: Stančić, Bralić. Digital Archives Relying on Blockchain: Overcoming the Limitations of Data Immutability. 2021.

Naziv sustava	Arhivski sustav	Centraliziran sustav	Korištena tehnologija
ARCHAIN	Da	Da	Postojeći državni arhivski sustav
Cilegon E-Archive	Da	Ne	Interplanetary file system (IPFS)
Lekana	Da	Ne	Apache Cassandra (noSQL)
BigchainDB	Ne	Ne	RethinkDB i MongoDB (noSQL)
ChainSQL	Ne	Ne	Bilo koja baza podataka (SQL or noSQL)
EthernityDB	Ne	Ne	Ethereum sustav ulančanih blokova
Mystiko	Ne	Ne	Apache Cassandra (noSQL)

Iz tablice 11 je vidljivo da su distribuirani noSQL sustavi dominantni u konkurenciji razmotrenih sustava te se u ovom poglavlju predlažu konkretne podatkovni modeli temeljeni na Apache Cassandra i MongoDB sustavima. Na osnovu gornjeg pregleda, moj zaključak je da bi pri implementaciji TrustChain modela trebalo koristiti jedan od ova dva sustava. Svejedno, za razliku od rada koji je izvorno predstavio podatkovne modele pomoćne baze, ovdje će biti razrađen i relacijski model. Razlozi za dodavanje relacijskog (SQL) modela su:

a) Relacijski modeli su uobičajen način prikaza podataka. Predstavljanje modela će početi s njim da olakša čitatelju razumijevanje modela i da omogući lakšu usporedbu kasnijih noSQL modela, baziranih na dokumentima, sa standardnim relacijskim modelom.

b) Iako preporučujem upotrebu Apache Cassandra ili MongoDB sustava, zbog postojeće infrastrukture u ustanovama koje sudjeluju u TrustChain sustavu moguće je da će se kasnije, iz ekonomskih razloga, SQL baze podataka pokazati kao isplativije rješenje.

Iduća poglavlja razrađuju navedene podatkovne modele prema arhivističkim zahtjevima i pratećem logičkom modelu, te prema (novim) potrebama detaljnih procesa iz prethodnih poglavlja.

6.3.1. SQL varijanta pomoćnog sustava

U ovom poglavlju razrađen je pojednostavljeni relacijski model TrustChain pomoćne baze podataka. Ne preporučujem upotrebu relacijske SQL baze podataka za implementaciju pomoćne baze podataka. Istraživanje postojećih sustava prezentirano u Computers časopisu[418] jasno je pokazalo da su za ovakav sustav pogodnije distribuirane noSQL baze podataka.[419] Kasnija poglavlja daju dva primjera upotrebe ovakve tehnologije za implementaciju TrustChain pomoćne baze podatka upotrebom distribuiranih noSQL tehnologija – MongoDB i Apache Cassandra.

Svrha relacijskog SQL modela predstavljenog u ovom poglavlju je jasnija prezentacija podataka potrebnih za funkcioniranje pomoćne baze podataka. Relacijski modeli najsličniji su logičkim modelima podatka te je na njihovo čitanje naviknut veći broj stručnjaka (nego na čitanje relativno novih noSQL modela). Osim ovoga, noSQL baze podataka nemaju standardiziran način grafičkog prikaza odnosa između njihovih entiteta. One ni same nisu standardizirane. Ovdje predstavljeni MongoDB model se bazira na noSQL bazi podataka temeljenoj na dokumentima dok je Apache Cassandra model temeljen na upitima. Svrha izrade relacijskog modela je pokušaj standardizacije prikazanih modela.

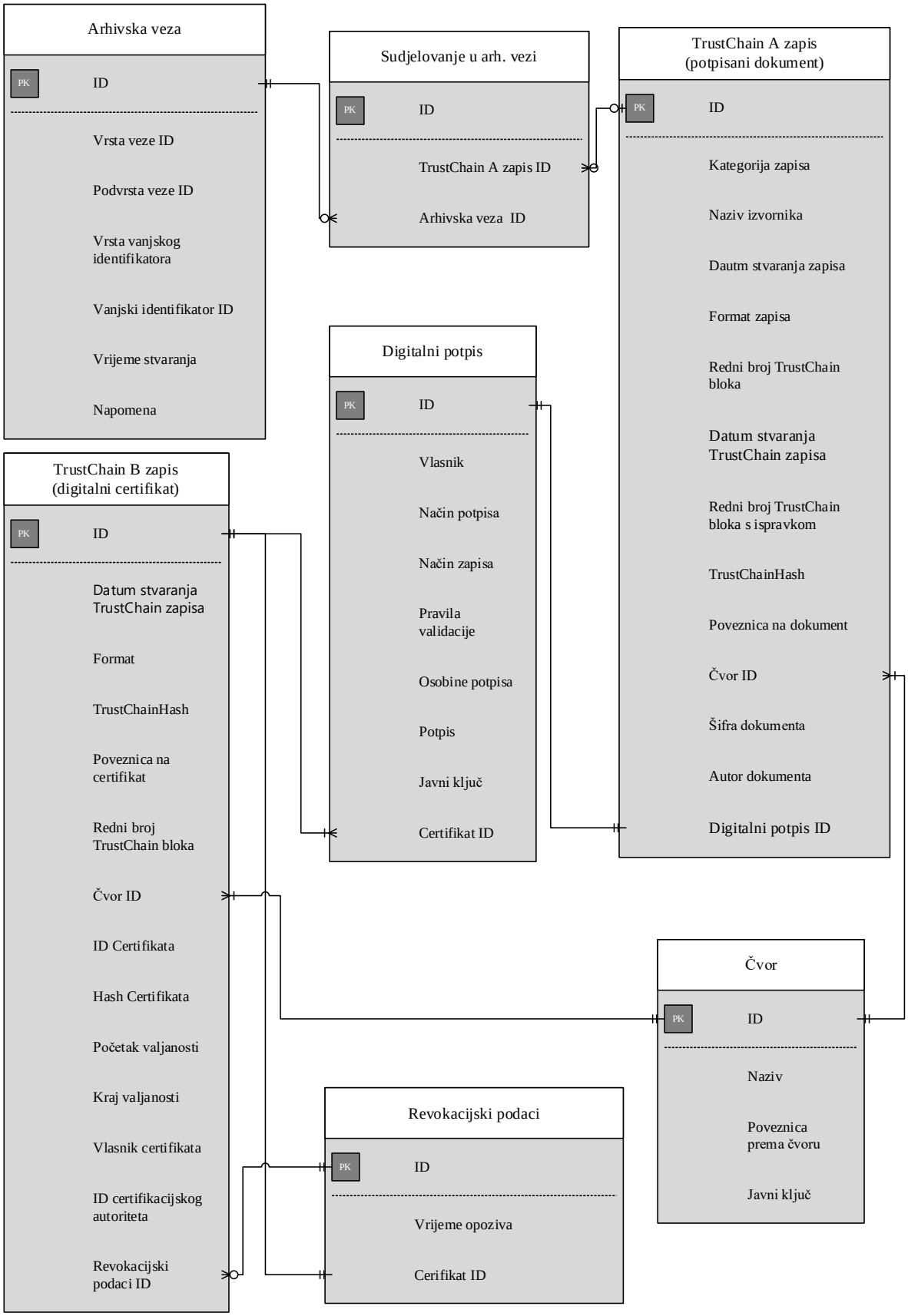
Sukladno cilju pojašnjenja podatkovnog modela, relacijski model prikazan na slici 37 koristi notaciju vranine noge (engl. *crow's foot*), vjerojatno najkorišteniji i jedan od najstarijih načina prikaza relacijskog podatkovnog modela.[420] Model je baziran na logičkom modelu razrađenom u drugom poglavlju te je korišten kao osnova za kasnije prikazane MongoDB i Apache Cassandra noSQL podatkovne modele.

---

[418] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.

[419] Han, J., Haihong, E., Le, G., & Du, J. (2011). Survey on NoSQL database. *2011 6th international conference on pervasive computing and applications* (str. 363-366). IEEE. Preuzeto 26. 11. 2021. s http://faculty.washington.edu/wlloyd/courses/tcss562/papers/Spring2017/team7_NOSQL_DB/Survey%20on%20NoSQL%20Database.pdf

[420] Everest, G. C. (1976). BASIC DATA STRUCTURE MODELS EXPLAINED WITH A COMMON EXAMPLE. Computing Systems 1976, Proceedings Fifth Texas Conference on Computing Systems (str. 39-46). Austin: IEEE Computer Society Publications Office. Preuzeto 26. 11. 2021. s https://www.researchgate.net/profile/Gordon-Everest-2/publication/291448084_BASIC_DATA_STRUCTURE_MODELS_EXPLAINED_WITH_A_COMMON_EXAMPLE/links/57affb4b08ae95f9d8f1ddc4/BASIC-DATA-STRUCTURE-MODELS-EXPLAINED-WITH-A-COMMON-EXAMPLE.pdf



Slika 37. SQL model pomoćne baze podataka

Relacijski model na slici 37 je pojednostavljen, nedostaje više relacija, čak i neke koje sam prikazani model referencira. Na primjer, nedostaju kataloške relacije "Vrsta arhivske veze" i "Podvrsta arhivske veze". Također nedostaje relacija s popisom certifikacijskih autoriteta i relacija koji bilježi čvor koji je na redu za stvaranje novog bloka. Osim ovakvih relacija, izravno povezanih s TrustChain podacima nedostaje i veliki broj pomoćnih relacija koje su potrebne za normalno funkcioniranje informacijskog sustava. Na primjer, nedostaju relacije s popisom i ovlastima korisnika TrustChain sustava. Sve ove relacije (i druge) su preskočene jer na razini modela informacijskog sustava nisu kritične (ova disertacija nema za cilj sastaviti detaljnu projektu dokumentaciju) te bi zakomplicirale prikaz i otežale razumijevanje podatkovnih struktura koje jesu kritične za TrustChain model.

6.3.2. MongoDB varijanta pomoćnog sustava

MongoDB je sustav za pohranu podataka temeljen na "dokumentima".[421] MongoDB dokument označava osnovni organizacijski skup podataka u sustavu, slično terminu relacija u relacijskim bazama koji je često kolokvijalno zvan i "tablica". Termin "dokument" je poprilično problematičan za ovu disertaciju jer se termin već (često) koristi kao osnova jedinca pohrane u TrustChain A modulu. Zbog ovog razloga u ovom poglavlju će se posebno naglašavati kada se termin dokument koristi za digitalno potpisanu datoteku (što je slučaj u ostatku disertacije). U ostalim slučajevima termin se koristi za oznaku osnovne organizacijske jedinice MongoDB sustava.

Zapisi u MongoDB sustavu prate JSON standard te su izuzetno pogodni za upotrebu u modeliranju TrustChain sustava jer se isti standard koristio i u prethodnim poglavljima pri modeliranju zapisa sadržanih u TrustChain lancu blokova. Na ovaj način omogućuje se dosljednost pri pisanju i čitanju podatkovnih struktura. JSON zapis se pri uključivanju u MongoDB sustav prevodi u binarni zapis te je karakterističan za MongoDB sustav – BSON (engl. *Binary JavaScript Object Notation*).[422] BSON zapisi se danas koriste i izvan MongoDB sustava te su njegove implementacije dostupne u većini programskih jezika.[423]

Osim toga, u usporedbi s konkurentom tehnologijom, Apache Cassandra distribuiranim sustavom za pohranu podatka, MongoDB je pokazao nešto slabije performanse u istraživanju

[421] MongoDB Inc. (2020). MongoDB Documentation, n. dj.
[422] MongoDB Inc. (2020). JSON and BSON. Preuzeto 19. 11. 2022. s MongoDB: https://www.mongodb.com/json-and-bson
[423] *BSON Implementations*. (2021). Preuzeto 24. 11. 2021. s BSON: https://bsonspec.org/implementations.html

koje su proveli Abramova i Bernardino 2013.[424] Do sličnih zaključaka 2016. došli su Ramesh i suradnici[425] iako se njihovi zaključci odnose prvenstveno na upite koji čitaju podatke te predlažu upotrebu tehnologije pri analizi velike količine podatka. Iste godine Haughian i suradnici[426] su došli do zaključka da MongoDB sustav na učinkovitiji, to jest brži način provodi replikaciju podatka među čvorovima koji sudjeluju u sustavu. S obzirom na to je TrustChain arhivski sustav, koji barata s velikim brojem dokumenata koji se više aktivno ne koriste, smatram da broj upita koji se postavlja TrustChain pomoćnoj bazi podataka neće biti velik (u kratkim vremenskim periodima), što donekle umanjuje prepoznate prednosti Cassandra Apache sustava. Čak i ako jest, povećanje učinkovitosti prilikom korištenja noSQL baza podataka (Cassanda, MongoDB) u odnosu na klasične relacijske SQL baze podataka (Microsoft, Oracle) je značajno prema svim navedenim istraživanjima.

Zbog navedenih razloga (dosljednost modela i rasprostranjenost tehnologije) MongoDB je logičan izbor te će u konačnici biti i preporučen kao temeljna tehnologija za pomoćnu bazu podataka u budućim implementacijama TrustChain modela. Usprkos tome bolje performanse sustava moguće je postići upotrebom Cassandra Apache modela pa će biti predstavljen i model baziran na toj tehnologiji.

Podatkovne strukture prikazane u radu iz 2021.[427] pokazale su se nedostatne za model koji je razvijen u ovoj disertaciji. I sam izvorni rad priznaje da će u končanom podatkovnom modelu biti sadržani neki elementi koje on ne obuhvaća, ali se pokazalo da ni to nije dovoljno već je potrebna izmjena postojećih dokumenta (MongoDB podatkovnih struktura) i razvoj novih koje odgovaraju elementima uvedenim u ranijim poglavljima. Prije svega treba razviti novi MongoDB dokument koji odgovara zapisu iz TrustChain B modula. Modeli koji se navode u nastavku ove disertacije se temelje na onima iz izvornog rada, ali su značajno razrađeni te su dodane i nove strukture koje odgovaraju novim potrebama sustava.

Dokument prikazan u programskom kodu 22 najsličniji je izvornom radu te je u ovoj disertaciji modificiran na način da ostvari cilj indeksiranja i pohrane metapodataka zapisa

---

[424] Abramova, V., & Bernardino, J. (2013). NoSQL databases: MongoDB vs cassandra. *Proceedings of the international C* conference on computer science and software engineering*, (str. 14-22). Preuzeto 7. 1. 2022. s http://web.cs.wpi.edu/~cs585/s17/StudentsPresentations/This%20Year/Week14/mongodb_vs_cassandra.pdf

[425] Ramesh, D., Sinha, A., & Singh, S. (2016). Data modelling for discrete time series data using Cassandra and MongoDB. *2016 3rd international conference on recent advances in information technology (RAIT)* (str. 598-601). Dhanbad, India : IEEE. doi: https://doi.org/10.1109/RAIT.2016.7507966

[426] Haughian, G., Osman, R., & Knottenbelt, W. (2016). Benchmarking replication in cassandra and mongodb nosql datastores. *International Conference on Database and Expert Systems Applications* (str. 152-166). Cham: Springer. doi: https://doi.org/0.1007/978-3-319-44406-2_12

[427] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.

stvorenih u TrustChain A modulu. Dokument sadrži sve metapodatke iz izvornog rada, koji su utvrđeni na osnovu istraživanja o korištenim metapodacima u arhivskim standardima razrađenom u drugom poglavlju.

Programski kod 22. MongoDB dokument s podacima TrustChain A zapisa

```
{
          "_id": "<identifikator zapisa>",
          "Category": "<kategorija zapisa>",
          "date": "<vrijeme stvaranja zapisa>",
          "Format": "<format zapisa>",
          "TrustChainHash": "<hash dok. (koji je dodan u TC)>",
          "docLnk": "<poveznica na dokument (koji je dodan u TC)>",
          "blockID": "<identifikator bloka>",
          "blockCorrID": [
               "blockID": "<identifikator bloka>",
               ...]
          "nodeID": "<identifikator čvora koji je stvorio blok>",
          "digitalSignature":[{
               "sigEncode": "<način zapisa>",
               "signer": "<vlasnik potpisa>",
               "sigMethod": "<način potpisa>",
               "sigValue": "<vrijednost potpisa>",
               "sigValid": "<pravila provjere ispravnosti>",
               "sigProp": "<osobine potpisa>",
               "key": "<javni ključ>",
               "certAuthName": "<naziv cert. autoriteta>",
               "certAuthID": "<id certifikacijskog autoriteta>",
               "certAuthApiLink": "<poveznica cert. aut.>"},...]
          "metadata": {
               "docRefCode": "<šifra dokumenta>",
               "docTitle": "<naziv dokumenta>",
               "docCreator": "<naziv ili ime autora>",
               "docCreationDate":"<datum stvaranja>"
                }
          "arcBond": [{
               "type": "<vrsta_arhivske_veze>",
               "subtype": "<podvrsta_arhivske_veze>",
               "id": {
                    "idBondType": "<tip poveznice pov. obj.>",
                    "idBondValue": "<id_povezanog_objekta>",   }
               "bondSeqNumber": "<redni_broj_povezanog_objekta>"
               },...]
}
```

Gore prikazani MongoDB dokument u potpunosti zadovoljava potrebe zapisa proizašlog iz TrustChain A modula. U dokument su dodane zbirke `"digitalSignature"` i `"arcBond"` koje predstavljaju kolekcije ugniježđenih dokumenata te omogućavaju:

a) da sustav sačuva informaciju o višestrukim digitalnim potpisima što je sasvim sigurno slučaju kod dokumenata koji su prije ulaska u TrustChain sustav bili dio arhiva koji je dugotrajno očuvanje autentičnosti garantirao na ovaj način,

b) da sustav sačuva informaciju o višestrukom sudjelovanju u arhivskim vezama.

Osim ovih zbirki dodana je i posebna zbirka s referencama na zapise s ispravkom navoda u kasnijim blokovima. Proces stvaranja bloka s ispravkom svodi se na kopiranje postojećih podataka, ispravak pogrešnih navoda i stvaranje novog zapisa.

S obzirom na to da MongoDB sustav, za razliku od suvremenih relacijskih bazi podataka ne garantira referencijalni integritet te se pretraga koja uspoređuje više različitih upita, što je ekvivalent SQL *join* (engl.) upita,[428] mora riješiti na logičkoj razini sustava na ovaj način (ugnježđivanjem dokumenta) omogućuje se stvaranje i učinkovita pretraga povezanih podatkovnih struktura.

Osim podataka o arhivskoj vezi sadržanih u svakom zapisu o dokumentu, MongoDB pomoćni sustav može sadržavati još jedan dokument koji je fokusiran na samu arhivsku vezu te sadrži podatke o njoj i zbirku identifikatora MongoDB dokumenata koji u njoj sudjeluju. Taj dokument prikazan je u programskom kodu 23.

Programski kod 23. MongoDB dokument s podacima o arhivskoj vezi

```
{
          _id: "<identifikator arhivske veze>",
          type: "<kategorija arhivske veze>",
          name: "<naziv arhivske veze>",
          date: "<vrijeme stvaranja arhivske veze>",
          note: "<napomena>",
          records:[
               recordId: "<identifikator zapisa>",
               ...
               ]
}
```

[428] MS SQL *join* (engl.)dokumentacija: https://docs.microsoft.com/en-us/sql/relational-databases/performance/joins?view=sql-server-ver15

Iako dodavanje ovakvog dokumenta rezultira djelomičnim repliciranjem podataka u pomoćnoj bazi smatram da je ono opravdano jer se omogućava novi način pretrage pomoćne baze podataka i samog TrustChain lanca blokova. Dodavanjem ovakvog dokumenta, kao što je predloženo u izvornom prijedlogu pomoćne baze podatka[429] omogućuje se nova početna točka pretrage pomoćne baze podatka. Bez nje svaka pretraga mora početi od nekog zapisa (dokumenta ili certifikata) te je u slučaju zapisa o dokumentu (TrustChain A zapis) moguće nastaviti pretragu prema drugim zapisima prateći arhivske veze u kojima dokument sudjeluje. Ovaj dodatak omogućava da arhivska veza, u slučaju kada je ona u početku poznata, bude početna točka pretege pomoćne baze podatka. Smatram da je replikacija podatka (koja rezultira kompliciranijim i dužim postupkom zapisa podataka) opravdana dodavanjem ove nove funkcionalnosti.

U odnosu na izvorni model pomoćne baze podataka[430] podatkovna struktura prikazana u programskom kodu 24, je potpuno nova. U ovom MongoDB dokumentu sadržani su podaci o zapisima proizašlim iz TrustChain B modula, dakle zapisima o digitalnim certifikatima koji su pohranjeni u sustavu.

Programski kod 23. MongoDB dokument s podacima TrustChain B zapisa

```
{
            "_id": "<identifikator_zapisa>",
            "Category": "<kategorija_zapisa>",
            "date": "<vrijeme_stvaranja_zapisa>",
            "Format": "<format_zapisa>",
            "TrustChainHash": "<hash TurstVhain zapisa>",
            "certLnk": "<poveznica na certifikat>"
            "blockID": "<identifikator_bloka>",
            "nodeID": "<identifikator cvora koji je stvorio blok>",
            "recordData":
                  {
                  "timestamp": "<vrijeme stvaranja zapisa>",
                  "certData": {
                        "CertificateID": "<ID certifikata>",
                        "CertificateHash": "<hash certifikata>",
                        "ValidityStart": "<vrijeme od kada vrijedi>",
                        "ValidityEnd": "<vrijeme do kada vrijedi>",
                        "CertificateOwner": "<naziv vlasnika>",
                        "CertificateIssuerID":"<CA identifikator>" }
            "RevocationData": {
                        "RevocationTime": "<vrijeme revokacije>",
                        "CRLlink": "<id revokacijske liste>" }
```

[429] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.
[430] Ibid.

```
            "recordsUsingCert":[
                  "recordId": "<id zapisa koji koristi certifikat>",
                  ...
                  ]
}
```

Novi MongoDB dokument uključuje sve podatke koji su sadržani u samom zapisu o certifikatu u TrustChain lancu blokova te dodaje zbirku `"recordsUsingCert"`. Ova zbirka sadrži identifikatore TrustChain A zapisa koji koriste certifikat na koji se ovaj MongoDB dokument odnosi. Dakle, sadrži poveznice prema zapisima o dokumentima (PDF ili drugim datotekama) koje su potpisane ovim digitalnim certifikatom. Na ovaj način povećana je učinkovitost pretrage pomoćne baze podataka kada je cilj pretrage dokument koji je potpisan upotrebom certifikata koji je početna točka pretrage.

Osim gornjeg zapisa pomoćna baza podataka sadrži i sam certifikat. Podatkovna struktura koja sadrži certifikat prikazana je u programskom kodu 25. Ovaj MongoDB dokument nije zamišljen kao početna točka pretrage baze podatka pa stoga, za razliku od ranije prikazanih dokumenata ne sadrži nikakve ugniježđene zbirke.

Programski kod 25. MongoDB dokument s digitalnim certifikatom

```
{
            "_id": "<identifikator_certifikata>",
            "CertVersion": "<kategorija_zapisa>",
            "CertSerialN": "<vrijeme_stvaranja_zapisa>",
            "SigningAlgorithmID": "<format_zapisa>",
            "CertificateIssuer": "<CA naziv>",
            "ValidityStart": "<vrijeme od kada vrijedi>",
            "ValidityEnd":"<vrijeme do kada vrijedi>",
            "CertificateOwner": "<naziv vlasnika>",
            "PublicKeyAlgorithm": "<naziv algoritma javnog ključa>",
            "PublicKey": "<sadržaj javnog ključa>",
            "CertificateIssuerID": "<CA identifikator>",
            "CertificateOwnerID": "<identifikator vlasnika>",
            "Extensions": {
                  "CDP": "<lokacija distribucijske točke CRL listi>",
                  "AIA": "<lokacija pristupne točke CA informacija>",
                  "Other": "<drugi (opcionalni) podaci>"}
            "CAPublicKeyAlgorithm": "<algoritam javnog ključa CA>",
            "CAPublicKey": "<sadržaj javnog ključa CA>"
}
```

MongoDB zapis prikazan programskim kodom 25 sličan je zapisu o digitalnom certifikatu (MongoDB dokument o TrustChain B zapisu) ali sadrži samo digitalni certifikat, to

je JSON zapis standardnog x.509 certifikata (opisan u trećem poglavlju). Ova distinkcija je važna jer je ovo set podatka na koji se odnosi certifikacijski hash pohranjen u TrustChain lanac blokova te kao takav (iako za to ne postoji tehnička zapreka) nije predviđen za izmjene. Ranije prikazani zapis se može mijenjati, na primjer u slučaju dodavanja novih zapisa koji koriste konkretni certifikat. U slučaju izmjene ovog dokumenta njegov hash više neće biti identičan hashu pohranjenom u TrustChain lancu blokova te će za provjeru certifikata biti potrebno, na osnovu pohranjenih metapodataka (u nepromjenjivom dijelu TrustChain sustava), pronaći izvornu verziju certifikata iz drugog izvora.

6.3.3. Apache Cassandra varijanta pomoćnog sustava

Apache Cassandra[431] distribuirani je sustav za pohranu podatka temeljen na stupcima (engl. *column*).[432] U osnovi ovo znači da je riječ o sustavu koji grupira podatke na osnovu pripadnosti stupaca određenoj relaciji, to jest rječnikom Cassandra sustava, obitelji stupaca (engl. *column family*). Više ovakvih obitelji čini prostor ključa (engl. *key space*) koji definira pravila replikacije pohranjenih podataka kroz više čvorova.

Zbog svoje distribuirane prirode sustav pomoćne TrustChain baze podatka baziran na Apache Cassandra tehnologiji bio bi vrlo jednostavan za implementaciju. TrustChain *keyspace* (engl.) se jednostavno replicira preko svih TrustChain čvorova. Arhitektura ovog sustava je potpuno kompatibilna s predloženim TrustChain modelom.

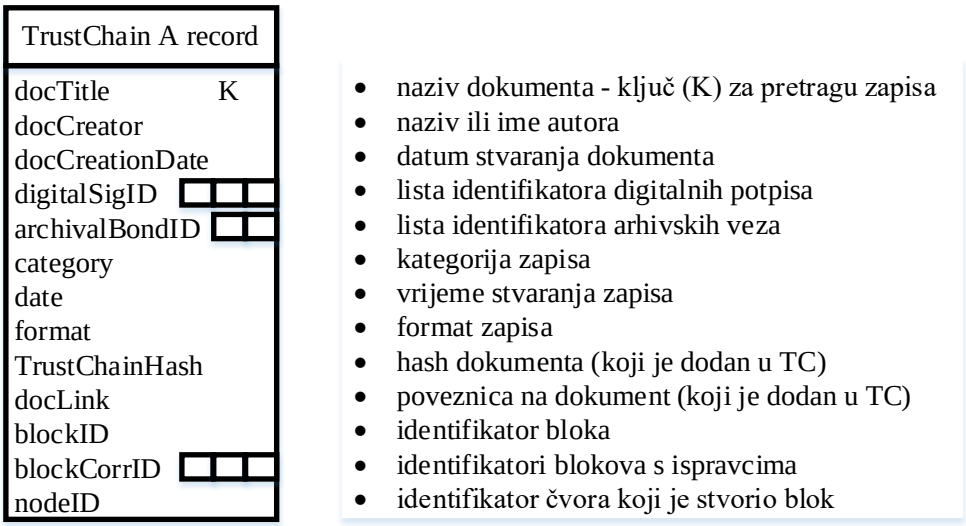
Do komplikacije pri upotrebi Apache Cassandra sustava dolazi zbog nemogućnosti povezivanja više obitelji stupaca pri upitu. Ovo je problem sličan i nedostatku MongoDB sustava koji ne podržava upite preko više dokumenata te zbog toga dolazi do dupliciranja podataka ili potrebe da se upiti riješe programski na logičkoj razini aplikacije. U slučaju Apache Cassandra sustava moguće je ugraditi funkcionalnosti pretrage, to jest povezivanja podataka pri upitima korištenjem dodatnih alata, poput DataStax ODBC (engl. *Open Database*

[431] Apache Cassandra URL: https://cassandra.apache.org/_/index.html
[432] Lakshman, A., & Malik, P. (2010). Cassandra: a decentralized structured storage system. *ACM SIGOPS Operating Systems Review*, *44*(2), 35-40. Preuzeto 7. 1. 2022. s https://www.cs.cornell.edu/projects/ladis2009/papers/lakshman-ladis2009.pdf

Connectivity)[433] i Apache SparkSQL.[434] Upotreba ovih sustava nije predviđena te podatkovni modeli prikazani u nastavku podrazumijevaju upotrebu isključivo Apache Cassandra sustava.

S obzirom na to da Cassandra ne podržava nikakve upite, njezine obitelji stupaca (tablice) je potrebno dizajnirati na način da što je više moguće odgovaraju upitima koji će se tražiti od baze podataka tijekom njezine upotrebe. Ako je ovaj dizajn uspješan rezultat je izuzetno učinkovita baza podataka. U slučaju da nije tako upite koji bili predviđeni treba rješavati na logičkom sloju aplikacije ili koristiti ranije navedene vanjske alate. Iz ovog razloga, kao što je i izvorni prijedlog upotrebe Apache Cassandra sustava uočio[435], izuzetno je teško dizajnirati obitelji stupaca bez konkretnih primjera upotrebe. Usprkos tome moguće je pretpostaviti više obitelji, koje djelomično odgovaraju dokumentima iz MongoDB primjera u prethodnom poglavlju. S obzirom na velike razlike koje su uvedene u sve TrustChain podatkovne modele u odnosu na izvorne modele i ovdje razrađene podatkovne strukture značajno odstupaju od onih predstavljenih 2021.[436]

Slika 38 prikazuje moguću podatkovnu strukturu TrustChain A modela. Kao što je navedeno, ovakav osnovni model vjerojatno nije konačni rezultat izgleda obitelji stupaca u Cassandra sustavu ali je dobra početna točka.

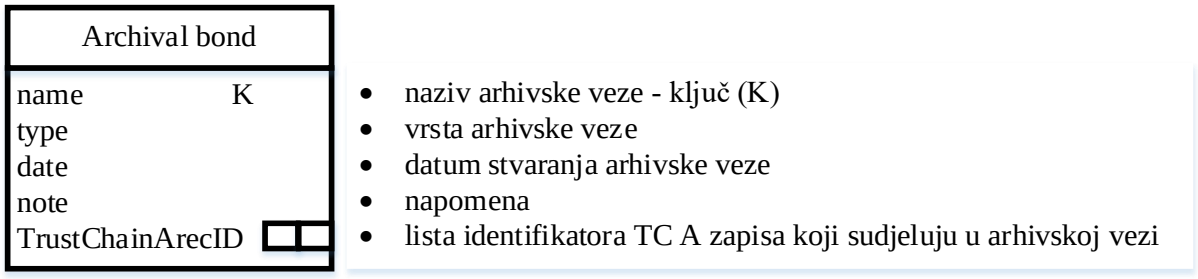


Slika 38. TrustChain Cassandra pomoćna baza podataka – TC A zapis

[433] Kondylakis, H., Fountouris, A., & Plexousakis, D. (2016). Efficient Implementation of Joins over Cassandra DBs. *EDBT—International Conference on Extending Database Technology*, (str. 666-667). doi: http://dx.doi.org/10.5441/002/edbt.2016.77
[434] Armbrust, M., Xin, R., Lian, C., Huai, Y., Liu, D., Bradley, J., . . . Zaharia, M. (2015). Relational data processing in spark. *Proceedings of the 2015 ACM SIGMOD international conference on management of data*, (str. 1383-1394). Preuzeto 27. 11. 2021. s https://dl.acm.org/doi/pdf/10.1145/2723372.2742797
[435] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.
[436] Ibid.

Cassandra shema TrustChain A zapisa sadrži stupce "digitalSigID", `"archivalBondID" i "blockCorrID"` koji sadrže liste identifikatora drugih obitelji, u shemi ovo je naglašeno dodavanjem niza kvadrata iza imena stupca. Ovo je nužno jer dokument može biti potpisan s više potpisa, može sudjelovati u više arhivskih veza te može biti ispravljen u više kasnijih zapisa. Ovo povezivanje treba odraditi na logičkoj razini sustava. Alternativa tome bi bila da se svi relevantni podaci uključe u prikazanu obitelj.

Na slici 39 prikazana je obitelj stupaca koja u sustavu temeljenom na Cassandra Apache tehnologiji omogućava zapis podataka o arhivskoj vezi. Ranije je navedeno da je ovaj zapis bilo moguće uključiti i u samu obitelj stupaca o TrustChain A zapisu (slika 38). Ovo bi rezultiralo dupliciranjem podataka jer je obitelj stupaca o arhivskoj vezi nužna kako bi se ona mogla koristiti kao početna točka pretrage pomoćne baze podataka, što smatram da će biti čest slučaj. Alternativa bi bila i da se potpuni podaci o svakom TrustChain A zapisu koji sudjeluje u arhivskoj vezi uključe u obitelj stupaca koja sadrži podatke o Arhivskoj vezi. Ovo bi rezultiralo još većim dupliciranjem podataka.



Slika 39. TrustChain Cassandra pomoćna baza podataka – zapis o arhivskoj vezi

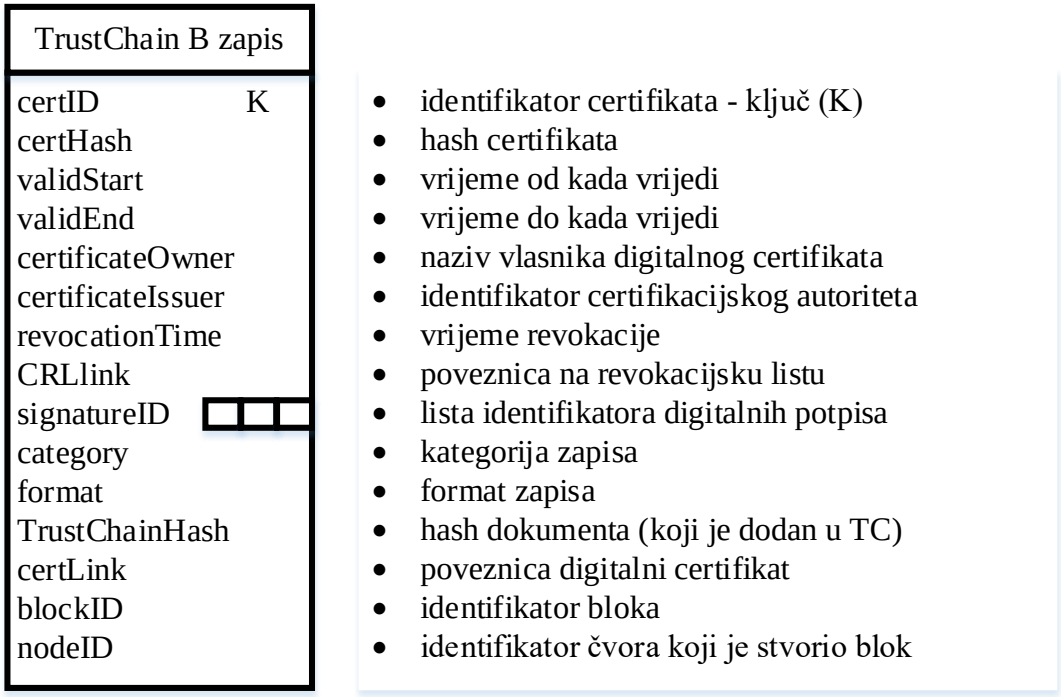
Spomenuti način dupliciranih zapisa, koji je u slučaju Apache Cassandra modela izbjegnut, primijenjen je u ranije prikazanom MongoDB modelu. Tamo ugniježđeni dokumenti često sadrže duplicirane podatke.

Ovo je svjesna odluka potaknuta željom da se paralelno prikazani modeli razlikuju i da budu specijalizirani u skladu s prepoznatim osobinama tehnologije. S obzirom na to da je Apache Cassandra u prethodnim istraživanjima pokazao bolje performanse u slučaju zapisa podataka[437] ovakav način modeliranja obitelji stupaca optimizira upravo proces zapisa novih podataka (jer se minimizira broj dupliciranih podatka koji su karakteristični za noSQL sustave) te naglašava prednosti Cassandra sustava. MongoDB model nije optimiziran na ovaj način, on

[437] Abramova, V., & Bernardino, J. (2013). NoSQL databases: MongoDB vs cassandra. *Proceedings of the international C* conference on computer science and software engineering*, (str. 14-22). Preuzeto 7. 1. 2022. s http://web.cs.wpi.edu/~cs585/s17/StudentsPresentations/This%20Year/Week14/mongodb_vs_cassandra.pdf

je optimiziran za brže pretrage (umjesto brzih zapisa). Vrijeme zapisa će u MongoDB sustavu biti duže a zbog dupliciranja podataka njihova ukupna količina će biti veća. S obzirom da je MongoDB sustav pokazao bolje performanse pri repliciranju podataka[438] on je i pogodan za ovakvu optimizaciju.

TrustChain B Cassandra zapis, prikazan na slici 40 vrlo je sličan ekvivalentnom MongoDB zapisu, ali zbog ranije objašnjenog razloga ne uključuje veliki broj dupliciranih podatka. `"signatureID"` stupac sadrži liste poveznica prema potpisima koji koriste certifikat na koji se zapis odnosi. Ostali podaci su već više puta objašnjeni pa ne zahtijevaju posebna objašnjena.

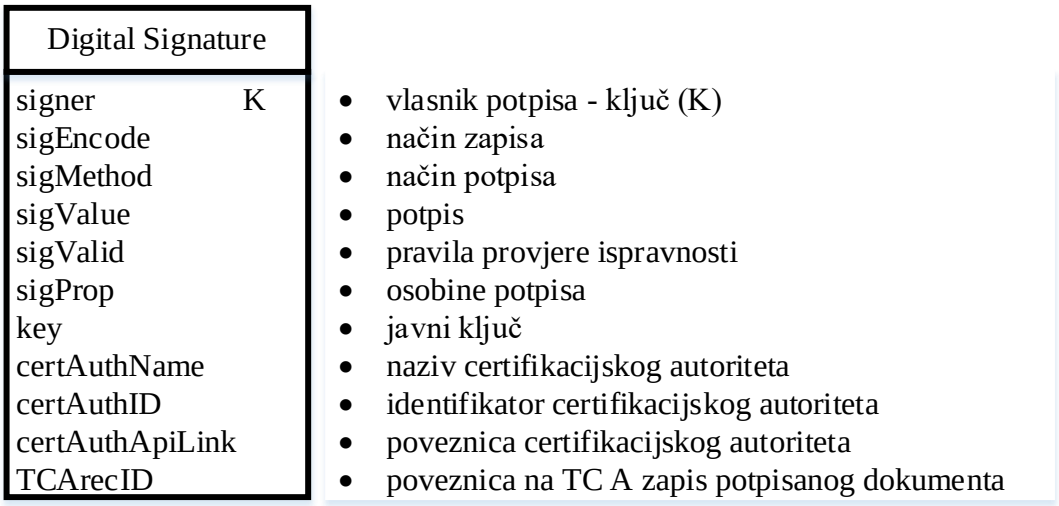


Slika 40. TrustChain Cassandra pomoćna baza podataka – TC B zapis

Ranije spomenuti prikaz obitelji stupaca koji se odnosi na zapis konkretnog digitalnog potpisa prikazan je na slici 41. Ovaj zapis uključuje podatke o digitalnom zapisu. Iako je svaki potpis vezan uz konkretni dokument ovaj zapis je isključen iz obitelji stupaca koja sadrži podatke o TrustChain A zapisu (slika 38) jer svaki dokument može biti potpisan s više potpisa. Iako ovdje nije riječ o dupliciranju podataka (svaki digitalni potpis bi i dalje bio prisutan samo jednom u sklopu jedne obitelji stupaca) ovo izdvajanje je svejedno napravljeno da bi se smanjio zapis i bolje organizirali podaci. Konkretni podaci o digitalnim potpisima se u ovom slučaju

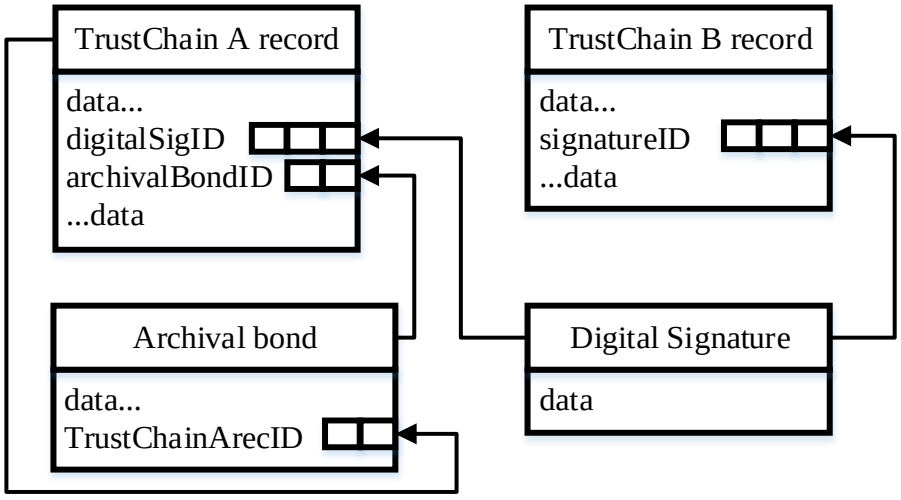
[438] Haughian, G., Osman, R., & Knottenbelt, W. (2016). Benchmarking replication, n. dj.

dohvaćaju kroz odvojeni upit ili na osnovu pregleda samog dokumenta i TrustChain lanca blokova (koji sadrže sve relevantne podatke).



Slika 41. TrustChain Cassandra pomoćna baza podataka – zapis digitalnog potpisa

Odnos ranije objašnjenih obitelji stupaca prikazan je na slici 42. Iako ovaj prikaz podsjeća na relacijski model s referentnim ključevima, ovo su funkcionalnosti koje će morati biti ugrađene u programsku logiku konačnog informacijskog sustava ili realizirane upotrebom pomoćnih (djelomično komercijalnih) sustava.



Slika 42. TrustChain Cassandra pomoćna baza podataka – odnos obitelji stupaca na logičkoj razini

Ovime završava pregled podatkovnih struktura u TrustChain modelu pomoćne baze podatka realizirane upotrebom Cassandra Apache tehnologije. Za razliku od MongoDB modela koji naglašava brzinu pretrage pomoćne baze podataka ovaj model je optimiziran za brzinu zapisa novih i izmjenu postojećih podataka.

6.4. Novi TrustChain model sustava za dugotrajnu pohranu digitalno potpisanih dokumenata i digitalnih certifikata

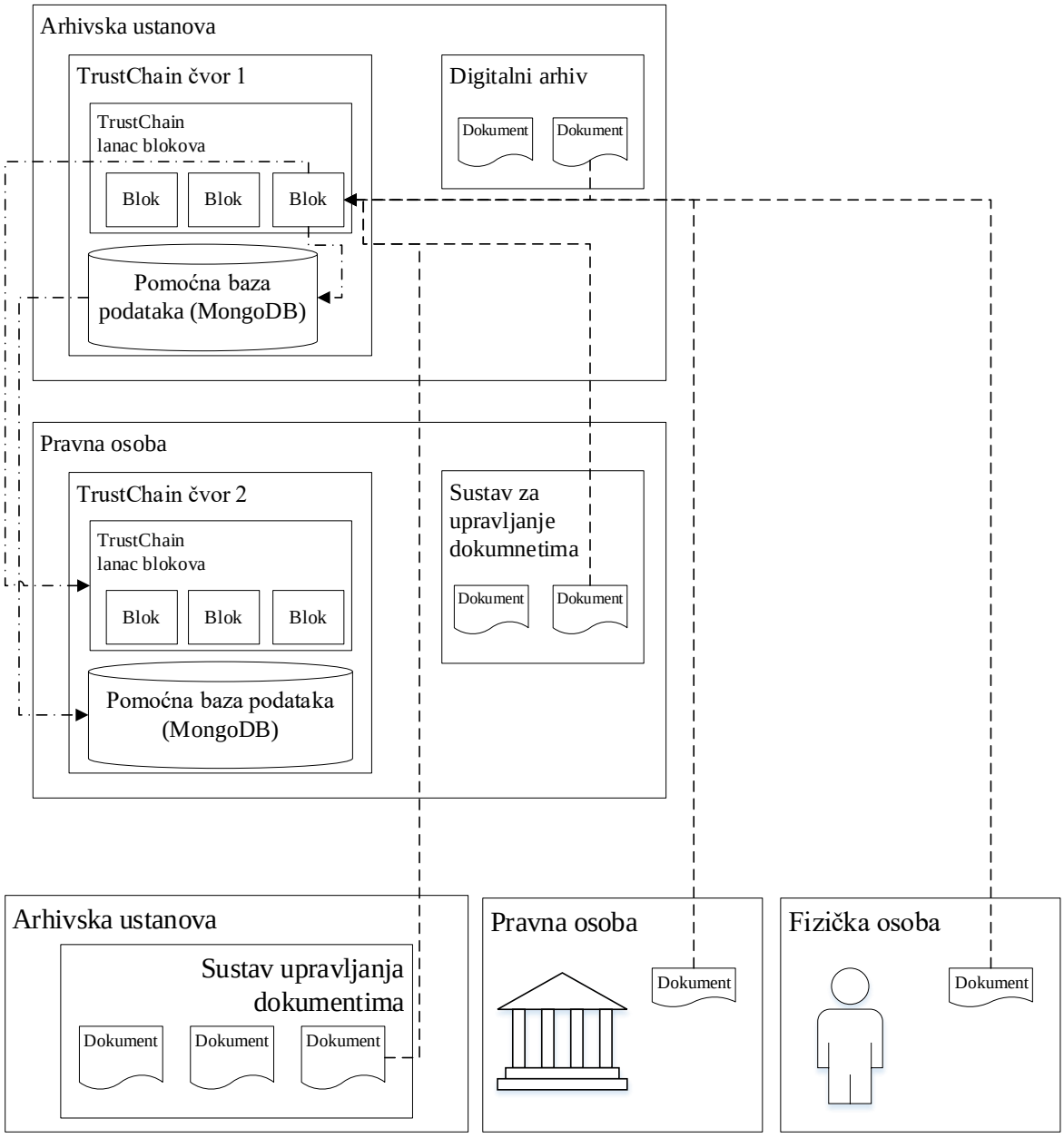
Ovo poglavlje razrađuje končani TrustChain model. Model je dizajniran spajanjem dvaju prethodno prikazanih modela, TrustChain A i TrustChain B u jedan, koji može biti iskorišten kao predložak za projekt informacijskog sustava za dugotrajnu pohranu digitalno potpisanih dokumenta (i digitalnih certifikata). S obzirom da su TrustChain A i B modeli već detaljno predstavljeni u obliku modula ovo poglavlje se fokusira na njihovu interakciju i daje kratku rekapitulaciju i objašnjenje važnijih dijelova TrustChain modela umjesto da u potpunosti predstavlja novi model koji se sastoji od ranije prikazanih dijelova.

Inicijalna premisa koja je potakla razvoj TrustChain B modela, tada zvanog TrustChain 2[439] jest da je TrustChain A model[440] neprilagođen upotrebi u državnim arhivima (ili drugim institucijama koje arhiviraju povjerljive dokumente). Ovo je istina ali sam nakon razvoja TrustChain B modela zaključio da on ne može biti adekvatna zamjena. TrustChain B model je u svojoj srži arhiv digitalnih certifikata, on tek posredno i samo djelomično ispunjava osnovni cilj TrustChain modela koji je: dugotrajna pohrana digitalno potpisanih dokumenata. Kao takav, TrustChain B model je dopuna A modela koja dozvoljava da se djelomično provjeri i djelomično garantira integritet zapisa koji inače ne mogu biti podvrgnuti provjeri od strane više ustanova. Osnovna funkcionalnost digitalnog arhivskog sustava koji čuva autentičnost i integritet digitalno potpisanih dokumenta je sadržana u TrustChain A modelu, TrustChain B model je odličan dodatak osnovnom modelu koji omogućava rješavanje rubnih slučajeva (vezanih uz povjerljivost dokumenata) te dodaje arhiv digitalnih certifikata. Taj arhiv digitalnih certifikata omogućuje funkcioniranje TrustChain B modula ali je i izvrstan dodatak funkcionalnosti osnovnog TrustChain A modela koji sada više ne pohranjuje samo digitalni potpis već i pripadajući certifikat te na taj način u potpunosti čuva podatke koji su izvorno korišteni za stvaranje arhiviranog dokumenta. Zajedno, na način koji je razrađen u ovoj

[439] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving, n. dj.
[440] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.

disertaciji, modeli čine cjelinu sa značajno više funkcionalnosti nego što su pružali u svojem odvojenom obliku.

Shema na slici 43 kroz apstrakciju toka podataka prikazuje osnovno funkcioniranje i svrhu TrustChain sustava.



Slika 43. TrustChain model – osnovni tok podataka

Shema na slici 43 prikazuje odnos institucija izvan i u TrustChain sustavu, to jest tok podataka od izvornog dokumenta do upisa u TrustChain lanac blokova. Isprekidane linije prikazuju proces skupljanja dokumenata za koje je traženo uvrštavanje u TrustChain lanac blokova. U ovom slučaju TrustChain čvor 1, koji je smješten u arhivskoj ustanovi, skuplja dokumente te, koristeći ranije opisane procese formira novi TrustChain blok. Na osnovu ovog

zapisa stvaraju se i zapisi u pomoćnoj bazi podataka te se svi podaci (TrustChain lanac blokova i pomoćna baza podataka) repliciraju na druge čvorove (isprekidana linija s točkama).

Shema prikazuje tri sudionika procesa koji nisu dio TrustChain sustava. Dizajnirani model može biti korišten kao zatvoreni sustav, koji koriste isključivo ustanove koje sudjeluju kao čvorovi, ili kao sustav u kojem TrustChain čvorovi obrađuju vlastite zahtjeve, ali i zahtjeve ustanova, pravnih i fizičkih osoba koje ne sudjeluju u TrustChain sustavu. Tri ovakva sudionika prikazana su na donjem dijelu slike 43. Ovakav, otvoreni pristup, otvara prostor za komercijalno iskorištavanje sustava. Ovisno o načinu financiranja uspostave i održavanja sustava ovo će možda biti potrebno te model podržava, ali ne zahtijeva takvu primjenu. Niti je nužno da sustav temeljen na TrustChain modelu bude otvoren za vanjske zahtjeve, niti je nužno da takvi zahtjevi budu naplaćeni. Ovo je ostavljeno kao mogućnost o kojoj će se odlučiti pri projektiranju konkretnih implementacija modela.

Prethodna shema navodi pomoćnu bazu podataka temeljenu na MongoDB tehnologiji. Odlučio sam se za ovu opciju jer smatram da TrustChain sustav u dogledno vrijeme neće doseći milijune transakcija po sekundi koji bi bili potrebni da se vidi razlika u performansama koju su pokazala prethodna istraživanja.[441, 442] Umjesto toga smatram da su konzistencija zapisa i brža replikacija pomoćne baze, koju postiže MongoDB sustav,[443] važniji. Konzistencija zapisa između pomoćne baze i lanca blokova može biti postignuta upotrebom UBJSON[444] (engl. *Universal Binary JavaScript Object Notation*) ili CBOR[445] (engl. *Concise Binary Object Representation*) standarda za formiranje konačnih datoteka koje tvore lanac blokova. Oba standarda omogućuju jednostavnu serijalizaciju JSON formata u binarne datoteke. Ovakva serijalizacija podataka uvelike će smanjiti veličinu datoteka u usporedbi sa standardnim JSON tekstualnim zapisom. Ovo je opravdan razlog za upotrebu, na taj način se omogućuje kompaktnija pohrana najvažnijeg dijela sustava (ulančanih blokova). UBJSON je logičan izbor jer se temelji je na BSON standardu[446] kojeg koristi MongoDB baza podataka koja je preferirani izbor za pomoćnu bazu podataka. Nažalost, iako je u širokoj upotrebi, te je već primijenjen u sustavima koji koriste ulančane blokove,[447] UBJSON još uvijek nije potpuno standardiziran. Iz

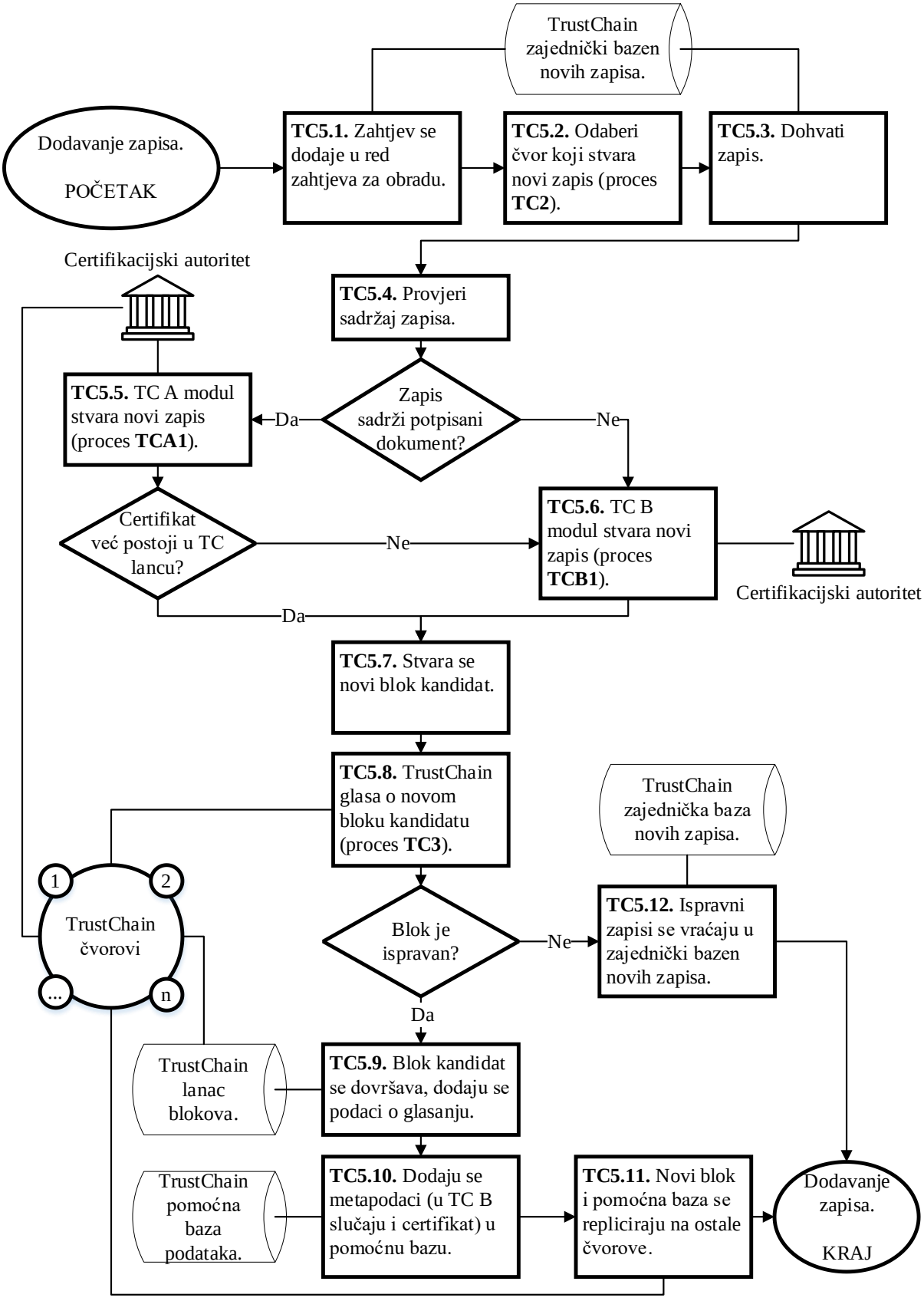
[441] Abramova, V., & Bernardino, J. (2013). NoSQL databases: MongoDB vs cassandra, n. dj.
[442] Ramesh, D., Sinha, A., & Singh, S. (2016). Data modelling for discrete time series, n. dj.
[443] Haughian, G., Osman, R., & Knottenbelt, W. (2016). Benchmarking replication, n. dj.
[444] UBJSON URL: https://ubjson.org/
[445] CBOR URL: https://cbor.io/
[446] Universal Binary JSON Specification. (2021). *Universal Binary JSON Specification*. (U. B. Specification, Producent) Preuzeto 11. 29. 2021. s https://ubjson.org/
[447] Ismailisufi, A., Popović, T., Gligorić, N., Radonjic, S., & Šandi, S. (2020). A private blockchain implementation using multichain open source platform. *2020 24th International Conference on Information*

ovog razloga odabir konkretnog standarda ostavljen je za trenutak izrade konkretnog TrustChain projektnog rješenja. U slučaju formalne standardizacije UBJSON zapisa, kroz objavljeni RFC ili ISO standard, preporuča se njegova upotreba. U suprotnom, nešto stariji CBOR je također dobar izbor. CBOR je zamišljen kao univerzalni način serijalizacije JSON datoteka[448] koji nije temeljen na BSON-u, ali je i dalje dovoljno sličan te bi i njegova upotreba dozvolila konzistenciju u zapisima.

Na slici 44 prikazan je proces TC5 koji opisuje dodavanje novog bloka s gledišta procesa za razliku od slike 43 koja stavlja naglasak na tok podatka. Slika opisuje proces dodavanja novog zapisa koji obuhvaća dodavanje zapisa u TrustChain A i B modul. Ovaj proces u svojim postupcima referencira prethodno opisane procese. Na primjer, postupak TC5.5 referencira ranije opisani proces TCA1 te je njegova apstrakcija. Na ovaj način u relativno kompaktnom dijagramu toka je prikazan potpuni proces dodavanja zapisa u novi TrustChain model kao i interakcija ranije prikazanih procesa i elemenata sustava (pomoćna baza, lanac bloka, TrustChain čvorovi i drugi) na visoko apstraktnoj razini. S obzirom na to da su pri ranijoj razradi TrustChain modula procesi detaljno objašnjeni za konačni model preostao je samo ovakav prikaz interakcije dvaju modela i njihovih procesa.

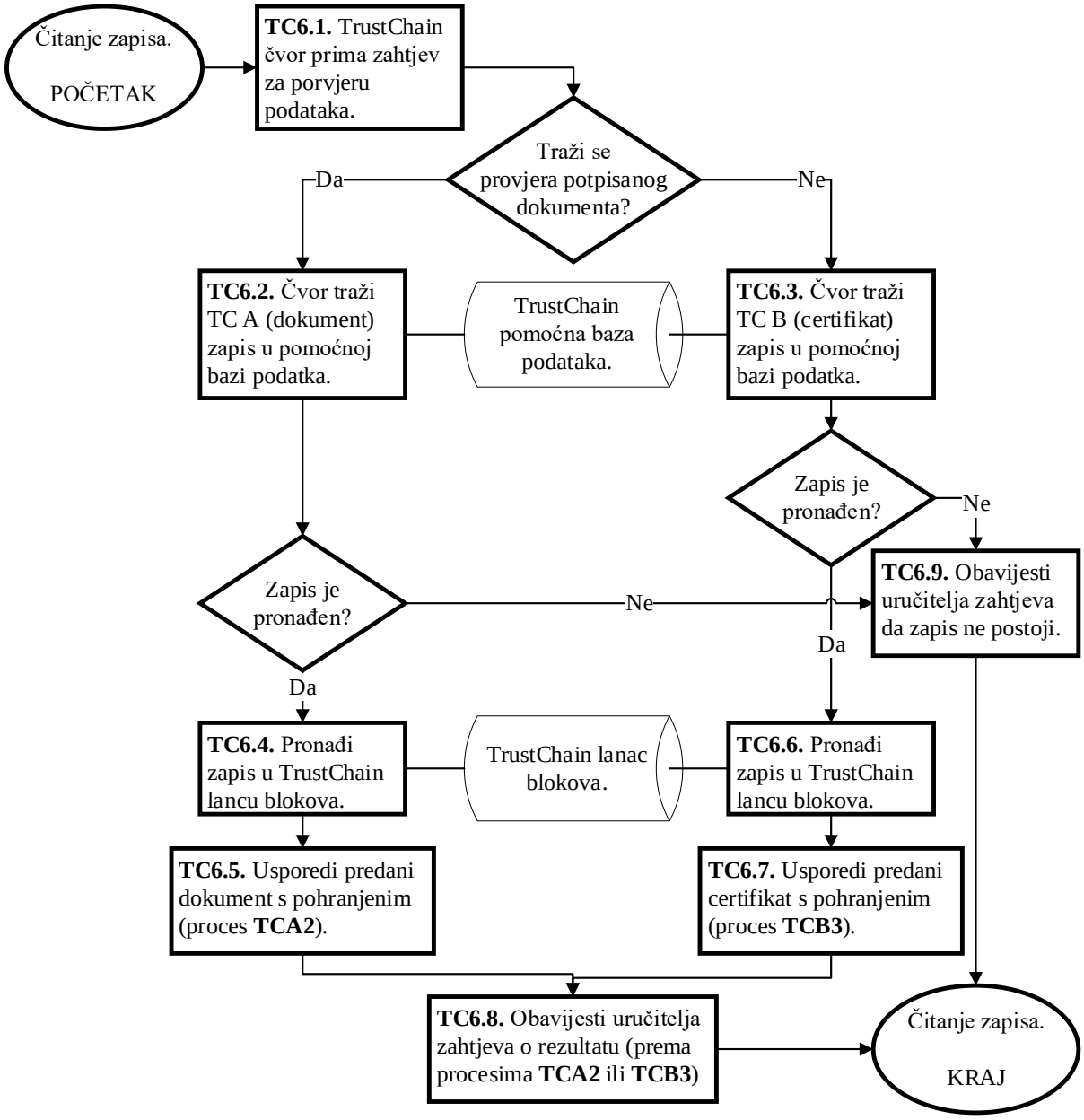
Technology (IT) (str. 1-4). Žabljak: IEEE. Preuzeto 29. 11. 2021. s https://www.researchgate.net/profile/Nenad-Gligoric/publication/339442490_A_Private_Blockchain_Implementation_Using_Multichain_Open_Source_Platform/links/5e73dd6d458515c677c6213d/A-Private-Blockchain-Implementation-Using-Multichain-Open-Source-Platform.pdf

[448] Bormann, C. a. (12 2020). *RFC 8949: Concise binary object representation (cbor).* Preuzeto 11. 29. 2021. s IETF: https://datatracker.ietf.org/doc/html/rfc8949



Slika 44. Proces TC5 – TrustChain proces dodavanja novog bloka

Slika 45 na istoj, visokoj razini prikazuje proces čitanja podatka iz novog TrustChain modela koji obuhvaća oba modula i pomoćnu bazu podataka.

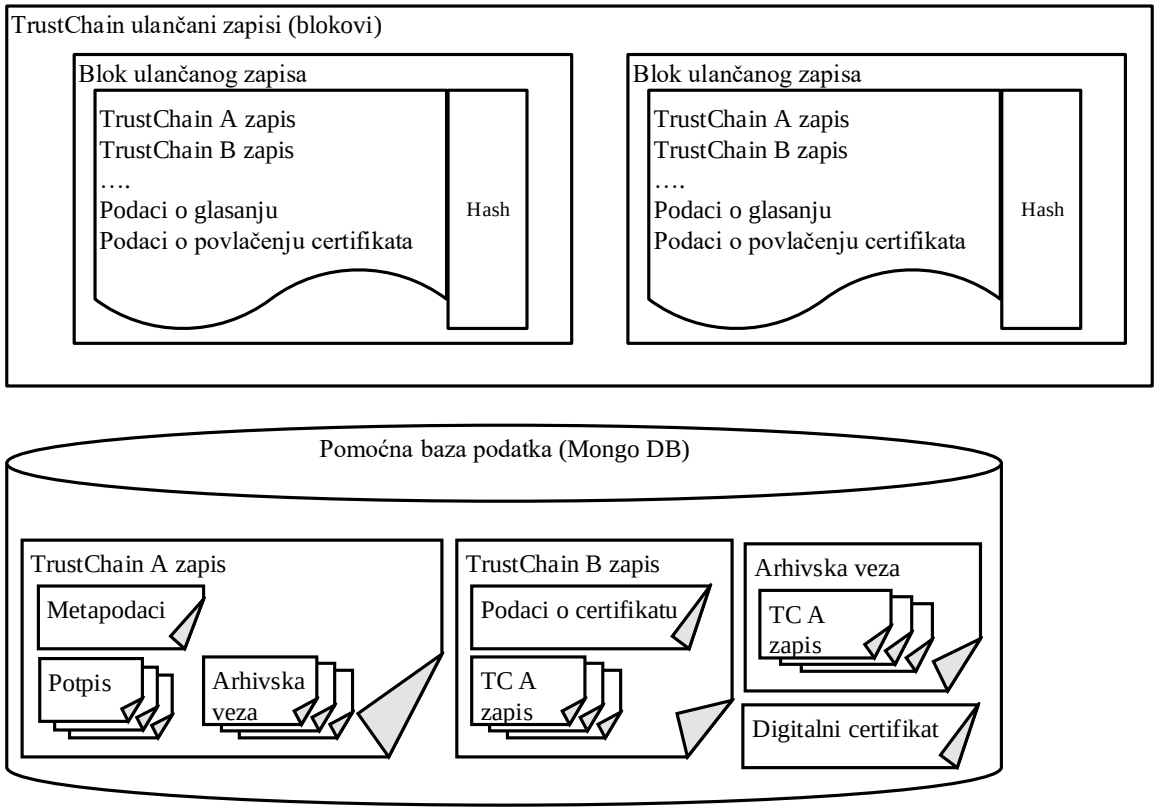


Slika 45. Proces TC6 – čitanje zapisa iz TrustChain sustava

Proces TC6 objašnjava, na najvišoj razini, način na koji novi TrustChain model čita podatke, to jest odgovara na zahtjeve provjere autentičnosti dokumenta ili ispravnosti digitalnog certifikata u nekom trenutku u prošlosti. Za pretpostaviti je da provjera autentičnosti zapisa ili certifikata u konačnici neće biti jedina vrsta zahtjeva na koju informacijski sustav temeljen na TrustChain modelu može odgovoriti. S obzirom na količinu pohranjenih podataka i metapodataka gotovo sigurno će se s vremenom pojaviti i potreba za jednostavnijim upitima. Slično u slučaju ranije prikazanog procesa zapisa podatka nisu pokriveni slučajevi u kojima se

stvaraju novi zapisi. Sasvim sigurno nedostaju procesi koji mijenjaju samo podatke u pomoćnoj bazi podataka (bez stvaranja novih ulančanih blokova). Svi ovi procesi su ovisni o konkretnoj implementaciji i previše detaljni za razinu modela sustava te će biti razrađeni u projektnoj dokumentaciji eventualne implementacije TrustChain modela.

Osim procesa pisanja i čitanja podataka na najvišoj razini treba prikazati i raspraviti konačnu podatkovnu strukturu novog TrustChain modela. O podatkovnim strukturama TrustChain modela već je bilo dosta riječi pa će se ovdje prokomentirati samo njihove najvažnije elemente. Shema podatkovnih struktura prikazana na slici 46 pretpostavlja da je prema ranijoj preporuci upotrijebljen MongoDB distribuirani sustav za pohranu podatka.



Slika 46. TrustChain prikaz podatkovnih struktura

Podaci prikazani na slici 46 podijeljeni su u dvije nezavisne cjeline:

- TrustChain ulančani blokovi. Ovo je podatkovna struktura temeljena na principu ulančanih zapisa opisanom u četvrtom poglavlju te realizirana upotrebom UBJSON ili CBOR standarda za binarni zapis JSON podataka. TrustChain lanac blokova sadrži najvažniji dio TrustChain modela: hasheve digitalno potpisanih dokumenta i digitalnih certifikata. Ova nepromjenjiva podatkovna struktura je relativno mala te je za očekivati da će i nakon godina upotrebe sustava temeljenog na TrustChain modelu ostati kompaktna i jednostavna za replikaciju, prijenos i sigurnu pohranu na razne načine.

- Pomoćna baza podataka. Ova distribuirana noSQL baza podataka sadrži širi set metapodataka od TrustChain lanca blokova koji omogućuju proširivanje i izmjenu metapodataka TrustChain lanca blokova (zbog ljudske greške ili zahtjeva evidencije podataka o arhivskoj vezi) te brže pretraživanje podataka nego što je to moguće koristeći isključivo TrustChain lanac blokova. Osim navedenog pomoćna baza podataka sadrži i digitalne certifikate čiji su hashevi zapisani u TrustChain lanac blokova kao i sve ostale podatke koji su potrebni za funkcioniranje sustava temeljnog na TrustChain modelu, a koji ovdje nisu prikazani jer ovise o konkretnoj implementaciji modela.

Obje cjeline se u potpunosti repliciraju na svim čvorovima koji sudjeluju u TrustChain sustavu. Pomoćna baza podataka temeljena na MongoDB tehnologiji već posjeduje sve mehanizme potrebne za ovakvu replikaciju, a replikacija TrustChain lanca blokova će biti riješena na programskoj razini sustava temeljenog na ovom modelu. Druga opcija je da se i sami TrustChain blokovi zapišu u MongoDB sustav te da se na taj način uštedi na količini programskog koda koji je potrebno stvoriti pri implementaciji modela. Nisam sklon ovom rješenju jer stvara preveliku ovisnost sustava o MongoDB tehnologiji. Ako je komponenta ulančanih blokova ostavljena izvan MongoDB sustava onda su ti podaci i dostupniji drugim sustavima koji mogu biti u interakciji s TrustChain sustavom. Ovisno o konačnoj implementaciji modela, to jest o odluci koliko zapisa (iz zajedničkog bazena novih zapisa) će biti uključeno u svaki blok moguće je da će individualni blokovi biti veći od 16MB, što je maksimalna veličina BSON datoteke koju MongoDB sustav može pohraniti.[449] Pohrana datoteka većih od 16MB podrazumijeva da će one biti podijeljene u manje dijelove.

6.5. Zaključak

U poglavlju je detaljno razrađen novi model TrustChain sustava za dugotrajnu pohranu digitalno potpisanih zapisa i digitalnih certifikata. Razrađeni model razvijen je na osnovu prethodnih istraživanja koja su rezultirala TrustChain 1 modelom,[450] koji je ovdje značajno unaprijeđen i prilagođen radu u sustavu koji podržava i funkcionalnosti TrustChain 2 varijante modela.[451] S obzirom na to da je TrustChain 1 model nužan da bi arhivski zahtjevi za ovakav sustav bili potpuno zadovoljeni, a TrustChain 2 model, iako nije savršen, je jedino moguće

[449] MongoDB Inc. (2020). MongoDB Documentation, n. dj.
[450] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.
[451] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving, n. dj.

rješenje (u kontekstu TrustChain modela, to jest sustava ovisnog o glasanju više čvorova) za dugotrajno očuvanje autentičnosti povjerljivih dokumenta oba modela su značajno modificirana te je na osnovu njih razvijen novi TrustChain model koji obuhvaća funkcionalnosti obaju sustava.

Novi model zadržao je fleksibilnost starih modela. Kao i ranije sustav je moguće implementirati u samo jednoj (arhivskoj) ustanovi, ali u tom slučaju se toj ustanovi treba u potpunosti vjerovati pri provjeri ispravnosti potpisa i certifikata. Pouzdanost procesa provjere ispravnosti digitalnog certifikata dolazi do izražaja kada se koristi sustav glasanja, to jest kada u sustavu sudjeluje više ustanova. Ne postoji maksimalni broj ustanova koje mogu sudjelovati u sustavu prema razvijenom modelu. Model omogućava zatvoren ili otvoren sustav. Ovisno o željama i potrebama ustanova koje sudjeluju u njemu sustav može ali ne mora opsluživati zahtjeve za dodavanjem novih i provjerom starih zapisa od strane pravnih i fizičkih osoba koje ne sudjeluju u TrustChain sustavu. U slučaju otvorenog sustava ovakve, vanjske zahtjeve je moguće naplaćivati i na taj način stvoriti implementaciju modela koja je komercijalna.

Razvijeni model nije potpuna projektna dokumentacija potrebna za razvoj novog informacijskog sustava. Razvoj konkretnog informacijskog sustava temeljenog na ovom modelu zahtjeva razradu značajnog broja procesa i podatkovnih struktura koje ovdje nisu pokrivene kao i donošenje odluka o postojećim elementima modela koje trenutno nisu moguće. Na primjer, veličina TrustChain ulančanog bloka ovisi o broju zapisa u njemu. Ova veličina bi trebala biti konstantna i izravno ovisi o potrebama i broju ustanova koje sudjeluju u TrustChain modelu. Ovakve informacije nisu dostupne prije identifikacije ustanova koje žele sudjelovati u razvoju sustava. Svrha modela je da pruži početnu, konceptualnu točku za razvoj takvog projekta i da pokaže da je moguće upotrijebiti ulančane blokove da bi se produžila mogućnost provjere autentičnosti digitalno potpisanih zapisa. To jest, da potvrdi ili opovrgne hipoteze istraživanja.

7. Zaključak i rasprava

Cilj ovog istraživanja bio je razviti novi model za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva. Svrha ovog modela je unapređenje arhivistike na način da se njeni koncepti prilagode digitalno potpisanom gradivu te da se pokaže na koji način je moguće upotrijebiti kriptografske algoritme da bi se zadovoljilo zahtjeve arhivistike za dokazivanje autentičnosti.

Uz ovo, postavljene su dvije hipoteze koje su trebale biti potvrđene ili opovrgnute novim modelom i pratećim istraživanjima. One su:

[1] Upotrebom ulančanih zapisa produžuje se dokazivost autentičnosti digitalno potpisanoga arhivskoga gradiva.

[2] Sustavi za pohranu digitalno potpisanoga arhivskoga gradiva utemeljeni na ulančanim zapisima omogućuju očuvanje arhivske veze.

Obje hipoteze povezane su uz ulančane blokove, posebnu podatkovnu strukturu koja ima dugu tradiciju upotrebe u sustavima za vremenske žigove. Svrha prve hipoteze je pokazati je li moguće sličan princip primijeniti na digitalno potpisane zapise (a ne samo na vremenske žigove) i time omogućiti dugotrajno očuvanje autentičnosti ovih zapisa. Druga hipoteza postavlja pitanje omogućuje li ta ista podatkovna struktura (ulančani blokovi) razvoj sustava koji omogućuje očuvanje arhivske veze.

Prije pristupanja razvoju novog modela provedeno je opsežno istraživanje postojeće literature iz područja arhivistike, diplomatike, kriptografije, industrijskih i zakonskih normi povezanih s digitalnim certifikatom i potpisom te postojećih sustava. Svrha ovog teorijskog istraživanja bila je:

- pokazati koje zahtjeve arhivistika postavlja modelu za dugoročnu pohranu digitalno potpisanoga arhivskoga gradiva,
- ispitati je li dokazivanje autentičnosti (prema zahtjevima arhivistike) uopće moguće upotrebom digitalnog potpisa,
- istražiti na koji način podatkovna struktura ulančanih blokova omogućuje dugoročno očuvanje autentičnosti te

- istražiti zašto postojeća suvremena rješenja za dugoročnu pohranu u potpunosti ne zadovoljavaju zahtjeve arhivistike.

U prvom dijelu teorijskog istraživanja razmotreni su temeljni koncepti arhivistike i diplomatike. U ovom dijelu utvrđeni su kriteriji za dokazivanje autentičnosti arhivskog zapisa, to jest što mora omogućiti model za dugotrajnu pohranu svih, a ne samo digitalno potpisanih zapisa. Ovi zahtjevi (dokazivanja autentičnosti) su:

1) da zapis sadrži vrijeme nastanka,
2) da zapis sadrži potpis koji identificira autora,
3) da je zapis od trenutka potpisivanja ostao nepromijenjen i
4) da je održano sudjelovanje zapisa u arhivskim vezama.

Ako informacijski sustav temeljen na novom modelu može zadovoljiti zahtjeve 1 do 3 možemo smatrati da je prva hipoteza potvrđena. Posredno, osim novog modela ove zahtjeve mora ispuniti i suvremeni sustav za digitalne potpise. Ako novi model zadovolji i 4. zahtjev i druga hipoteza će biti potvrđena. Tijekom ove faze utvrđeno je da je arhivska veza "…mreža odnosa svakog zapisa s drugim zapisima iz istog skupa"[452] te da je ona dinamična i da se formira dok god se stvaraju novi zapisi koji su relevantni za kontekst u kojem je nastao zapis. Iako se još nije prešlo na razvoj novog modela, već u ovoj ranoj fazi pojavila se skepsa prema sustavu temeljenom na ulančanim blokovima u pogledu ispunjavanja ovog uvjeta. Ulančani blokovi su (trajno) statična podatkovna struktura, a arhivska veza je dinamični odnos.

U istoj fazi istraživanja razmotrena je situacija sa zakonskim zahtjevima vezanima uz rok čuvanja zapisa i trajanjem digitalnih certifikata te je ukazano na drastičnu razliku u ovim vremenskim rasponima. Neke zapise nužno je čuvati 10 ili više godina, ali potpisni digitalni certifikati koje izdaju agencije Republike Hrvatske (službeni certifikaciji autoriteti) traju dvije ili, u posebnim slučajevima, tri godine. Istovremeno je pokazano i da je Republika Hrvatska od Europske Unije prihvatila legislativu prema kojoj je digitalni potpis (reguliran Uredbom eIDAS) istovjetan vlastoručnom. Istekli certifikat znači da više ne postoji dokazivosti identiteta autora te zapis više nije autentičan. Zbog ovih razloga nužno je bilo pronaći adekvatno rješenje za dugoročnu pohranu. Dok trenutno rješenje predstavlja upotreba vremenskih žigova, ovo istraživanje ponudilo je model novog sustava.

[452] Duranti, L. (1997). The archival bond, n. dj.

Teorijski pregled arhivističke teorije, u svom zadnjem dijelu, pružio je i sastav minimalnog skupa metapodataka koje novi model mora uključiti. Ovaj skup stvoren je na način da je kompatibilan s više priznatih arhivskih standarda uključujući PREMIS, ISAD(G) i Dublin Core.

Na ovaj način prva faza istraživanja postavila je temelje za daljnja razmatranja. Postalo je jasno zašto je nužan novi model (nesrazmjer u životnom vijeku digitalnog certifikata i zahtjev za dugotrajnim čuvanjem digitalno potpisanih zapisa) i koje funkcionalnosti on mora pružiti (zahtjevi arhivistike navedeni su ranije).

U idućem dijelu istraživanja razmotreni su kriptografski algoritmi koji omogućuju digitalni potpis (SHA hash i RSA algoritam za asimetrično šifriranje) i ulančane blokove (SHA hash algoritmi). Uz njih razmotren je sadržaj standardnog digitalnog certifikata, x.509, i standardi kojima je prema Uredbi eIDAS regulirana primjena digitalnog potpisa pri potpisivanju digitalnih zapisa. Na osnovu ovih saznanja moguće je:

- tvrditi da postojeći kriptografski algoritmi, regulirani eIDAS i drugim uredbama, omogućavaju dokazivanje autentičnosti digitalnog zapisa prije isteka digitalnog certifikata jer ispunjavaju uvjete definirane u prvoj fazi istraživanja u točkama 1 do 3,
- postaviti temelje novog modela, prije svega podatkovnih struktura koje on koristi, na osnovu prikazanog sadržaja digitalnog certifikata te
- pokazati na koji način se provjerava ispravnost digitalnog potpisa i na koji način se on može izlučiti iz ostatka zapisa. Ovi postupci su temeljne pretpostavke funkcioniranja novog modela.

Četvrto poglavlje završilo je teorijski dio istraživanja proučavanjem literature vezane uz ulančane blokove, temeljnu podatkovnu strukturu novog modela. Pokazano je da ova struktura, u kombinaciji s ranije raspravljenim hash algoritmima, garantira nepromjenjivost prethodno stvorenih zapisa te tako pomaže ispuniti jedan od zahtjeva arhivistike, da je zapis od trenutka potpisivanja ostao nepromijenjen. U ovom slučaju to se odnosi na nepromjenjivost od trenutka ulaska zapisa u arhivski sustav temeljen na novom modelu. Prije toga nepromjenjivost je garantirana digitalnim potpisom. U tijeku ovog istraživanja odabrana je lista hasheva kao podvrsta ulančanih blokova koja čini temelj za nepromjenjivu podatkovnu strukturu novog modela. Osim ove podvrste raspravljene su i druge uobičajene podvrste, ali je utvrđeno da sve

imaju za cilj poboljšanje performansi sustava što u ovom modelu nije nužno. Arhivski sustavi ne trebaju biti dizajnirani za brzi pristup (ili čak pristup u realnom vremenu) te su komplikacije koje dolaze sa složenijim podatkovnim strukturama poput stabla hasheva izbjegnute.

Osim literature koja razmatra ulančane blokove u univerzalnom kontekstu, istražena su i prethodna istraživanja upotrebe ulančanih blokova u arhivske svrhe. Ulančani blokovi oduvijek imaju mjesto u arhivskim sustavima kroz sustave za vremenske žigove, kako su izvorno predložili Haber i Stornetta,[453] ali je nedavno započet cijeli niz istraživanja, potaknutih uspjehom primjene ulančanih zapisa na kriptovalute, koja razmatraju ima li ova podatkovna struktura veći potencijali u arhivskim sustavima. Razvoj TrustChain modela samo je jedno u nizu istraživanja novih načina upotrebe ove stare tehnologije u arhivskim sustavima.

Zadnji korak prije prelaska na razvoj novog modela bila je analiza postojećih sustava za vremenske žigove, koji su trenutno najbolje dostupno rješenje problema dugoročnog očuvanja digitalno potpisanoga arhivskoga gradiva. Tijekom analize ovih sustava utvrđeno je da oni garantiraju nepromjenjivost podataka kroz vrijeme, kao i njihovo postojanje u točno određenoj vremenskoj točki te su neki od njih prilagođeni Uredbi eIDAS i uspješno upotrijebljeni u razne arhivske svrhe. Najbolji primjer ovako uspješnog sustava za vremenske žigove je GuardTime[454] sustav koji je uspješno implementiran u više državnih institucija Estonije.[455] Iako je ovaj sustav odličan dokaz primjenjivosti tehnologije ulančanih blokova u dugotrajnoj pohrani, nažalost, nije dostupna dovoljno detaljna dokumentacija njegovog funkcioniranja. Uz to, GuardTime, kao ni ostali sustavi za vremenske žigove ne omogućuju potpunu provjeru i očuvanje autentičnosti, jer kao što je ranije navedeno, oni zanemaruju identitet autora zapisa. Kao što je navedeno u prvom dijelu istraživanja, prema Stančiću:

> *"Provjera autentičnosti bi, kad je riječ o uključivanju zapisa u sustav za očuvanje na dulji vremenski rok, trebala biti omogućena najmanje u dvije situacije. Jednom, prilikom prihvata kako bi se provjerio identitet i integritet dostavljenih zapisa, te drugi puta na zahtjev korisnika prilikom pristupa očuvanim zapisima."* (Stančić H., 2006)[456]

Provjera autentičnosti mora biti omogućena i prilikom ulaska zapisa u arhiv i u naknadnim provjerama koje mogu biti inicirane bilo kada od strane korisnika. Vremenski žigovi

[453] Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document, n. dj.
[454] GuardTime AS. (2021). *GuardTime,* n. dj.
[455] Vatsa, V. R., & Chhaparwal, P. (2021, July). Estonia's e-governance, n. dj.
[456] Stančić, H. (2006). Teorijski model postojanog očuvanja autentičnosti, n. dj.

ovo ne omogućuju jer ne potvrđuju identitet autora zapisa te je ovo, u slučaju njihove upotrebe, prepušteno drugim mehanizmima arhivske ustanove. TrustChain model uključuje ovu funkcionalnost.

Usprkos ovim nedostatcima, nedvojbena je pozitivna uloga sustava za vremenske žigove pri dugoročnom očuvanju arhivske građe. Njihova uloga u dokazivanju integriteta podataka je neosporna i regulirana Uredbom eIDAS. Možemo zaključiti da ovi sustavi (ponovno) dokazuju staru tvrdnju Habera i Storentte o mogućnosti upotrebe ulančanih zapisa kako bi trajno zaključali podatke i na taj način osigurali njihovu nepromjenjivost.

Osim toga, vremenski žigovi su od izuzetne vrijednosti za ovo istraživanje jer ih je moguće uključiti u novi model. Upotrebom eIDAS prilagođenog vremenskog žiga može se značajno povećati povjerenje u vjerodostojnost TrustChain blokova te eliminirati potrebu za razvojem posebnih, vlastitih mehanizama koji bi garantirali ispravnost navedenog vremena stvaranja ulančanih blokova.

Zadnje važno saznaje iz ove faze istraživanja je potvrda osnovne arhitekture TrustChain modela koji se pri dokazivanju autentičnosti ne oslanja samo kriptografske mehanizme (kao što je slučaj sa sustavima za vremenske žigove) već i u povjerenje u (arhivske) institucije koje konsenzusom potvrđuju ispravnost digitalnih zapisa i potpisa. Vigil i suradnici proveli su veliku komparativnu analizu sustava za dugotrajnu pohranu te na kraju utvrdili da se: "…pronalazak rješenja koja zadovoljavaju dugoročne potrebe za povjerljivošću, druge zaštitne ciljeve te balansa između kriptografske zaštite i oslanjanja na ustanove kojima se vjeruje se doima kao zanimljiv i važan pravac istraživanja."[457]

Zadnji dio istraživanja dovršio je razvoj novog modela za dugoročnu pohranu digitalno potpisanoga arhivskoga gradiva. Ovaj model, koji se temelji na ranijim modelima sjedinjuje i razrađuje izvorne ideje modela. Novi model jasno dijeli dvije izvorne ideje te pruža mehanizme za njihovu interakciju. Novi model podijeljen je na TrustChain A i TrustChain B modul.

Prvi način rada novog modela, TrustChain A modul, provjerava ispravnost potpisa i identitet autora upotrebom kriptografskih rješenja te u slučaju postojanja konsenzusa o ispravnosti podataka, izlučeni potpis pohranjuje u nepromjenjivu podatkovnu strukturu – ulančane zapise. Ovaj način rada temeljen je na izvornom TrustChain modelu.[458]

[457] Vigil, M., Cabarcas, D., Wiesmaier, A., & Buchmann, J. (2011). Authenticity, integrity and proof of existence, n. dj.
[458] Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation, n. dj.

Ovaj modul, nakon što se uzme u obzir ostatak ovog istraživanja koji je zorno pokazao funkcioniranje relevantnih tehnoloških rješenja i pružio dokaze njihove prethodne uspješne upotrebe, potvrđuje prvu hipotezu (H1) istraživanja. Možemo tvrditi da se upotrebom ulančanih zapisa produžuje dokazivost autentičnosti digitalno potpisanoga arhivskoga gradiva. Ovo je uvjetovano provjerom identiteta autora prije unosa zapisa u sustav temeljen na ulančanim zapisima što TrustChain model radi. Jednom dokazani identitet zauvijek je zapisan u nepromjenjivu podatkovnu strukturu kao dio digitalnog potpisa te skupa s njim može biti iskorišten za naknadno dokazivanje autentičnosti zapisa.

Drugi dio modela, TrustChain B modul, temeljen na kasnijem TrustChain modelu[459], omogućuje upotrebu sličnih mehanizama, prije svega mehanizma za postizanje konsenzusa sudjelujućih ustanova, da bi se dokazala ispravnost digitalnog certifikata. Nakon ovog dokaza certifikat se pohranjuje u nepromjenjivu podatkovnu strukturu na isti način kao što je u prethodnom dijelu pohranjen digitalni potpis.

Ovaj dio sustava ne doprinosi potvrdi hipoteza istraživanja. Pohrana digitalnog certifikata sama po sebi ne ostvaruje dokaz autentičnosti (preciznije rečeno, pouzdanosti) nekog zapisa iz dva razloga:

1) Certifikat u sustavu nije povezan s arhivskim zapisima. Moguće da ni ne postoje takvi arhivski zapisi.

2) Ako se naknadno pojave arhivski zapisi koji koriste certifikat koji je dio TrustChain sustava, TrustChain i dalje ne može garantirati za njihovu autentičnost. Sustav može dokazati da je certifikat bio ispravan u nekom trenutku u prošlosti i da je ispravno zabilježen njegov vlasnik, ali s obzirom da sustavu na provjeru nije dan sam zapis koji koristi (sada istekli) certifikat TrustChain ne može dokazati da je zapis pouzdan. Moguće je da je privatni ključ vezan uz certifikat u međuvremenu ukraden, ili da je uspješno izveden napad preslikom.

TrustChain model se oslanja na konsenzus više ustanova koje sve moraju provjeriti ispravnost potpisa pri ulasku zapisa u sustav. Ovo nije moguće kad je zapis povjerljive prirode. Zbog tog razloga razvijen je ovaj, dodatni model koji ne ispunjava zahtjeve arhivistike, ali pruža djelomičnu zaštitu dokumenata čije potpise nije bilo moguće pohraniti u TrustChain sustav.

[459] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving, n. dj.

Kao takav on je izuzetno praktičan dodatak izvornom modelu koji u potpunosti ispunjava zahtjeve arhivistike.

Zanimljivo je primijetiti da ovaj model funkcionira upravo suprotno od sustava za vremenske žigove. Vremenski žig može garantirati integritet zapisa, ali ne i identitet autora. TrustChain B modul može garantirati identitet vlasnika certifikata, ali ne garantira integritet zapisa koji su njime potpisani. Ustanova koja barata s velikim brojem povjerljivih zapisa može upotrijebiti kombinaciju vremenskog žiga i TrustChain B modula da bi osigurala dokazivost autentičnosti svojih zapisa. Ovo i dalje nije potpuno pouzdan postupak, jer će se u slučaju povjerljivog zapisa vremenski žig sasvim sigurno formirati na osnovu isključivo hasha zapisa, ali ipak pruža veću razinu pouzdanosti nego upotreba samo vremenskog žiga ili samo TrustChain B modula.

Konačno, detaljno je razrađena pomoćna baza podataka koja omogućava unos više metapodataka nego što je moguće s TrustChain ulančanim blokovima, omogućava naknadne izmjene metapodataka (u skladu s Općom uredbom o zaštiti podataka) i omogućava pohranu podataka o arhivskoj vezi. Ova baza podataka temeljena je na prethodnom istraživanju Stančića i Bralića.[460] Pomoćna baza podataka je nužna, ne samo iz prethodno navedenih razloga, već i zbog činjenice da ona omogućava učinkovitu pretragu podataka pohranjenih u ulančane blokove. Ovi podaci bez pomoćne baze uopće nisu indeksirani i svaka pretraga tih podataka morala bi se svesti na linearni pregled svih blokova. Osim strukture podataka koji moraju biti pohranjeni u ovoj bazi podataka istražene su i tehnologije koje su kandidati za nju te je dana preporuka upotrebe distribuiranog noSQL sustava MongoDB i, kao alternativa, Apache Cassandra sustava.

Implementacija podataka o arhivskoj vezi u pomoćnu bazu podataka znači da druga hipoteza (H2) nije dokazana. S obzirom na to da je pokazano da je arhivska veza dinamična mreža odnosa zapisa te da je ona nužna za svaki zapis, dakle nije riječ o rijetkoj pojavi poput potrebe za ispravkom određenih zapisa u ulančanim blokovima, očuvanje takvih podataka, prema mojim zaključcima, nije kompatibilno s ulančanim blokovima. Zbog ovog razloga ne možemo tvrditi da sustavi za pohranu digitalno potpisanoga arhivskoga gradiva utemeljeni na ulančanim zapisima omogućuju očuvanje arhivske veze. TrustChain sustav omogućava očuvanje arhivske veze, ali to ne čini u dijelu koji je temeljen na ulančanim zapisima. Sustav

[460] Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain, n. dj.

namjerno, zbog njihove prirode, prepušta očuvanje arhivskih veza klasičnom sustavu za pohranu podataka, poput MongoDB sustava.

Istraživanje je dalo znanstveni doprinos arhivistici kroz detaljnu analizu problema vezanih uz sve učestaliju upotrebu digitalnog potpisa kao mehanizma dokazivanja autentičnosti zapisa te je ponudilo rješenje problema u obliku novog modela za dugotrajnu pohranu digitalno potpisanoga arhivskoga gradiva, TrustChain modela. Novi model može biti uporabljen kao početna točka razvoja informacijskog sustava temeljenog na njemu. Alternativno, model se može iskoristi kao referenti model koji ukazuje na zahtjeve za dokazivanje autentičnosti digitalno potpisanih zapisa te predlaže mehanizme koje bi trebalo uključiti u postojeće standarde ili informacijske sustave da bi oni bili dovedeni na razinu koja je u skladu sa zahtjevima arhivistike.

8. Popis literature

Abramova, V., & Bernardino, J. (2013). NoSQL databases: MongoDB vs cassandra. *Proceedings of the international C* conference on computer science and software engineering*, (str. 14-22). Preuzeto 7. 1. 2022. s http://web.cs.wpi.edu/~cs585/s17/StudentsPresentations/This%20Year/Week14/mongodb_vs_cassandra.pdf

Adams, C., Cain, P., Pinkas, D., & Zuccherato, R. (2001). *RFC 3161: Internet X.509 Public Key Infrastructure Time-Stamp Protocol (TSP).* Preuzeto 6. 17. 2019. s Internet Engineering Task Force (IETF) : https://www.ietf.org/rfc/rfc3161.txt

Agencija za komercijalnu djelatnost. (2021). *Proizvodi i usluge*. Preuzeto 13. 12. 2021. s id.hr: https://www.id.hr/hr/proizvodi-i-usluge

Almgren, H. (2015). *Br. patenta WO 2015/020599 Al.* Preuzeto 1. 12. 2021. s https://patentimages.storage.googleapis.com/7a/25/e6/d6ba657b93aa99/WO2015020599A1.pdf

Amy, M., Di Matteo, O., Gheorghiu, V., Mosca, M., Parent, A., & Schanck, J. (2016). Estimating the cost of generic quantum pre-image attacks on SHA-2 and SHA-3. *International Conference on Selected Areas in Cryptography* (str. 317-337). Cham: Springer. Preuzeto 14. 5. 2020. s https://arxiv.org/pdf/1603.09383.pdf

ANSI. (2013). *ANSI/NISO Z39.85 – The Dublin Core Metadata Element Set.* Preuzeto 14. 12. 2021. s http://www.niso.org/publications/ansiniso-z3985-2012-dublin-core-metadata-element-set

ANSI. (2016). *ANSI X9.95-2016 Financial Services – Trusted Time Stamp Management And Security.* Preuzeto 19. 11. 2022. s American National Standards Institute : https://infostore.saiglobal.com/en-gb/Standards/ANSI-X9-95-2016-1894464/

Armbrust, M., Xin, R., Lian, C., Huai, Y., Liu, D., Bradley, J., . . . Zaharia, M. (2015). Relational data processing in spark. *Proceedings of the 2015 ACM SIGMOD international conference on management of data*, (str. 1383-1394). Preuzeto 27. 11. 2021. s https://dl.acm.org/doi/pdf/10.1145/2723372.2742797

Bandara, E., Liang, X., Shetty, S., Ng, W. K., Foytik, P., Ranasinghe, N., . . . Larsson, D. (2020). Lekana-Blockchain Based Archive Storage for Large-Scale Cloud Systems. *International Conference on Blockchain* (str. 169-184). Cham: Springer. Preuzeto 7. 1. 2022. s https://www.researchgate.net/publication/344372675_Lekana_-_Blockchain_Based_Archive_Storage_for_Large-Scale_Cloud_Systems

Bandara, E., Ng, W. K., Zoysa, D., K., F., N., T., S., M. P., & Jayasuriya, N. (2018). Mystiko—blockchain meets big data. *2018 IEEE International Conference on Big Data*. IEEE. Preuzeto 7. 1. 2022. s https://www.researchgate.net/publication/330632586_Mystiko-Blockchain_Meets_Big_Data

Bayer, D., Haber, S., & Stornetta, W. (1993). Improving the efficiency and reliability of digital time-stamping. *Sequences II* (str. 329-334). New York, NY: Springer. Preuzeto 1. 12. 2021. s http://www.math.columbia.edu/~bayer/papers/Timestamp_BHS93.pdf

Becker, G. (2008). *Merkle signature schemes, merkle trees and their cryptanalysis*. Bocuhm: Ruhr-Universit¨at Bochum. Preuzeto 7. 12. 2021. s https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.392.7879&rep=rep1&type=pdf

Bellare, M., & Rogaway, P. (1994). Optimal asymmetric encryption. *Workshop on the Theory and Application of of Cryptographic Techniques – EUROCRYPT 1994: Advances in Cryptology — EUROCRYPT'94* (str. 92-111). Berlin, Heidelberg: Springer. doi: https://doi.org/10.1007/BFb0053428

Benet, J. (2014). *IPFS – Content Addressed, Versioned, P2P File System.* Preuzeto 19. 11. 2022. s arxiv.org: https://arxiv.org/pdf/1407.3561.pdf

Bertoni, G., Daemen, J., Hoffert, S., Peeters, M., Van Assche, G., & Van Keer, R. (2021). *Keccak specifications summary*. Preuzeto 28. 12. 2021. s Team Keccak: https://keccak.team/keccak_specs_summary.htm l

Bertoni, G., Daemen, J., Peeters, M., & Van Assche, G. (2013). Keccak. *Annual international conference on the theory and applications of cryptographic techniques* (str. 313-314). Berlin, Heidelberg: Springer. Preuzeto 14. 5. 2020. s https://link.springer.com/content/pdf/10.1007/978-3-642-38348-9_19.pdf

Blanchette, J.-F. (2016). The digital signature dilemma. *Annales des télécommunications, 61*(7), 908-923. Preuzeto 18. 12. 2021. s https://escholarship.org/content/qt1kt3f8hx/qt1kt3f8hx.pdf

Blum, M., & Micali, S. (1984). How to generate cryptographically strong sequences of pseudorandom bits. *SIAM journal on Computing, 13*(4), 850-864. Preuzeto 1. 12. 2021. s https://dl.acm.org/doi/abs/10.1145/3335741.3335751

Boritz, J. E. (2003). *IS Practitioners' Views on Core Concepts of Information Integrity.* Preuzeto s https://web.archive.org/web/20111005085820/http://www.fdewb.unimaas.nl/marc/ecais_new/files/boritz.doc

Bormann, C. a. (12 2020). *RFC 8949: Concise binary object representation (cbor).* Preuzeto 11. 29. 2021. s IETF: https://datatracker.ietf.org/doc/html/rfc8949

Bralić, V., Kuleš, M., & Stančić, H. (2017). A model for long-term preservation of digital signature validity: TrustChain. *INFuture2017 conference proceedings*, (str. 89-113). Zagreb. doi: https://doi.org/10.17234/INFUTURE.2017.10

Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving: digital signature certification chain preservation. *Records Management Journal, 30*(3), 345-362. Preuzeto 4. 5. 2021. s https://www.emerald.com/insight/content/doi/10.1108/RMJ-08-2019-0043/full/html

Brothman, B. (1992). Isad (g): general international standard archival description. *Archivaria, 34*, 17-32. Preuzeto 7. 1. 2022. s https://archivaria.ca/index.php/archivaria/article/view/11838/12790

Brzica, H. (2018). *KONCEPT USPOSTAVE ELEKTRONIČKOGA ARHIVA U JAVNOJ UPRAVI.* Zagreb: FFZG. Preuzeto 14. 12. 2021. s http://darhiv.ffzg.unizg.hr/id/eprint/10282/1/Doktorski_rad_-_Koncept_uspostave_elektroni%C4%8Dkoga_arhiva_u_javnoj_upravi.pdf

Brzica, H., Herceg, B., & Stančić, H. (2013). Long-term Preservation of Validity of Electronically Signed Records. *INFuture2013: Information Governance* (str. 147-158). Zagreb: Filozofski Fakultet Sveučilišta u Zagrebu. Preuzeto 20. 12. 2021. s http://darhiv.ffzg.unizg.hr/id/eprint/8291/1/4-

03%20Brzica%2C%20Herceg%2C%20Stancic%2C%20LTP%20of%20Validity%20of%20Electronically%20Signed%20Records.pdf

BSON Implementations. (2021). Preuzeto 24. 11. 2021. s BSON: https://bsonspec.org/implementations.html

Buldas, A., Kroonmaa, A., & Laanoja, R. (2013). How to build global distributed hash-trees. *Nordic Conference on Secure IT Systems* (str. 313-320). Springer. Preuzeto 3. 12. 2021. s http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.434.790&rep=rep1&type=pdf

Caplan, P. (2009). Understanding Premis. Washington DC: USA: Library of Congress. Preuzeto 7. 1. 2022. s https://www.loc.gov/standards/premis/understanding-premis.pdf

Carucci, P. (1987). *Il documento contemporaneo. Diplomatica e criteri di edizione.* Rome, Italy: La Nuova Italia scientifica.

CCSDS. (2012). *REFERENCE MODEL FOR AN OPEN ARCHIVAL INFORMATION SYSTEM (OAIS).* Washington, DC, USA: Consultative Committee for Space Data Systems Secretaiat. Preuzeto 14. 12. 2021. s https://public.ccsds.org/Pubs/650x0m2.pdf

Cencetti, G. (1985). La preparazione dell'archivista. *Antologia di Scritti Archivistici*, 283-313. Preuzeto 17. 12. 2021. s http://2.42.228.123/dgagaeta/dga/uploads/documents/Saggi/543ba04bdab11.pdf

Chen, L., Jordan, S., Liu, Y.-K., Moody, D., Peralta, R., Perlner, R., & Smith-Tone, D. (2016). *NISTIR 8105 – Report on Post-Quantum Cryptography.* National Institute of Standards and Technology, U.S. Department of Commerce. National Institute of Standards and Technology. doi: http://dx.doi.org/10.6028/NIST.IR.8105

Chen, P. P.-S. (1976). The entity-relationship model—toward a unified view of data. *ACM transactions on database systems (TODS)*, 9-36. Preuzeto 14. 12. 2021. s https://dspace.mit.edu/bitstream/handle/1721.1/47638/entityrelationsh00chen.pdf?s

Chen, T., & Lin, F. (2011). Electronic medical archives: a different approach to applying re-signing mechanisms to digital signatures. *Journal of medical systems, 35*(4), 735-742. doi: https://doi.org/10.1007/s10916-009-9414-2

ChromaWay. (2017). *Postchain*. Preuzeto 9. 12. 2021. s Postchain: https://postchain-docs.readthedocs.io/en/latest/

ChromaWay. (2021). *ChromaWay Technology*. Preuzeto 9. 12. 2021. s ChromaWay : https://chromaway.com/technology

Cooper, D., Regenscheid, A., Souppaya, M., Bean, C., Boyle, M., Cooley, D., & Jenkins, M. (2018). *Security considerations for code signing.* NIST. Preuzeto 20. 12. 2021. s https://www.encryptionconsulting.com/wp-content/uploads/2020/01/NIST-code-sign-whitepaper.pdf

Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., & Polk, W. (2008). *RFC 5280: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile.* Preuzeto 6. 17. 2019. s Internet Engineering Task Force (IETF) : https://tools.ietf.org/html/rfc5280

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms.* Massachusetts Institute of Technology. Preuzeto 7. 12. 2021. s http://139.59.56.236/bitstream/123456789/106/1/Introduction%20to%20Algorithms%20by%20Thomas%20%20H%20Coremen.pdf

Cullen, C., Hirtle, P., Levy, D., & Lynch, C. R. (2000). *Authenticity in a Digital Environment.* Council on Library and Information Resources. Preuzeto 11. 12. 2021. s https://www.clir.org/wp-content/uploads/sites/6/pub92.pdf

Custódio, R., Vigil, M., Romani, J., Pereira, F., & da Silva Fraga, J. (2008). Optimized certificates–A new proposal for efficient electronic document signature validation. *European Public Key Infrastructure Workshop* (str. 49-59). Berlin, Heidelberg: Springer. Preuzeto 5. 12. 2021. s ftp://nozdr.ru/biblio/kolxoz/Cs/CsLn/P/Public%20Key%20Infrastructure,%205%20conf...%20Theory%20and%20Practice,%20EuroPKI%202008(LNCS5057,%20Springer,%202008)(ISBN%209783540694847)(247s).pdf#page=57

Dai, W. (1998). B-money proposal. Preuzeto 18. 5. 2020. s https://web.archive.org/web/20180328204908/http://www.weidai.com/bmoney.txt

Damgård, I. B. (1989). A design principle for hash functions. *Conference on the Theory and Application of Cryptology* (str. 416-427). New York, NY: Springer. Preuzeto 5. 5. 2020. s https://link.springer.com/content/pdf/10.1007/0-387-34805-0_39.pdf

Das, M., & Samdaria, N. (2014). On the security of SSL/TLS-enabled applications. *Applied Computing and informatics*, 68-81. doi: https://doi.org/10.1016/j.aci.2014.02.001

Diffie, W., & Hellman, M. (1976). New directions in cryptography. *IEEE transactions on Information Theory, 22*(6), 644-654. Preuzeto 20. 12. 2021. s https://caislab.kaist.ac.kr/lecture/2010/spring/cs548/basic/B08.pdf

Doyle, M. D. (1998). *Sjedninje Američke Države Br. patenta US6381696B1.* Preuzeto 3. 12. 2021. s https://patents.google.com/patent/US6381696B1/en

Duranti, L. (1989). Diplomatics: New Uses for an Old Science, Part I. *Archivaria, 28*, 7-27. Preuzeto 16. 12. 2021. s https://www.researchgate.net/publication/225035619_Diplomatics_New_Uses_for_an_Old_Science

Duranti, L. (1995). Reliability and Authenticity: The Concepts and Their Implications. *Archivaria, 39*, 5-10. Preuzeto 7. 1. 2022. s https://archivaria.ca/index.php/archivaria/article/view/12063/13035

Duranti, L. (1997). The archival bond. *Archives and Museum Informatics, 11*, 213-218. Preuzeto 16. 12. 2021. s https://www.researchgate.net/publication/226554280_The_Archival_Bond

Duranti, L., & Macneil, H. (1996). The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project. *Archivaria, 42*, 46-67. Preuzeto 7. 1. 2022. s https://archivaria.ca/index.php/archivaria/article/view/12153/13158

Duranti, L., Gilliland-Swetland, A., Guercio, M., Hamidzadeh, B., Iacovino, L., Lee, B., . . . Ross, S. (2001). *Authenticity Task Force Final Report.* InterPARES. Preuzeto 10. 12. 2021. s http://www.interpares.org/book/interpares_book_d_part1.pdf

Dworkin, M. J. (2015). *SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions.* doi: https://doi.org/10.6028/NIST.FIPS.202

eIDAS eID Technical Subgroup. (2019). eIDAS Cryptographic Requirements for the Interoperability Framework. Preuzeto 20. 12. 2021. s https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiP3uPQl_L0AhV5S_EDHeP1BS4QFnoECAwQAQ&url=https

%3A%2F%2Fec.europa.eu%2Fcefdigital%2Fwiki%2Fdownload%2Fattachments%2F82773108%2FeIDAS%2520Cryptographic%2520Requirement%

Emmadi, N., & Narumanchi, H. (2017). Reinforcing Immutability of Permissioned Blockchains with Keyless Signatures' Infrastructure. *Proceedings of the 18th International Conference on Distributed Computing and Networking*, (str. 1-6). Preuzeto 3. 12. 2021. s https://www.researchgate.net/profile/Harika-Narumanchi/publication/311806965_Reinforcing_Immutability_of_Permissioned_Blockchains_with_Keyless_Signatures_Infrastructure/links/5a7492f1aca2720bc0de2181/Reinforcing-Immutability-of-Permissioned-Blockchains-wi

Enigio Time AB. (2021). *Enigio*. Preuzeto 2. 12. 2021. s time:beat: https://enigio.com/timebeat/

Etingof, I. (2020). *ASN.1 modules for Python*. Preuzeto 26. 4. 2020. s GitHub: https://github.com/etingof/pyasn1-modules

European Commision. (2021). *EU Trust Services Dashboard*. Preuzeto 4. 12. 2021. s eSignature page of the European Commision: https://esignature.ec.europa.eu/efda/tl-browser/#/screen/tl/EE

European Parliament. (2014). *eIDAS.* Preuzeto 4. 12. 2021. s https://www.eid.as/home/

European Parliament and Council. (23. 7 2014). REGULATION (EU) No 910/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Preuzeto 6. 2. 2020. s https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2014.257.01.0073.01.ENG

European Parliament and Council. (2016). *Regulation (EU) 2016/679 – General Data Protection Regulation.* Preuzeto 7. 1. 2022. s https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504

European Parliament and Council. (2018). *Uredba (EU) 2016/679 Europskog parlamenta i Vijeća od 27. travnja 2016. o zaštiti pojedinaca u vezi s obradom osobnih podataka i o slobodnom kretanju takvih podataka te o stavljanju izvan snage Direktive 95/46/EZ.* Preuzeto 12. 12. 2021. s https://www.zakon.hr/z/1021/Op%C4%87a-uredba-o-za%C5%A1titi-podataka---Uredba-%28EU%29-2016-679

European Telecommunications Standards Institute. (2002). *ETSI TS 101 903 V1.1 XML Advanced Electronic Signatures (XAdES).* Preuzeto 6. 1. 2022. s ETSI: https://uri.etsi.org/01903/v1.1.1/ts_101903v010101p.pdf

European Telecommunications Standards Institute. (2009). *ETSI TS 102 778-1 V1.1 Electronic Signatures and Infrastructures (ESI); PDF Advanced Electronic Signature Profiles; Part 1: PAdES Overview – a framework document for PAdES.* Preuzeto 6. 1. 2022. s ETSI: https://www.etsi.org/deliver/etsi_ts/102700_102799/10277801/01.01.01_60/ts_10277801v010101p.pdf

European Telecommunications Standards Institute. (2013). *ETSI TS 101 733 – Electronic Signatures and Infrastructures (ESI); CMS Advanced Electronic Signatures (CAdES).* Preuzeto 19. 11. 2022. s ETSI: http://www.etsi.org/deliver/etsi_ts/101700_101799/101733/02.02.01_60/ts_101733v020201p.pdf

European Telecommunications Standards Institute. (2013). *ETSI TS 102 918 – Electronic Signatures and Infrastructures (ESI);Associated Signature Containers (ASiC.* Preuzeto 9. 5. 2020. s ETSI: https://www.etsi.org/deliver/etsi_ts/102900_102999/102918/01.03.01_60/ts_102918v010301p.pdf

European Telecommunications Standards Institute. (2013). *ETSI TS 103 172 V2 Electronic Signatures and Infrastructures (ESI); PAdES Baseline Profile.* Preuzeto 6. 1. 2022. s https://www.etsi.org/deliver/etsi_ts/103100_103199/103172/02.02.02_60/ts_103172v020202p.pdf

European Telecommunications Standards Institute. (2021). *ETSI EN 319 102-1 Electronic Signatures and Infrastructures (ESI) – Procedures for Creation and Validation of AdES Digital Signatures – Part 1: Creation and Validation.* Preuzeto 20. 11. 2022. s https://www.etsi.org/deliver/etsi_en/319100_319199/31910201/01.03.01_60/en_31910201v010301p.pdf

European Telecommunications Standards Institute. (2016). *ETSI EN 319 142-1 Electronic Signatures and Infrastructures (ESI);PAdES digital signatures Part 1: Building blocks and PAdES baseline signatures.* Preuzeto 17. 6. 2019. s ETSI:

https://www.etsi.org/deliver/etsi_en/319100_319199/31914201/01.01.01_60/en_31914201v010101p.pdf

European Telecommunications Standards Institute. (2021). *ETSI EN 319 132-1: Electronic Signatures and Infrastructures (ESI); XAdES digital signatures; Part 1: Building blocks and XAdES baseline signatures.* Preuzeto 13. 2. 2020. s ETSI: https://www.etsi.org/deliver/etsi_en/319100_319199/31913201/01.02.00_20/en_31913201v010200a.pdf

Everest, G. C. (1976). BASIC DATA STRUCTURE MODELS EXPLAINED WITH A COMMON EXAMPLE. *Computing Systems 1976, Proceedings Fifth Texas Conference on Computing Systems* (str. 39-46). Austin: IEEE Computer Society Publications Office. Preuzeto 26. 11. 2021. s https://www.researchgate.net/profile/Gordon-Everest-2/publication/291448084_BASIC_DATA_STRUCTURE_MODELS_EXPLAINED_WITH_A_COMMON_EXAMPLE/links/57affb4b08ae95f9d8f1ddc4/BASIC-DATA-STRUCTURE-MODELS-EXPLAINED-WITH-A-COMMON-EXAMPLE.pdf

Feistel, H. (1973). Cryptography and computer privacy. *Scientific american*, *228*(5), 15-23. Preuzeto 15. 5. 2020. s http://www.apprendre-en-ligne.net/crypto/bibliotheque/feistel/

Financijska Agencija. (2021). *Cijene digitalnih certifikata i vremenskih žigova*. Preuzeto 13. 12. 2021. s Financijska Agencija: https://www.fina.hr/cijene

Financijska agencija. (2021). *Osobni soft certifikat – FinaSoftCert*. Preuzeto 20. 12. 2021. s Financijska agencija: https://www.fina.hr/osobni-soft-certifikat-finasoftcert

Financijska agencija. (2021). *Rješenja za elektronički potpis*. Preuzeto 13. 12. 2021. s Financijska Agencija: https://www.fina.hr/rjesenja-za-elektronicki-potpis

FM4DD. (2020). *X509 certificate examples for testing and verification – Public Key Infrastructure and Digital Certificates*. Preuzeto 26. 4. 2020. s FM4DD.com – Information Technology Security, Programming, Software: http://fm4dd.com/openssl/certexamples.htm

Galiev, A., Prokopyev, N., Ishmukhametov, S., Stolov, E., Latypov, R., & Vlasov, I. (2018). Archain: a novel blockchain based archival system. *2018 Second World Conference on Smart Trends in Systems, Security and Sustainability* (str. 84-89). IEEE. Preuzeto 7. 1. 2022. s https://arxiv.org/ftp/arxiv/papers/1901/1901.04225.pdf

Gipp, B., Meuschke, N., & Gernandt, A. (2015). Decentralized Trusted Timestamping using the Crypto Currency. *Proceedings of the iConference 2015.* Preuzeto 3. 12. 2021. s https://www.researchgate.net/profile/Thomas-Hepp-2/publication/329249467_OriginStamp_A_blockchain-backed_system_for_decentralized_trusted_timestamping/links/5c7e3313299bf1268d395112/OriginStamp-A-blockchain-backed-system-for-decentralized-trusted-timestam

Gondrom, T., Brandner, R., & Pordesch, U. (2007). *RFC4998: Evidence Record Syntax (ERS).* Preuzeto 5. 12. 2021. s IETF Datatracker: https://datatracker.ietf.org/doc/html/rfc4998

Gränström, C., Hornfeldt, T., Peterson, G., Rinaldi Mariana, M. P., Schäfer, U., & Zwicker, J. (2002). *Authenticity of Electronic Records, a report by ICA to UNESCO.* INTERNATIONAL COUNCIL ON ARCHIVES. Preuzeto 11. 12. 2021. s https://www.ica.org/sites/default/files/ICA_Study-13-1-Authenticity-of-electronic-records-ICA-Report-to-UNESCO_EN.pdf

GuardTime AS. (2021). *GuardTime.* Preuzeto 3. 12. 2021. s KSI blockchain timestamping: https://guardtime.com/timestamping

GuardTime Federal LLC Proprietary. (n.d.). *Keyless Signature Infrastructure® (KSI™) Technology.* Preuzeto 3. 12. 2021. s GuardTime Library – Whitepapers: http://blockchain.machetemag.com/wp-content/uploads/2017/11/Guardtime_WhitePaper_KSI.pdf

Haber, S., & Kamat, P. (2006). A content integrity service for long-term digital archives. *Archiving Conference, 2006,* str. 159-164. Preuzeto 5. 12. 2021. s http://www.hpl.hp.com/techreports/2006/HPL-2006-54.pdf

Haber, S., & Stornetta, W. S. (January 1990). How to time-stamp a digital document. *Journal of Cryptology, 3*(2), 99-111. doi: https://doi.org/10.1007/BF00196791

Han, J., Haihong, E., Le, G., & Du, J. (2011). Survey on NoSQL database. *2011 6th international conference on pervasive computing and applications* (str. 363-366). IEEE. Preuzeto 26. 11. 2021. s http://faculty.washington.edu/wlloyd/courses/tcss562/papers/Spring2017/team7_NOSQL_DB/Survey%20on%20NoSQL%20Database.pdf

Haughian, G., Osman, R., & Knottenbelt, W. (2016). Benchmarking replication in cassandra and mongodb nosql datastores. *International Conference on Database and Expert Systems Applications* (str. 152-166). Cham: Springer. doi: https://doi.org/0.1007/978-3-319-44406-2_12

Helmer, S., Roggia, M., El Ioini, N., & Pahl, C. (2018). Ethernitydb–integrating database functionality into a blockchain. *European Conference on Advances in Databases and Information Systems* (str. 37-44). Cham: Springer. Preuzeto 7. 1. 2022. s https://www.researchgate.net/publication/327309357_EthernityDB_-_Integrating_Database_Functionality_into_a_Blockchain

Hepp, T., Schoenhals, A., Gondek, C., & Gipp, B. (2018). OriginStamp: A blockchain-backed system for decentralized trusted timestamping. *it-Information Technology*, *60*(5-6), 273-281. doi: https://doi.org/10.1515/itit-2018-0020

Hepp, T., Wortner, P., Schönhals, A., & Gipp, B. (2018). Securing physical assets on the blockchain: Linking a novel object identification concept with distributed ledgers. *Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems* (str. 60-65). New York, USA: ACM Pres. doi: https://doi.org/10.1145/3211933.3211944

Hofman, D. L. (2017). Legally speaking: Smart contracts, archival bonds, and linked data in the blockchain. *26th International Conference on Computer Communication and Networks (ICCCN)* (str. 1-4). IEEE. Preuzeto 16. 12. 2021. s https://blockhack.osive.com/_downloads/d5a5c8e62a06c24e5791ab043a950796/61.pdf

Housley, R. (2009). *RFC 5626: Cryptographic Message Syntax (CMS).* Preuzeto 2. 5. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc5652

Hrvatski državni arhiv. (2021). Preuzeto 12. 12. 2021. s Nacionalni Arhivski Informacijski Sustav: http://arhinet.arhiv.hr/default.aspx

Hrvatski Sabor. (2005). *Zakon o Elektroničkoj Ispravi.* Preuzeto 13. 12. 2021. s https://www.zakon.hr/z/272/Zakon-o-elektroni%C4%8Dkoj-ispravi

Hrvatski Sabor. (2017). *Odluka o proglašenju zakona o provedbi uredbe (EU) br. 910/2014 Europskog Parlamenta i Vijeća od 23. srpnja 2014. o elektroničkoj identifikaciji i uslugama povjerenja za elektroničke transakcije na unutarnjem tržištu i stavljanju*

izvan snage direktive 199. Preuzeto 12. 12. 2021. s https://narodne-novine.nn.hr/clanci/sluzbeni/2017_06_62_1430.html

Hrvatski Sabor. (2018). *Odluka o proglašenju Zakona o provedbi opće uredbe o zaštiti podatka.* Preuzeto 13. 12. 2021. s https://narodne-novine.nn.hr/clanci/sluzbeni/2018_05_42_805.html

Hrvatski sabor. (2018). *Zakon o arhivskom gradivu i arhivima.* Preuzeto 19. 11. 2022. s https://www.zakon.hr/z/373/Zakon-o-arhivskom-gradivu-i-arhivima

Hrvatsko arhivsko vijeće. (2012). *Opći popis gradiva s rokovima čuvanja.* Preuzeto 12. 12. 2021. s ArhiNET: http://arhinet.arhiv.hr/_Download/PDF/Opci_popis_gradiva_s_rokovima_cuvanja.pdf

Hyla, T., W. Bielecki, W., & Pejaś, J. (2010). Non-repudiation of Electronic Health Records in distributed healthcare systems. *Pomiary Automatyka Kontrola, 56*(10), 1170-1173. Preuzeto 5. 12. 2021. s https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-article-BSW4-0086-0015/c/Hyla.pdf

IBM. (2007). *The History of Notes and Domino.* IBM. Preuzeto 20. 12. 2021. s https://www.notesmail.com/home.nsf/ls-NDHistory-pdf.pdf

International Organization for Standardization. (2008). *ISO 32000-1:2008 Document management — Portable document format — Part 1: PDF 1.7.* Preuzeto 1. 5. 2020. s ISO – International Organization for Standardization: https://www.iso.org/standard/51502.html

International Organization for Standardization. (2009). Preuzeto 19. 11. 2022. s ISO/IEC 18014-3:2009 Information technology -- Security techniques -- Time-stamping services -- Part 3: Mechanisms producing linked tokens: https://www.iso.org/standard/50457.html

International Organization for Standardization. (2017). *ISO 15836-1:2017 Information and documentation — The Dublin Core metadata element set — Part 1: Core elements.* ISO. Preuzeto 14. 12. 2021. s https://www.iso.org/standard/71339.html

International Organization for Standardization. (2018). *ISO 14721:2012 Space data and information transfer systems — Open archival information ystem (OAIS) — Reference*

model. International Organization for Standardization. Preuzeto 14. 12. 2021. s https://www.iso.org/standard/57284.html

Internet Engineering Task Force. (2017). *RFC8259: The JavaScript Object Notation (JSON) Data Interchange Format*. (T. Bary, Urednik) Preuzeto 18. 11. 2021. s Internet Engineering Task Force (IETF): https://datatracker.ietf.org/doc/html/rfc8259

InterPARES Project. (2001). *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project.* Preuzeto 7. 1. 2022. s http://www.interpares.org/book/

InterPARES Project. (2016). *InterPARES 2 Project Glossary.* Preuzeto s http://www.interpares.org/ip2/ip2_term_pdf.cfm?pdf=glossary

InterPARES Trust Project. (2018). *TRUSTER Preservation Model (EU31) – Case Study 1 – digitally signed retirement fund records.* Preuzeto 1. 7. 2022. s https://interparestrust.org/assets/public/dissemination/TRUSTERPreservationModel(EU31)-CaseStudy1v1_2.pd f

InterPARES Trust Project. (2018). *TRUSTER Preservation Model (EU31) – Case Study 2 – digitally signed e-tax records.* Preuzeto 7. 1. 2022. s https://interparestrust.org/assets/public/dissemination/TRUSTERPreservationModel(EU31)-CaseStudy2v1_2.pdf

InterPARES Trust Project. (2018). *TRUSTER Preservation Model (EU31) – Case Study 3 – digitally signed medical records, procurement and supplier contracts, official political political decisions and minutes of meetings.* Preuzeto 7. 1. 2022. s https://interparestrust.org/assets/public/dissemination/TRUSTERPreservationModel(EU31)-CaseStudy3v1_3.pdf

Ismailisufi, A., Popović, T., Gligorić, N., Radonjic, S., & Šandi, S. (2020). A private blockchain implementation using multichain open source platform. *2020 24th International Conference on Information Technology (IT)* (str. 1-4). Žabljak: IEEE. Preuzeto 29. 11. 2021. s https://www.researchgate.net/profile/Nenad-Gligoric/publication/339442490_A_Private_Blockchain_Implementation_Using_Multichain_Open_Source_Platform/links/5e73dd6d458515c677c6213d/A-Private-Blockchain-Implementation-Using-Multichain-Open-Source-Platform.pd

ITU. (2022). *X.680 : Information technology – Abstract Syntax Notation One (ASN.1): Specification of basic notation.* Preuzeto 5. 1. 2022. s International Communication Union: https://www.itu.int/rec/T-REC-X.680

Josefsson, S., & Leonard, S. (2015). *RFC 7468: Textual Encodings of PKIX, PKCS, and CMS Structures.* Preuzeto 25. 4. 2020. s Internet Engineering Task Force (IETF): https://tools.ietf.org/html/rfc7468

Kaliski, B. (1998). *RFC2315 – PKCS #7 – Cryptographic Message Syntax.* Preuzeto 2. 5. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc2315

Kaliski, B., & Staddon, J. (1998). *RFC 2437: PKCS #1: RSA Cryptography Specifications.* Preuzeto 26. 4. 2020. s Internet Engineering Task Force (IETF): https://tools.ietf.org/html/rfc2437

Kayser, R. F. (2007). *Announcing request for candidate algorithm nominations for a new cryptographic hash algorithm (SHA-3) family*. Preuzeto 14. 5. 2020. s Federal Register: https://www.federalregister.gov/documents/2007/11/02/E7-21581/announcing-request-for-candidate-algorithm-nominations-for-a-new-cryptographic-hash-algorithm-sha-3

Kondylakis, H., Fountouris, A., & Plexousakis, D. (2016). Efficient Implementation of Joins over Cassandra DBs. *EDBT—International Conference on Extending Database Technology*, (str. 666-667). doi: http://dx.doi.org/10.5441/002/edbt.2016.77

Krawczyk, H., Bellare, M., & Canetti, R. (Februray 1997). *RFC 2014 – HMAC: Keyed-Hashing for Message Authentication.* Preuzeto 9. 5. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc2104

Lakshman, A., & Malik, P. (2010). Cassandra: a decentralized structured storage system. *ACM SIGOPS Operating Systems Review*, *44*(2), 35-40. Preuzeto 7. 1. 2022. s https://www.cs.cornell.edu/projects/ladis2009/papers/lakshman-ladis2009.pdf

Lamport, L. (1981). Password authentication with insecure communication. *Communications of the ACM*, *24*(11), 770-772. Preuzeto 7. 12. 202.1 s http://merlot.usc.edu/cs530-s07/papers/Lamport81a.pdf

Lawson, N. (2009). *Why RSA encryption padding is critical*. Preuzeto 23. 4. 2020. s rdist: https://rdist.root.org/2009/10/06/why-rsa-encryption-padding-is-critical/

Leitold, H., & Konrad, D. (2019). Qualified Remote Signatures–Solutions, its Certification, and Use. *Proceedings of 29th SmartCard Workshop*, (str. 219-231). Preuzeto 29. 4. 2020. s https://pdfs.semanticscholar.org/3908/de8d4adb1dcbd88785e64aae37295147c58e.pdf

Lemieux, V. L. (2016). Trusting records: is Blockchain technology the answer? *Records Management Journal, Vol. 26 Issue: 2*, 110-139. Preuzeto 7. 1. 2022. s https://www.emerald.com/insight/content/doi/10.1108/RMJ-12-2015-0042/full/html

Lemieux, V. L. (2017). A typology of blockchain recordkeeping solutions and some reflections on their implications for the future of archival preservation. *2017 IEEE International Conference on Big Data (Big Data)* (str. 2271-2278). IEEE. Preuzeto 3. 12. 2021. s https://www.researchgate.net/profile/Victoria-Lemieux/publication/322511343_A_typology_of_blockchain_recordkeeping_solutions_and_some_reflections_on_their_implications_for_the_future_of_archival_preservation/links/5a9626ed45851535bcdcc1f4/A-typology-of-bl

Lemieux, V. L. (2017). Blockchain and distributed ledgers as trusted recordkeeping systems. *Future Technologies Conference (FTC).* Preuzeto 4. 12. 2021. s https://www.researchgate.net/profile/Victoria-Lemieux/publication/317433591_Blockchain_and_Distributed_Ledgers_as_Trusted_Recordkeeping_Systems_An_Archival_Theoretic_Evaluation_Framework/links/593aa6450f7e9b3317f4d860/Blockchain-and-Distributed-Ledgers-as

Lemieux, V. L. (2021). Blockchain and Recordkeeping: Editorial. *Computers, 10*, 135. doi: https://doi.org/10.3390/computers10110135

Lemieux, V. L., & Sporny, M. (2017). Preserving the Archival Bond in Distributed Ledgers: A Data Model and Syntax. *WWW '17 Companion Proceedings of the 26th International Conference on World Wide Web Companion*, (str. 1437-1443). Perth, Australia. Preuzeto 16. 12. 2021. s http://papers.www2017.com.au.s3.amazonaws.com/companion/p1437.pdf

Maaten, E. (2004). Towards remote e-voting: Estonian case. *Electronic voting in Europe-Technology, law, politics and society, workshop of the ESF TED programme together with GI and OCG*. Gesellschaft für Informatik eV. Preuzeto 4. 12. 2021. s

https://dl.gi.de/bitstream/handle/20.500.12116/29132/Proceeding.GI.47-9.pdf?sequence=1

Maetouq, A., Daud, S., Ahmad, N., Maarop, N., Sjarif, N., & Abas, H. (2018). Comparison of hash function algorithms against attacks: A review. *International Journal of Advanced Computer Science and Applications, 9*(8). Preuzeto 21. 12. 2021. s https://pdfs.semanticscholar.org/6ed2/50d11a5c80f550bd8efcc673606c3cae34b7.pdf

Maniatis, P., & Baker, M. (2002). Enabling the Archival Storage of Signed Documents. *FAST 2002 Proceedings*, (str. 31-46). Preuzeto 5. 12. 2021. s http://shiftleft.com/mirrors/www.hpl.hp.com/personal/Mary_Baker/publications/fast2002.pdf

Maniatis, P., Roussopoulos, M., Giuli, T., Rosenthal, D., & Baker, M. (2005). The LOCKSS peer-to-peer digital preservation system. *ACM Transactions on Computer Systems (TOCS), 23*(1), 2-50. Preuzeto 5. 12. 2021. s https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.60.2234&rep=rep1&type=pdf

McConaghy, T., Marques, R., Müller, A., Jonghe, D. D., McConaghy, T., McMullen, G., . . . Granzotto, A. (8. 6. 2016.). *Bigchaindb: a scalable blockchain database.* Preuzeto 19. 11. 2022. s BigChainDB: https://gamma.bigchaindb.com/whitepaper/bigchaindb-whitepaper.pdf

Mendel, F., Nad, T., & Schläffer, M. (2011). Finding SHA-2 characteristics: searching through a minefield of contradictions. *International Conference on the Theory and Application of Cryptology and Information Security* (str. 288-307). Berlin, Heidelberg: Springer. Preuzeto 5. 5. 2020. s https://link.springer.com/content/pdf/10.1007/978-3-642-25385-0_16.pdf

Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1996). *Handbook of Applied Cryptography.* CRC Press. Preuzeto 20. 12. 2021. s http://labit501.upct.es/~fburrull/docencia/SeguridadEnRedes/old/teoria/bibliography/HandbookOfAppliedCryptography_AMenezes.pdf

Merkle, R. C. (1979). *Secrecy, authentication, and public key systems.* Stanford University. Preuzeto 5. 5. 2020. s http://www.merkle.com/papers/Thesis1979.pdf

Merkle, R. C. (1980). Protocols for public key cryptosystems. *IEEE Symposium on Security and Privacy* (str. 122-134). IEEE. Preuzeto 6. 7. 2019. s https://www.researchgate.net/profile/Ralph_Merkle/publication/220713913_Protocols_for_Public_Key_Cryptosystems/links/00b495384ecda07784000000/Protocols-for-Public-Key-Cryptosystems.pdf

Merkle, R. C. (1982). *Washington, DC Br. patenta U.S. Patent No. 4,309,569*. Preuzeto 8. 12. 2021. s https://worldwide.espacenet.com/patent/search/family/022107098/publication/US4309569A?q=pn%3DUS4309569

Merkle, R. C. (1987). A digital signature based on a conventional encryption function. *Conference on the theory and application of cryptographic techniques* (str. 369-378). Berlin, Heidelberg: Springer. Preuzeto 1. 12. 2021. s https://link.springer.com/content/pdf/10.1007/3-540-48184-2_32.pdf

Metsallik, J., Ross, P., Draheim, D., & Piho, G. (2018). Ten years of the e-health system in Estonia. *CEUR Workshop Proceedings*, (str. 6-15). Preuzeto 4. 12. 2021. s http://ceur-ws.org/Vol-2336/MMHS2018_invited.pdf

MongoDB Inc. (2020). *JSON and BSON*. Preuzeto 19. 11. 2022. s MongoDB: https://www.mongodb.com/json-and-bson

MongoDB Inc. (2020). *MongoDB Documentation*. Preuzeto 19. 11. 2022. s MongoDB: https://docs.mongodb.com/

MongoDB Inc. (2021). *GridFS*. Preuzeto 30. 11. 2021. s MongoDb documentation: https://docs.mongodb.com/manual/core/gridfs/

Moriarty, K., B., K., Jonsson, J., & R Rusch, A. (2016). *RFC 8017 – PKCS #1: RSA Cryptography Specifications Version 2.2*. Preuzeto 23. 4. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc8017

Muzammal, M., Qu, Q., & Nasrulin, B. (2019). Renovating blockchain with distributed databases: An open source system. *Future generation computer systems*, 105-117. doi: https://doi.org/10.1016/j.future.2018.07.042

Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system.* Preuzeto 7. 6. 2019. s https://bitcoin.org/bitcoin.pdf

National Institute of Standards and Technology. (1995). *NIST FIPS 180-1 Secure Hash Standard.* doi: https://doi.org/10.6028/NIST.FIPS.180-1

National Institute of Standards and Technology. (2002). *NIST FIPS 180-2 Secure Hash Standard (SHS).* Preuzeto 15. 5. 2020. s National Institute of Standards and Technology: https://csrc.nist.gov/publications/detail/fips/180/2/archive/2002-08-01

National Institute of Standards and Technology. (2008). *NIST FIPS 180-3 Secure Hash Standard (SHS).* Preuzeto 15. 5. 2020. s National Institute of Standards and Technology: https://csrc.nist.gov/publications/detail/fips/180/3/archive/2008-10-31

National Institute of Standards and Technology. (2015). *NIST FIPS 180-4 Secure Hash Standard (SHS).* doi: https://doi.org/10.6028/NIST.FIPS.180-4

National Institute of Standards and Technology. (2019). *Transitioning the Use of Cryptographic Algorithms and Key Lengths.* Preuzeto 27. 12. 2021. s https://csrc.nist.gov/publications/detail/sp/800-131a/rev-2/final

Naz, M., Al-zahrani, F. A., Khalid, R., Javaid, N., Qamar, A. M., Afzal, M. K., & Shafiq, M. (2019). A secure data sharing platform using blockchain and interplanetary file system. *Sustainability, 11*(24), 7054. Preuzeto 7. 1. 2022. s https://www.mdpi.com/2071-1050/11/24/7054

NIST. (2021). *array*. Preuzeto 20. 12. 2021. s Dictionary of Algorithms and Data Structures: https://xlinux.nist.gov/dads/HTML/array.html

OASIS Open. (2015). *PKCS #11 Cryptographic Token Interface Base Specification Version 2.4.* Preuzeto 28. 4. 2020. s OASIS | Advancing open standards for th einformation society: http://docs.oasis-open.org/pkcs11/pkcs11-base/v2.40/os/pkcs11-base-v2.40-os.html

OriginStamp AG. (2021). *Blockchain Timestamps for Businesses*. Preuzeto 4. 12. 2021. s OriginStamp: https://originstamp.com/

Paar, C., & Pelzl, J. (2009). *Understanding cryptography: a textbook for students and practitioners.* Springer Science & Business Media. doi: https://doi.org/10.1007/978-3-642-04101-3

Paar, C., & Pelzl, J. (2010). SHA-3 and The Hash Function Keccak. U C. Paar, & J. Pelzl, *Understanding Cryptography-A Textbook for Students and Practitioners.* Springer.

Penard, W., & van Werkhoven, T. (2008). On the secure hash algorithm family. *Cryptography in Context*, (str. 1-18). Preuzeto 5. 5. 2020. s https://web.archive.org/web/20160330153520/http://www.staff.science.uu.nl/~werkh108/docs/study/Y5_07_08/infocry/project/Cryp08.pdf

Permatasari, I., Essaid, M., Kim, H., & Ju, H. (2020). Blockchain Implementation to Verify Archives Integrity on Cilegon E-Archive. *Applied Sciences, 10*(7). Preuzeto 7. 1. 2022. s https://www.mdpi.com/2076-3417/10/7/2621

Pharow, P., & Blöbel, B. (2005). Electronic signatures for long-lasting storage purposes in electronic archives. *International Journal of Medical Informatics, 74*(2-4), 279-287. Preuzeto 5. 12. 2021. s https://ebooks.iospress.nl/pdf/doi/10.3233/978-1-60750-939-4-316

Pinkas, D., Pope, N., & Ross, J. (February 2008). *RFC 5126: CMS Advanced Electronic Signatures (CAdES).* Preuzeto 2. 5. 2020. s Internet Engineering Task Force: https://tools.ietf.org/html/rfc5126

PREMIS Editorial Committee. (2015). *Premis data dictionary for preservation metadata.* Dohvaćeno s USA: Library of Congress: https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf

Proofspace. (2007). *Proofmark System Technical Overview.* ProofSpace. Preuzeto 3. 12. 2021. s http://fios.com/proofmarksystemtech.pdf

Rahim, R., Nurarif, S., Ramadhan, M., Aisyah, S., & Purba, W. (2017). Comparison searching process of linear, binary and interpolation algorithm. *Journal of Physics: Conference Series, 930*(1), 012007. doi: https://doi.org/10.1088/1742-6596/930/1/012007

Ramesh, D., Sinha, A., & Singh, S. (2016). Data modelling for discrete time series data using Cassandra and MongoDB. *2016 3rd international conference on recent advances in information technology (RAIT)* (str. 598-601). Dhanbad, India : IEEE. doi: https://doi.org/10.1109/RAIT.2016.7507966

Rivest, R., Shamir, A., & Adleman, L. (1977). *On Digital Signatures and Public-Key Cryptosystems.* MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTER SCIENCE. Preuzeto 4. 1. 2022. s https://apps.dtic.mil/sti/pdfs/ADA039036.pdf

Rogers, C. (2016). A literature review of authenticity of records in digital systems: from ‘machine-readable’to records in the cloud. *Acervo, 29*(2), 16-44. Preuzeto 16. 12. 2021. s https://www.researchgate.net/publication/320593846_A_Literature_Review_of_Authenticity_of_Records_in_Digital_Systems_From_%27Machine-Readable%27_to_Records_in_the_Cloud

Rudolph, K. (2011). Separated at appraisal: Maintaining the archival bond between archives collections and museum objects. *Archival Issues, 33*(1), 25-40. Preuzeto 16. 12. 2021. s https://minds.wisconsin.edu/bitstream/handle/1793/72333/AI_Vol33_No1_KatieRudolph1.pdf?sequence=1

Santesson, S., Myers, M., Ankney, R., Malpani, A., Galperin, S., & C., A. (2013). *RFC 6960: X.509 Internet Public Key Infrastructure Online Certificate Status Protocol – OCSP*. Preuzeto 11. 19. 2021. s Internet Engineering Task Force (IETF): https://tools.ietf.org/html/rfc6960

Schwalm, S. (2017). A service for the preservation of evidence and data–a key for a trustworthy & sustainable electronic business. *Open Identity Summit 2017.* Bonn: Gesellschaft für Informatik. Preuzeto 5. 12. 2021. s https://dl.gi.de/bitstream/handle/20.500.12116/3571/proceedings-10.pdf?sequence=1&isAllowed=y

Snow, P., Deery, B., Kirby, P., & Johnston, D. (2015). *Factom ledger by consensus.* Preuzeto 9. 12. 2021. s https://cryptochainuni.com/wp-content/uploads/Factom-Ledger-by-Consensus.pdf

Society of American Archivists. (2021). *Describing Archives: A Content Standard (DACS)*. Preuzeto 19. 11. 2022. s Society of American Archivists: https://www2.archivists.org/groups/technical-subcommittee-on-describing-archives-a-content-standard-dacs/describing-archives-a-content-standard-dacs-second-

Song, S., & JaJa, J. (2009). Techniques to audit and certify the long-term integrity of digital archives. *International Journal on Digital Libraries, 10*(2-3), 121-131. Preuzeto 5. 12. 2021. s https://drum.lib.umd.edu/bitstream/handle/1903/7130/ACE-techniques-UMIACS-TR-2007-38.pdf?sequence=1&isAllowed=y

Stančić, H. (2006). Teorijski model postojanog očuvanja autentičnosti elektroničkih informacijskih objekata. Zagreb: Filozofski fakultet. Preuzeto 7. 1. 2022. s https://www.bib.irb.hr/244465

Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain: Overcoming the Limitations of Data Immutability. *Computers, 10*(8). doi: https://doi.org/10.3390/computers10080091

Stančić, H., Rajh, A., & Brzica, H. (2015). Archival Cloud Services: Portability, Continuity, and Sustainability Aspects of Long-term Preservation of Electronically Signed Records. *Canadian Journal of Information and Library Science, 39*(2), 210-227. Preuzeto 5. 12. 2021. s https://muse.jhu.edu/article/590941

Stevens, M. (2013). New collision attacks on SHA-1 based on optimal joint local-collision analysis. *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (str. 245-261). Berlin, Heidelberg: Springer. Preuzeto 27. 12. 2021. s https://link.springer.com/content/pdf/10.1007/978-3-642-38348-9_15.pdf

Stevens, M., Bursztein, E., Karpman, P., Albertini, A. M., Bianco, A. P., & Baisse, C. (2017). *Announcing the first SHA1 collision*. Preuzeto 27. 12. 2022. s Google Security Blog: https://security.googleblog.com/2017/02/announcing-first-sha1-collision.html

Surety, I. (2021). *AbsoluteProof*. Preuzeto 2. 12. 2021. s Surety: http://www.surety.com/

The National Archives. (2021). *Our digital cataloguing practices*. Preuzeto 14. 12. 2021. s The National Archives: https://www.nationalarchives.gov.uk/about/our-role/plans-policies-performance-and-projects/our-plans/our-digital-cataloguing-practices/

Theuermann, K., Tauber, A., & Lenz, T. (2019). Mobile-only solution for server-based qualified electronic signatures. *CC 2019-2019 IEEE International Conference on Communications (ICC)* (str. 1-7). IEEE. Preuzeto 5. 1. 2022. s https://www.researchgate.net/profile/Kevin-Theuermann-3/publication/334485288_Mobile-Only_Solution_for_Server-Based_Qualified_Electronic_Signatures/links/5d95c72c92851c2f70e62ea4/Mobile-Only-Solution-for-Server-Based-Qualified-Electronic-Signatures.pdf

Thimblin, M., Kamisetty, N., Raman, P., & Paila, A. (2005). *Implementation of an Evidentiary Record Validation Utility and Security Analysis for Surety's AbsoluteProof*. George Mason University. Preuzeto 3. 12. 2005. s

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.136.7029&rep=rep1&type=pdf

UNESCO. (2021). *Organization of the Archives*. Preuzeto 14. 12. 2021. s UNESDOC Digital Library: https://unesdoc.unesco.org/archives/organization-of-the-archives

Universal Binary JSON Specification. (2021). *Universal Binary JSON Specification*. (U. B. Specification, Producent) Preuzeto 11. 29. 2021. s https://ubjson.org/

van Diessen, R. J., & van der Werf-Davelaar, T. (2002). *Authenticity in a Digital Environment, IBM / Koninklijke Bibliotheek Long-Term Preservation Study Report Series.* Amsterdam: IBM Netherlands.

Vatsa, V. R., & Chhaparwal, P. (2021, July). Estonia's e-governance and digital public service delivery solutions. In *2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT)* (pp. 135-138). IEEE. Preuzeto 4. 12. 2021. iz https://ieeexplore.ieee.org/abstract/document/9515004

Vigil, M., Cabarcas, D., Wiesmaier, A., & Buchmann, J. (2011). Authenticity, integrity and proof of existence for long-term archiving: a survey. *Cryptology EPrint Archive*. Preuzeto 4. 12. 2021. s https://d1wqtxts1xzle7.cloudfront.net/39270547/54920f5b0cf2484a3f3e092f-with-cover-page-v2.pdf?Expires=1638641185&Signature=Rkf~VTGJzZrTsQ~y3PXakhTN~TP6TwPbjzcJ4tvujltSnQQAy~EjNYoSuj1JqGwjYmdr0-z0P8o3TYeGGm7t9IL2ETrp~QRPmyF1msiTqLW5nz7Yhfd-EsDm3pjtEdf-njX

Volarević, I., & Stančić, H. (2016). Norme za elektroničke vremenske žigove i mogućnosti njihove primjene u arhivskoj struci. (S. Babić, Ur.) *Arhivi i domovinski rat*, 425-435. Preuzeto 18. 12. 2021. s https://www.researchgate.net/profile/Hrvoje-Stancic/publication/341287614_Norme_za_elektronicke_vremenske_zigove_i_mogucnosti_njihove_primjene_u_arhivskoj_struci_Standards_for_electronic_time_stamps_and_the_possibilities_for_their_application_in_archival_

Wang, X., Yu, H., & Yin, Y. (2005). Efficient collision search attacks on SHA-0. *Annual International Cryptology Conference* (str. 1-16). Berlin: Springer. Preuzeto 3. 1. 2022. s https://link.springer.com/content/pdf/10.1007/11535218_1.pdf

Webster, A. F., & Tavares, S. E. (1985). On the design of S-boxes. *Conference on the theory and application of cryptographic techniques* (str. 523-534). Berlin, Heidelberg: Springer. Preuzeto 15. 5. 2020. s https://link.springer.com/content/pdf/10.1007/3-540-39799-X_41.pdf

Wegner, P., & Reilly, E. D. (2003). Data structures. U A. Ralston, E. D. Reilly, & D. Hemmendinger (Ur.). John Wiley and Sons Ltd. Preuzeto 7. 1. 2022. s https://dl.acm.org/doi/pdf/10.5555/1074100.1074312

Weibel, S., & Baker, T. (2007). *RFC5013: Dublin Core Metadata for Resource Discovery.* Preuzeto 14. 12. 2021. s Internet Engineering Task Force: https://www.ietf.org/rfc/rfc5013.txt

Wood, G. (2014). *Ethereum: A secure decentralised generalised transaction ledger.* Preuzeto 23. 11. 2021. s Ethereum project yellow paper: https://files.gitter.im/ethereum/yellowpaper/VIyt/Paper.pdf

World Wide Web Consortium. (2008). *Extensible Markup Language (XML) 1.0 (Fifth Edition).* Preuzeto 2. 5 .2020. s World Wide Web Consortium (W3C): https://www.w3.org/TR/2008/REC-xml-20081126/

World Wide Web Consortium. (2013). *XML Signature Syntax and Processing Version 1.1.* Preuzeto 2. 5. 2020. s World Wide Web Consortium (W3C): https://www.w3.org/TR/xmldsig-core/

Yang, R., & LI, Y. (2020). *The Final Report for A Blockchain Explorer.* Hong Kong: Hong Kong University of Science and Technology. Preuzeto 22. 11. 2021. s https://file-1252789527.cos.ap-shenzhen-fsi.myqcloud.com/MSBD6000D-Blockchain/MSBD6000D%20Group%2016%20report.pdf

Yao, A. C. (1982). Theory and application of trapdoor functions. *3rd Annual Symposium on Foundations of Computer Science (SFCS 1982)* (str. 80-91). IEEE. doi: https://doi.org/10.1109/SFCS.1982.45

Yordzhev, K. Y. (2009). An example for the use of bitwise operations in programming. *Proceedings of the Thirty Eighth Spring Conference of the Union of Bulgarian Mathematicians*, (str. 196-202). Borovetz. Preuzeto 21. 12. 2021. s https://www.researchgate.net/publication/51978936_An_Example_for_the_Use_of_Bitwise_Operations_in_Programming

Zenner, E. (2009). Nonce generators and the nonce reset problem. *International Conference on Information Security* (str. 411-426). Berlin, Heidelberg: Springer. Preuzeto 3. 12. 2021. s https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1061.3009&rep=rep1&type=pdf

Zhong, Z., Wei, S., Xu, Y., Zhao, Y., Zhou, F., Luo, F., & Shi, R. (2020). SilkViser: A Visual Explorer of Blockchain-based Cryptocurrency Transaction Data. *2020 IEEE Conference on Visual Analytics Science and Technology (VAST)* (str. 95-106). IEEE. Preuzeto 22. 11. 2021. s https://arxiv.org/pdf/2009.02651.pdf

POPIS SLIKA

POPIS TABLICA

POPIS PROGRAMSKOG KODA

ŽIVOTOPIS AUTORA

Vladimir Bralić je rođen 1980. u Rijeci. Formalno obrazovanje započeo je u Osnovnoj školi Nikola Tesla u Rijeci i srednjoj školi, Prvoj sušačkoj hrvatskoj gimnaziji u Rijeci.

Završni rad obranio je 2013. godine na Veleučilištu Velika Gorica u Velikoj Gorici, na stručnom preddiplomskom studiju Održavanje računalnih sustava na temu "Anonymous napadi na internet infrastrukturu". Preddiplomski studij završio je s težinskim prosjekom ocjena 5,0 te mu je dodijeljena pohvalnica Veleučilišta Velika Gorica.

Diplomirao je u lipnju 2016. godine na sveučilišnom diplomskom studiju Informacijske znanosti, smjer Istraživačka informatika, na Filozofskom fakultetu Sveučilišta u Zagrebu, na temu "Lokacijski servisi na Android uređajima" s ukupnim težinskim prosjekom ocjena 4,805.

Poslijediplomski doktorski studij informacijskih i komunikacijskih znanosti na Filozofskom fakultetu Sveučilišta u Zagrebu upisuje 2017. godine.

Po završetku stručnog studija, od 2013. do 2016. godine radi na Veleučilištu Velika Gorica kao stručni suradnik u Odjelu za nastavnu djelatnost.

Od 2016. do 2019. radi kao asistent u nastavi na Veleučilištu Velika Gorica na predmetima Programiranje te Algoritmi i strukture podataka.

Od 2019. radi kao predavač na predmetima Programiranje te Algoritmi i strukture podataka. Od 2020. nositelj je predmeta Programiranje na Veleučilištu Velika Gorica.

Prije studija i zapošljavanja na Veleučilištu Velika Gorica (od 2006. do 2012.) pristupnik je zaposlen u Lungomare d.o.o. kao sistemski administrator te u Ting d.o.o. i Metornet d.o.o. kao mrežni tehničar.

Od 2016. do 2019. uključen je u međunarodni projekt InterPARES Trust. Pristupnik sudjeluje u TRUSTER radnoj skupini koja se bavi evaluacijom mogućnosti dugoročnog očuvanja digitalnih potpisa. Tijekom rada na projektu sudjelovao je u izradi tri studije slučajeva, stručnog rječnika na engleskom jeziku iz područja digitalnog potpisa i ulančanih zapisa te je započeo razvoj novog modela informacijskog sustava za očuvanje autentičnosti digitalno potpisanoga arhivskoga. Iz ovoga rada proizašla je i tema doktorske disertacije te suradnja s mentorom na njoj.

POPIS OBJAVLJENIH RADOVA

[1] Stančić, Hrvoje; Bralić, Vladimir. Digital Archives Relying on Blockchain: Overcoming the Limitations of Data Immutability // Computers, 10 (2021), 8; 91, 16 doi:10.3390/computers10080091.

[2] Bralić, Vladimir; Stančić, Hrvoje; Stengård, Mats. A blockchain approach to digital archiving: digital signature certification chain preservation // Records Management Journal, 30 (2020), 3; 345-362 doi:10.1108/RMJ-08-2019-0043.

[3] Bralić, Vladimir; Kuleš, Magdalena; Stančić, Hrvoje: A model for long-term preservation of digital signature validity: TrustChain, INFuture2017 Proceedings: The Future of Information Sciences / Atanassova, Iana ; Zaghouani, Wajdi ; Kragić, Bruno ; Aas, Kuldar ; Stančić, Hrvoje ; Seljan, Sanja (ur.). – Zagreb : Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia , 2017. 89-103.

[4] Filipović, Antun Matija; Bralić, Vladimir; Valić, Bruno: Društveni mediji u kriznim situacijama – naučene lekcije // 14. Znanstveno-stručna konferencija "Dani kriznog upravljanja 2021" – Zbornik radova / Toth, Ivan (ur.).Velika Gorica, 2022. str. 227-236.

[5] Bralić, Vladimir; Filipović, Antun Matija; Lozić, Davor: Pregled upotrebe algoritama za strojno učenje pri predviđanju potresa // 14. Znanstveno-stručna konferencija "Dani kriznog upravljanja 2021" – Zbornik radova / Toth, Ivan (ur.). Velika Gorica, 2022. str. 9-18.

[6] Bralić, Vladimir; Filipović, Antun Matija; Golubić, Iva. Primjena strojnog učenja u predviđanju kriznih događaja i prepoznavanju ugroza // 13. Znanstveno-stručna konferencija "dani kriznog upravljanja 2020" Zbornik radova / Toth, Ivan (ur.). Velika Gorica: Veleučilište Velika Gorica, 2020. str. 476-485.

[7] Bralić, Vladimir; Gaća, Krunoslav: Rizik od Anonymous napada na internet infrastrukture lokalne samouprave u Republici Hrvatskoj, 8. Međunarodna Znanstveno-stručna konferencija Dani kriznog upravljanja. Veleučilište Velika Gorica, Velika Gorica, 2015.

[8] Bralić, Vladimir; Kavran, Krešimir; Valić, Bruno. Automatic programming task grading – A case study in CodeRunner use at the University of applied sciences Velika Gorica // CIET 2020 Conference Proceedings / Kovačević, Tonko ; Akrap, Ivan (ur.). Split: University of Split – University Department of Professional Studies, 2020. str. 697-707.

[9] Filipović, Antun Matija; Bralić, Vladimir; Valić, Bruno. Osnovni IT sigurnosni testovi Kali Linuxom // Društvena i tehnička istraživanja, 1 (2020), 191-207.

[10] Bralić, Vladimir; Lozić, Davor; Valić, Bruno. Metode zaštite podataka na World Wide Webu // 13. Znanstveno-stručna konferencija "dani kriznog upravljanja 2020" zbornik radova / Toth, Ivan (ur.). Velika Gorica: Veleučilište Velika Gorica, 2020. str. 469-509.

[11] Filipović, Antun Matija; Bralić, Vladimir; Malešević, Nikola. Internet stvari i moguće ugroze // 12. Međunarodna znanstveno-stručna konferencija "Dani kriznog upravljanja 2019" Zbornik radova / Toth, Ivan (ur.). Velika Gorica: Veleučilište Velika Gorica, 2019. str. 808-819.

[12] Filipović, Antun Matija; Bralić, Vladimir; Obendorfer, Luka Dominic. Analiza ugroza putem društvenih mreža // 12. Međunarodna znanstveno-stručna konferencija "Dani kriznog upravljanja 2019" Zbornik radova / Toth, Ivan (ur.).Velika Gorica: Veleučilište Velika Gorica, 2019. str. 710-720.

[13] Bralić, Vladimir; Pavlović, Damir; Lebinac, Vladimir: Upravljanje nastavnim procesom – slučaj korištenja programskog rješenja LanSchool na Veleučilištu Velika Gorica, 15. CARNetova korisnička konferencija – CUC 2013, Zagreb: Hrvatska akademska i istraživačka mreža – CARNet, Zagreb, 2013.

[14] Topić, Tamara; Kožuh, Davor; Bralić, Vladimir: Primjena FMEA metode pri izradi analize rizika djelatnosti vezanih uz ionizirajuće zračenje, identifikacije mogućih izvanrednih događaja i evaluacije stupnja rizika. VI. međunarodna konferencija Dani kriznog upravljanja : Zbornik radova / Toth, Ivan (ur.). Velika Gorica : Veleučilište Velika Gorica, 2013.