Master’s Thesis

Peer Review Verification with Verifiable Credentials and Zero-Knowledge Proofs

Kaan Uzdogan

Supervisors: Prof. Jens Großklags, Prof. Ali Sunyaev

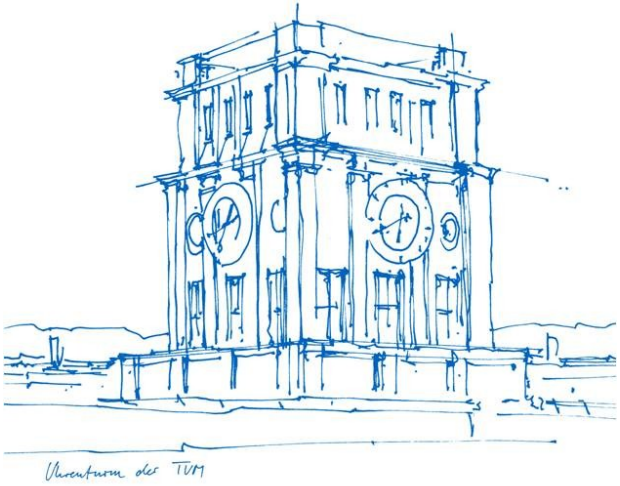
Advisors: Dr. Benjamin Sturm, Tibor Posa

Technische Universität München

Fakültät für Informatik

Lehrstuhl für Cyber Trust

1 February 2022



Problem Statement

- Peer review is essential for scientific publishing.
- Despite the importance many problems of peer review: bias, inconsistency, long publishing times, finding reviewers, review quality…
- Some of the problems due to the current incentives in academia to **publish more and review less.**
- Why?
  - Academic tenure, reputation etc. depend on publication based metrics
  - Scholars have limited time and reviews take time

Problem Statement

Closedness of Reviews Prevents Their Acknowledgement

- Nature of the reviews: mostly blinded and not published
- Unlike manuscripts, reviews are not publicly accessible, thus not possible to attribute.
- Even if a review is said to be done, it is **not possible to verify**.

Problem Statement

- Review showcase platforms tackling the problem of acknowledgement.
- Build a “peer review CV”. Forward the review email to them.
- Problems from an **open science perspective**:
  - Transparency: How are these reviews verified?
  - Access to data: API not providing much, scraping prohibited
  - Can be put behind a paywall, data access can be restricted
  - World’s peer review data effectively owned by a single for profit entity
- Due to them being **trusted third-parties**

Research Question

How can closed peer reviews be verified without trusted third-parties?

Research Problem

- Research mostly focused on incentivization of peer reviews
  - Reward/punishment
  - Reputation tokens
  - Bounties
- Existing works assume an open peer review system and don’t consider how to collect and verify peer review data.
- No research focusing to circumvent the third-parties in peer review verification

Goal

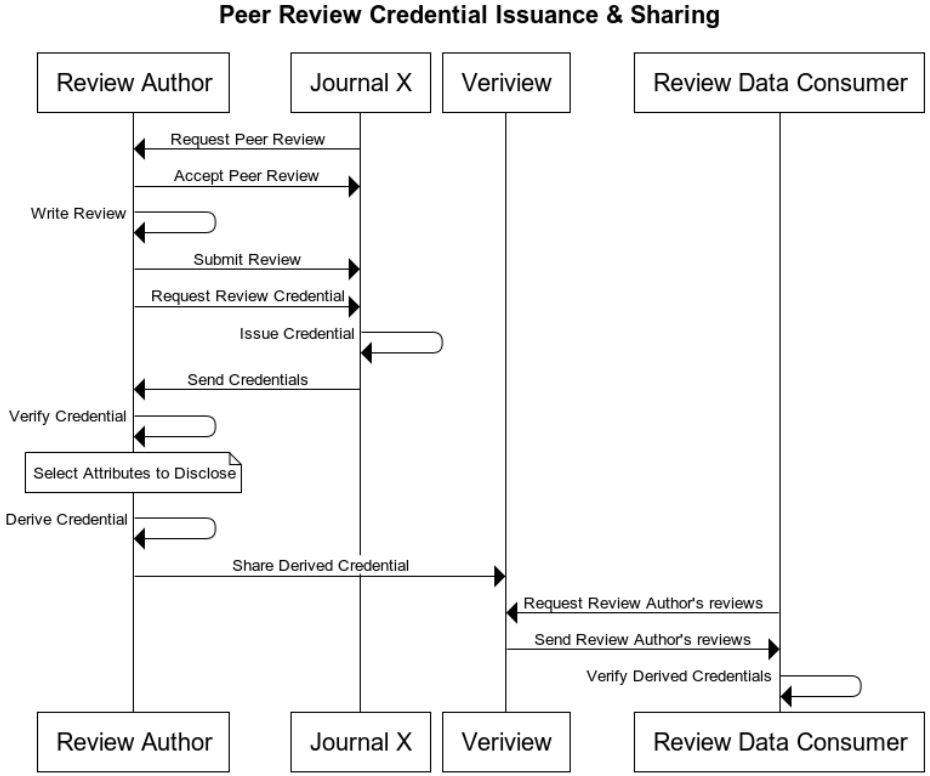
Design a system that will enable the verification of closed peer reviews without a trusted third-party

Design

Requirements based on problems stated:

- RQ1: Verifiability: Checking review authenticity
- RQ2: Transparency: How the verification is done
- RQ3: Open Data: Data stored = data public
- RQ4: Open Standards: Avoid lock-in, provide composability
- RQ5: Direct Trust: Avoid intermediaries
- RQ6: Selective Disclosure: Different data for different contexts, exclude identifying data in closed reviews.
- RQ7: Compatibility: Allow different review practices and different data schemas

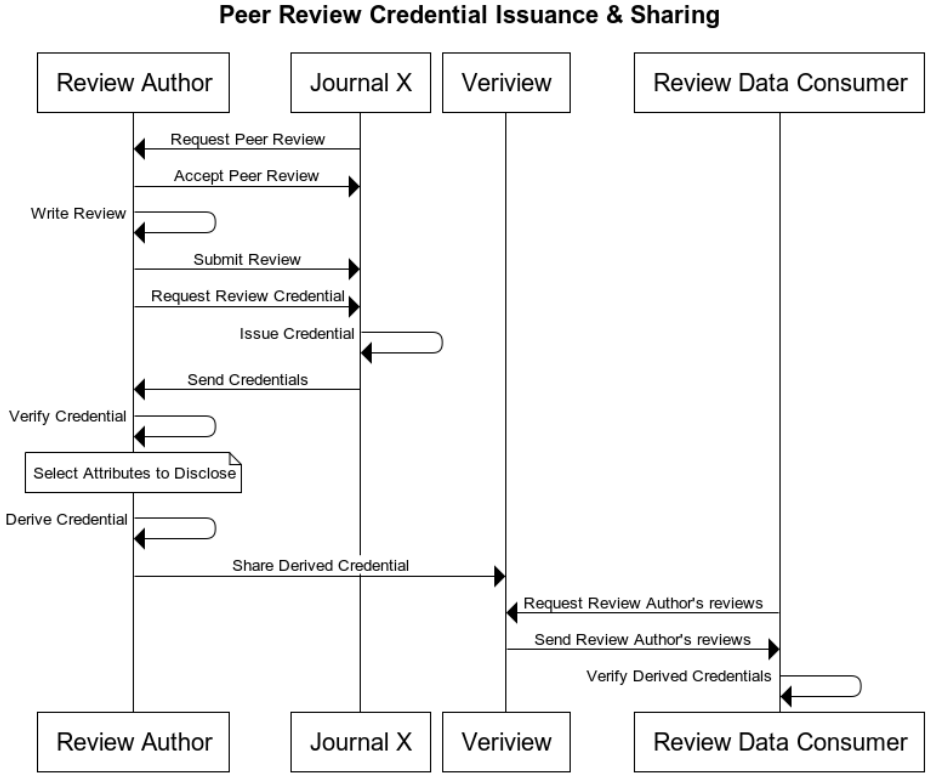
Design


Peer Review Credential Issuance & Sharing
Review Author
Journal X
Veriview
Review Data Consumer
Request Peer Review
Accept Peer Review
Write Review
Submit Review
Request Review Credential
Issue Credential
Send Credentials
Verify Credential
Select Attributes to Disclose
Derive Credential
Share Derived Credential
Request Review Author's reviews
Send Review Author's reviews
Verify Derived Credentials
Review Author
Journal X
Veriview
Review Data Consumer

Design

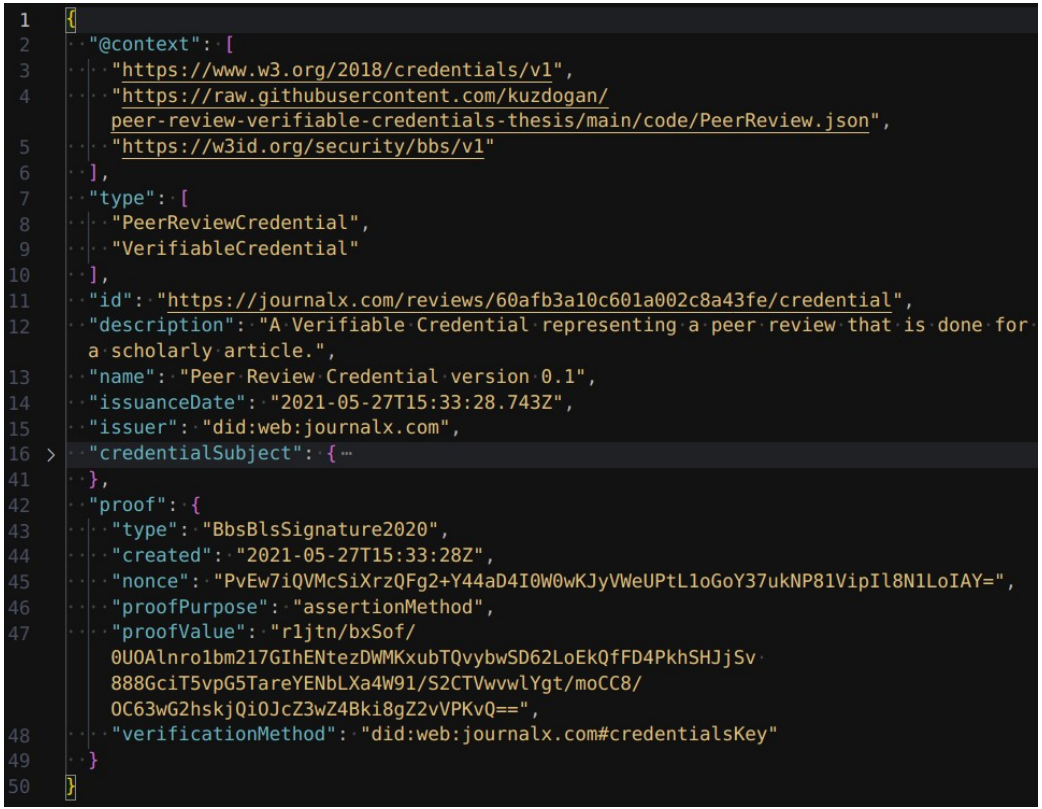
zk-proofs

Leave out identifying information e.g. title of the manuscript



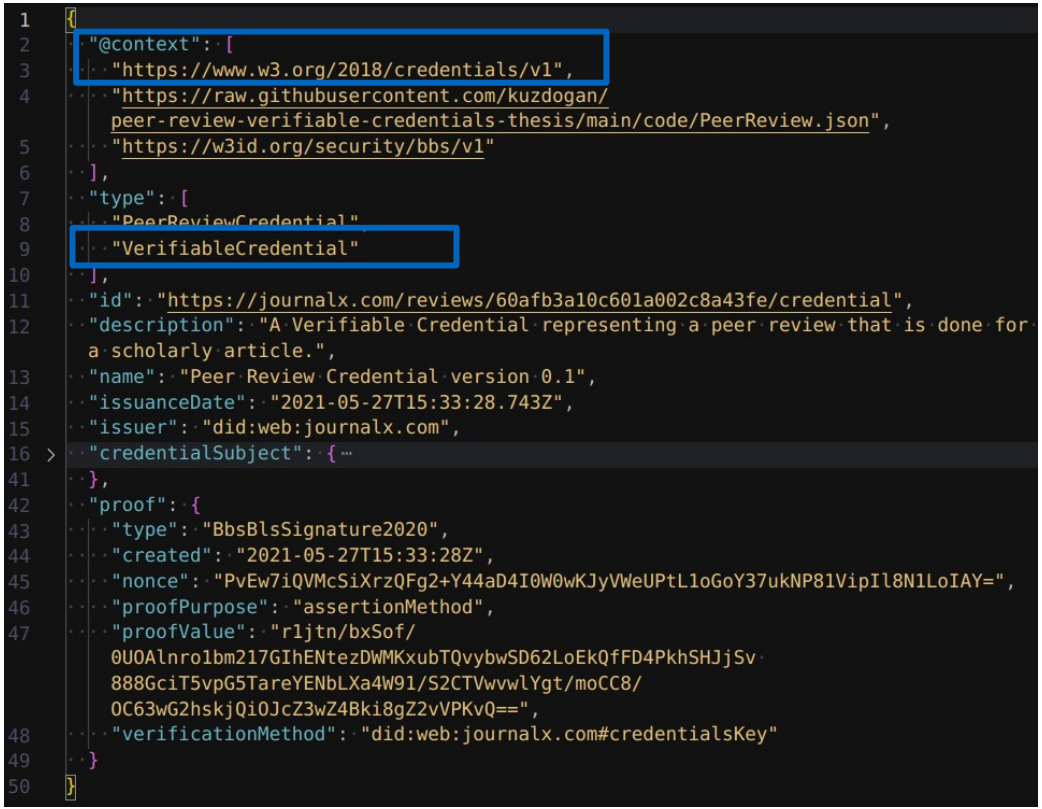
- Based on W3C’s Verifiable Credentials (VC) specification v1.0.
- Follows the JSON-LD (Linked Data) with Linked Proofs
  - Semantic Web
  - Vocabularies/Ontologies
- BBS+ signatures with BLS12-381 curve
  - Verification
  - Selective disclosure and zero-knowledge proofs

Example Credential Document



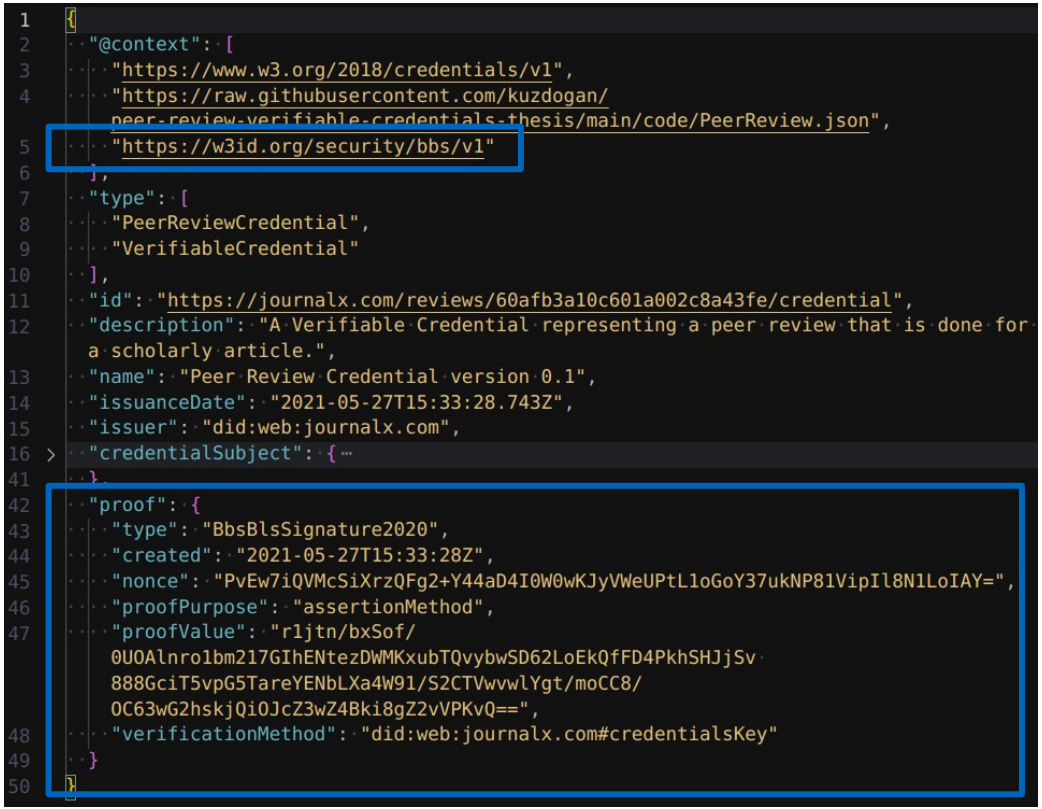
```
{
  "@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://raw.githubusercontent.com/kuzdogan/
    peer-review-verifiable-credentials-thesis/main/code/PeerReview.json",
    "https://w3id.org/security/bbs/v1"
  ],
  "type": [
    "PeerReviewCredential",
    "VerifiableCredential"
  ],
  "id": "https://journalx.com/reviews/60afb3a10c601a002c8a43fe/credential",
  "description": "A Verifiable Credential representing a peer review that is done for
  a scholarly article.",
  "name": "Peer Review Credential version 0.1",
  "issuanceDate": "2021-05-27T15:33:28.743Z",
  "issuer": "did:web:journalx.com",
  "credentialSubject": {…
  },
  "proof": {
    "type": "BbsBlsSignature2020",
    "created": "2021-05-27T15:33:28Z",
    "nonce": "PvEw7iQVMcSiXrzQFg2+Y44aD4I0W0wKJyVWeUPtL1oGoY37ukNP81VipIl8N1LoIAY=",
    "proofPurpose": "assertionMethod",
    "proofValue": "r1jtn/bxSof/
    0UOAlnro1bm217GIhENtezDWMKxubTQvybwSD62LoEkQfFD4PkhSHJjSv
    888GciT5vpG5TareYENbLXa4W91/S2CTVwvwlYgt/moCC8/
    OC63wG2hskjQiOJcZ3wZ4Bki8gZ2vVPKvQ==",
    "verificationMethod": "did:web:journalx.com#credentialsKey"
  }
}
```

Base Verifiable Credentials context

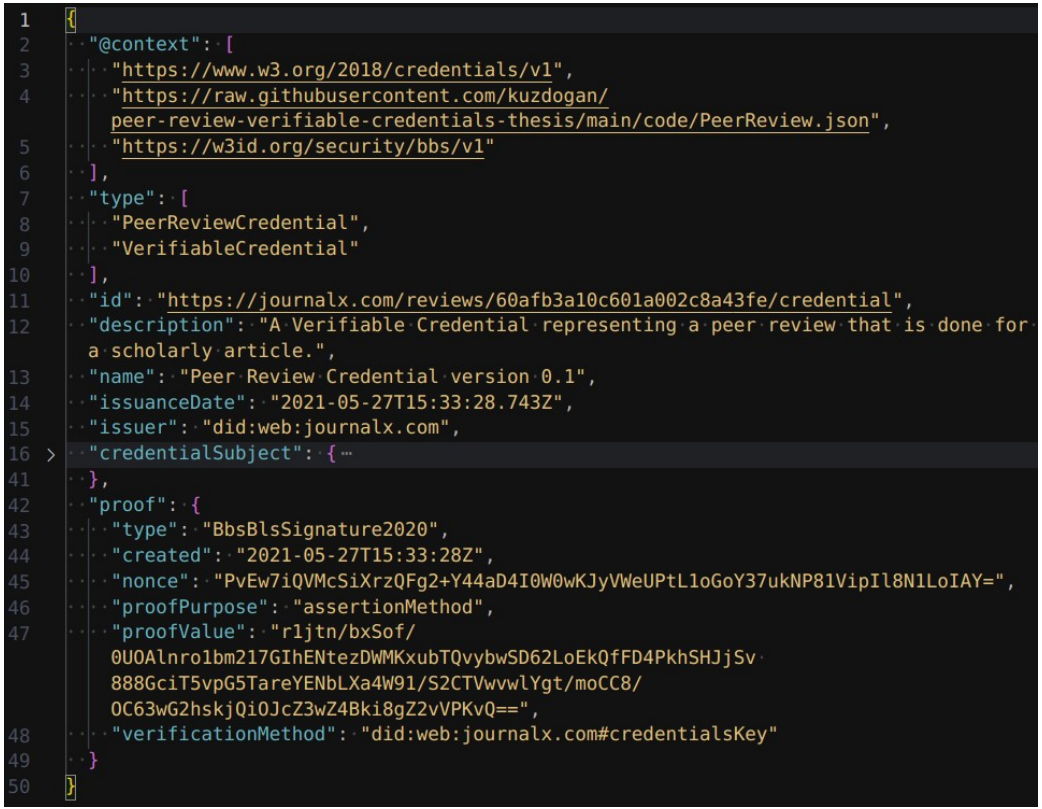


Linked Data signature definition

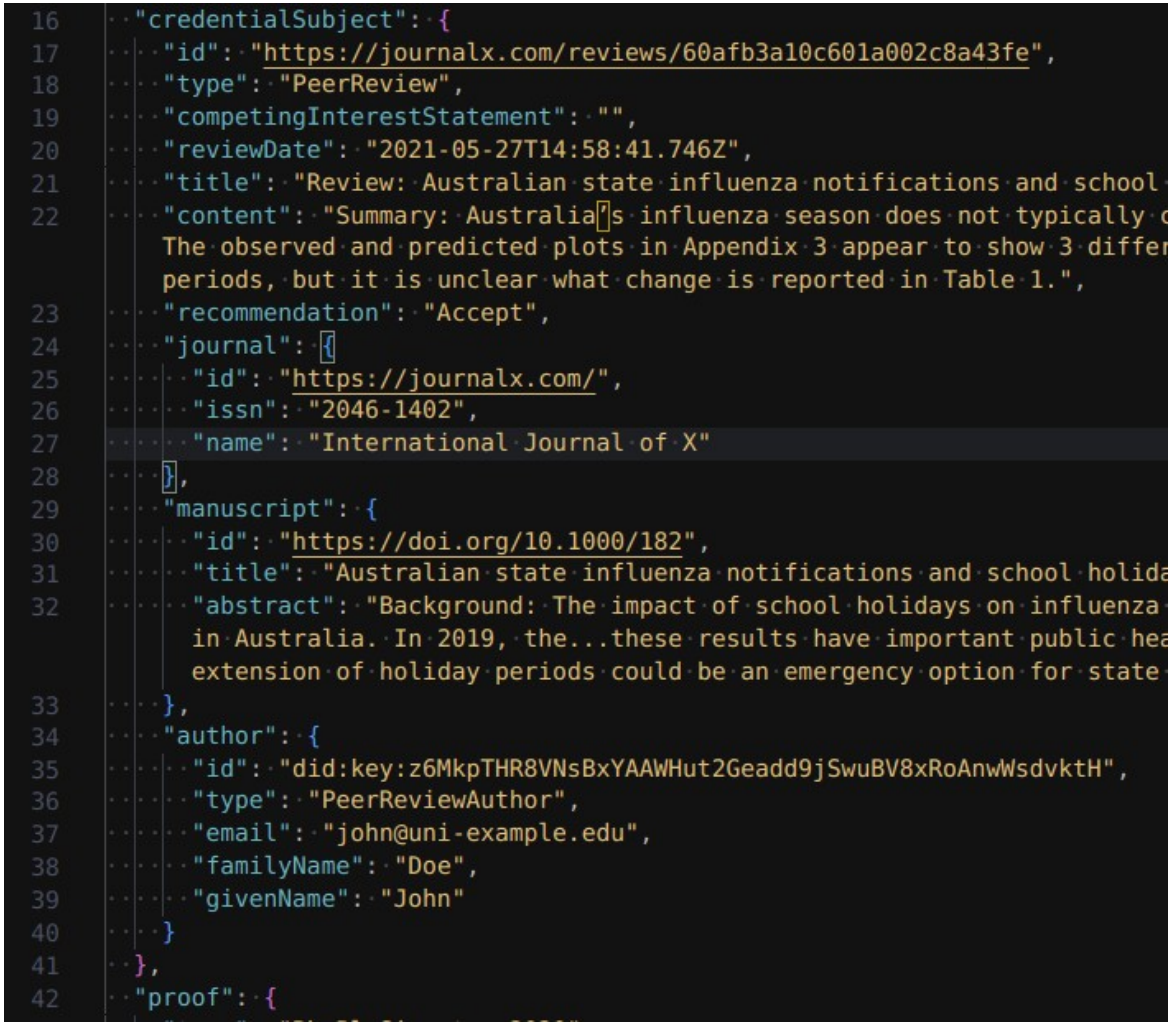
Linked Data signature



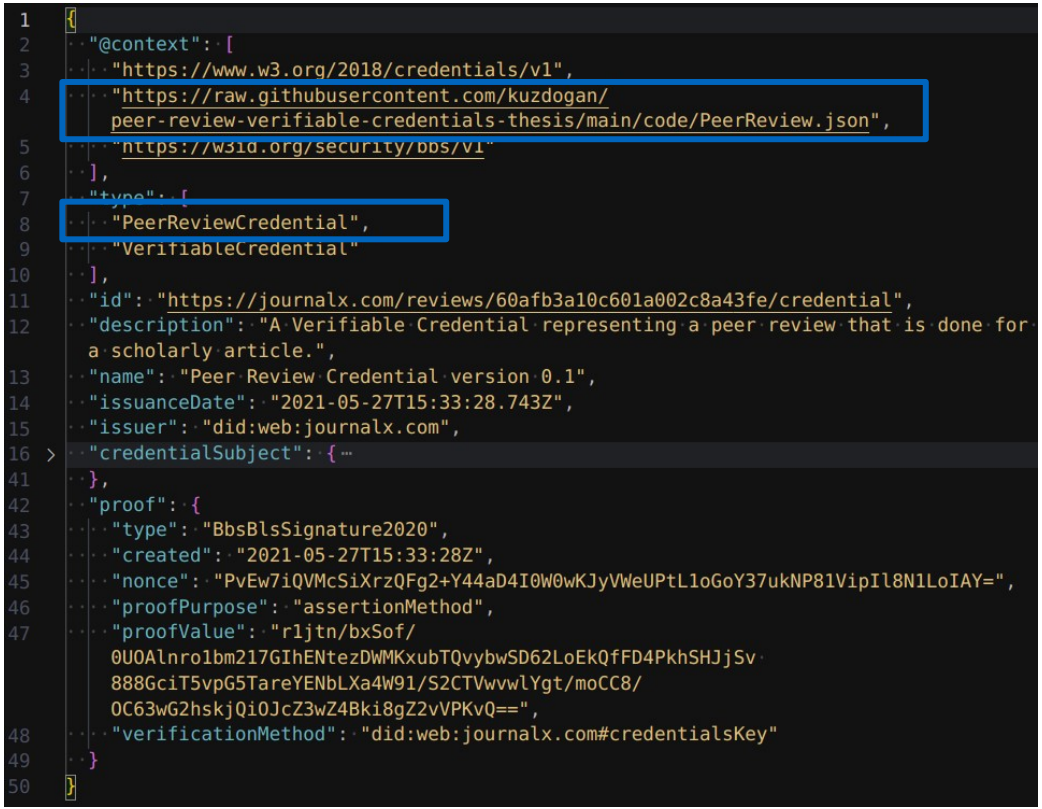
Expand



```
{
  "@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://raw.githubusercontent.com/kuzdogan/
    peer-review-verifiable-credentials-thesis/main/code/PeerReview.json",
    "https://w3id.org/security/bbs/v1"
  ],
  "type": [
    "PeerReviewCredential",
    "VerifiableCredential"
  ],
  "id": "https://journalx.com/reviews/60afb3a10c601a002c8a43fe/credential",
  "description": "A Verifiable Credential representing a peer review that is done for
  a scholarly article.",
  "name": "Peer Review Credential version 0.1",
  "issuanceDate": "2021-05-27T15:33:28.743Z",
  "issuer": "did:web:journalx.com",
  "credentialSubject": {…
  },
  "proof": {
    "type": "BbsBlsSignature2020",
    "created": "2021-05-27T15:33:28Z",
    "nonce": "PvEw7iQVMcSiXrzQFg2+Y44aD4I0W0wKJyVWeUPtL1oGoY37ukNP81VipIl8N1LoIAY=",
    "proofPurpose": "assertionMethod",
    "proofValue": "r1jtn/bxSof/
    0UOAlnro1bm217GIhENtezDWMKxubTQvybwSD62LoEkQfFD4PkhSHJjSv
    888GciT5vpG5TareYENbLXa4W91/S2CTVwvwlYgt/moCC8/
    OC63wG2hskjQiOJcZ3wZ4Bki8gZ2vVPKvQ==",
    "verificationMethod": "did:web:journalx.com#credentialsKey"
  }
}
```


"credentialSubject": {
"id": "https://journalx.com/reviews/60afb3a10c601a002c8a43fe",
"type": "PeerReview",
"competingInterestStatement": "",
"reviewDate": "2021-05-27T14:58:41.746Z",
"title": "Review: Australian state influenza notifications and school
"content": "Summary: Australia's influenza season does not typically
The observed and predicted plots in Appendix 3 appear to show 3 differ
periods, but it is unclear what change is reported in Table 1.",
"recommendation": "Accept",
"journal": {
"id": "https://journalx.com/",
"issn": "2046-1402",
"name": "International Journal of X"
},
"manuscript": {
"id": "https://doi.org/10.1000/182",
"title": "Australian state influenza notifications and school holida
"abstract": "Background: The impact of school holidays on influenza
in Australia. In 2019, the...these results have important public hea
extension of holiday periods could be an emergency option for state
},
"author": {
"id": "did:key:z6MkpTHR8VNsBxYAAWHut2Geadd9jSwuBV8xRoAnwWsdvktH",
"type": "PeerReviewAuthor",
"email": "john@uni-example.edu",
"familyName": "Doe",
"givenName": "John"
}
},
"proof": {

Defined peer review credential context

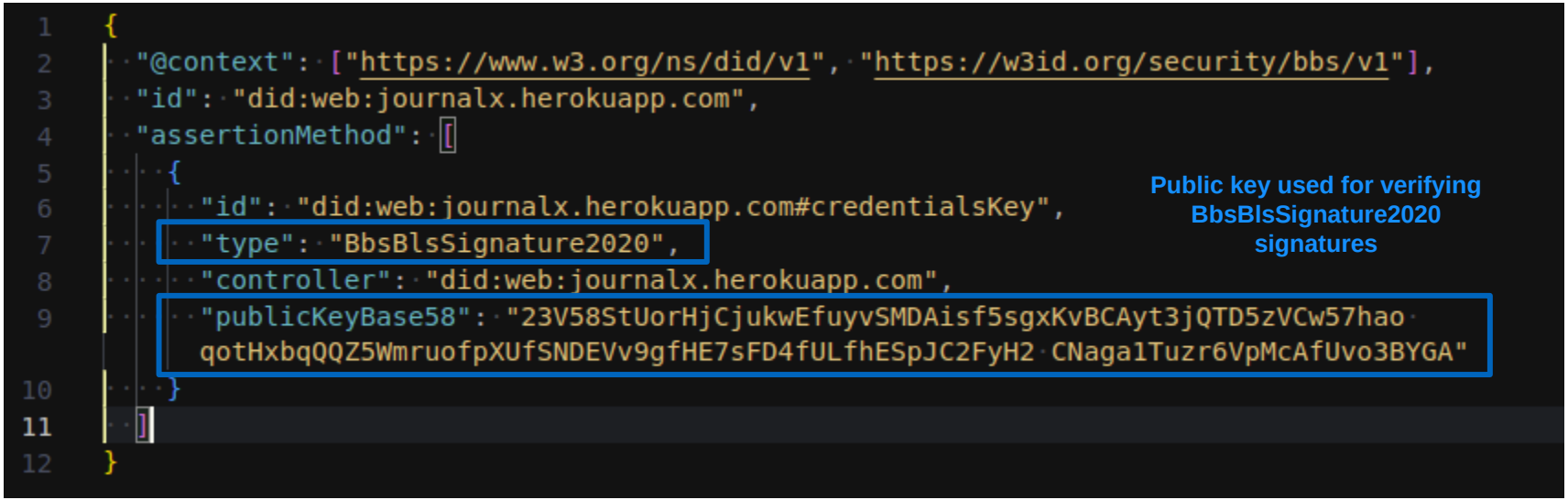


```
{
  "@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://raw.githubusercontent.com/kuzdogan/
    peer-review-verifiable-credentials-thesis/main/code/PeerReview.json",
    "https://w3id.org/security/bbs/v1"
  ],
  "type": [
    "PeerReviewCredential",
    "VerifiableCredential"
  ],
  "id": "https://journalx.com/reviews/60afb3a10c601a002c8a43fe/credential",
  "description": "A Verifiable Credential representing a peer review that is done for
  a scholarly article.",
  "name": "Peer Review Credential version 0.1",
  "issuanceDate": "2021-05-27T15:33:28.743Z",
  "issuer": "did:web:journalx.com",
  "credentialSubject": {…
  },
  "proof": {
    "type": "BbsBlsSignature2020",
    "created": "2021-05-27T15:33:28Z",
    "nonce": "PvEw7iQVMcSiXrzQFg2+Y44aD4I0W0wKJyVWeUPtL1oGoY37ukNP81VipIl8N1LoIAY=",
    "proofPurpose": "assertionMethod",
    "proofValue": "r1jtn/bxSof/
    0UOAlnro1bm217GIhENtezDWMKxubTQvybwSD62LoEkQfFD4PkhSHJjSv
    888GciT5vpG5TareYENbLXa4W91/S2CTVwvwlYgt/moCC8/
    OC63wG2hskjQiOJcZ3wZ4Bki8gZ2vVPKvQ==",
    "verificationMethod": "did:web:journalx.com#credentialsKey"
  }
}
```

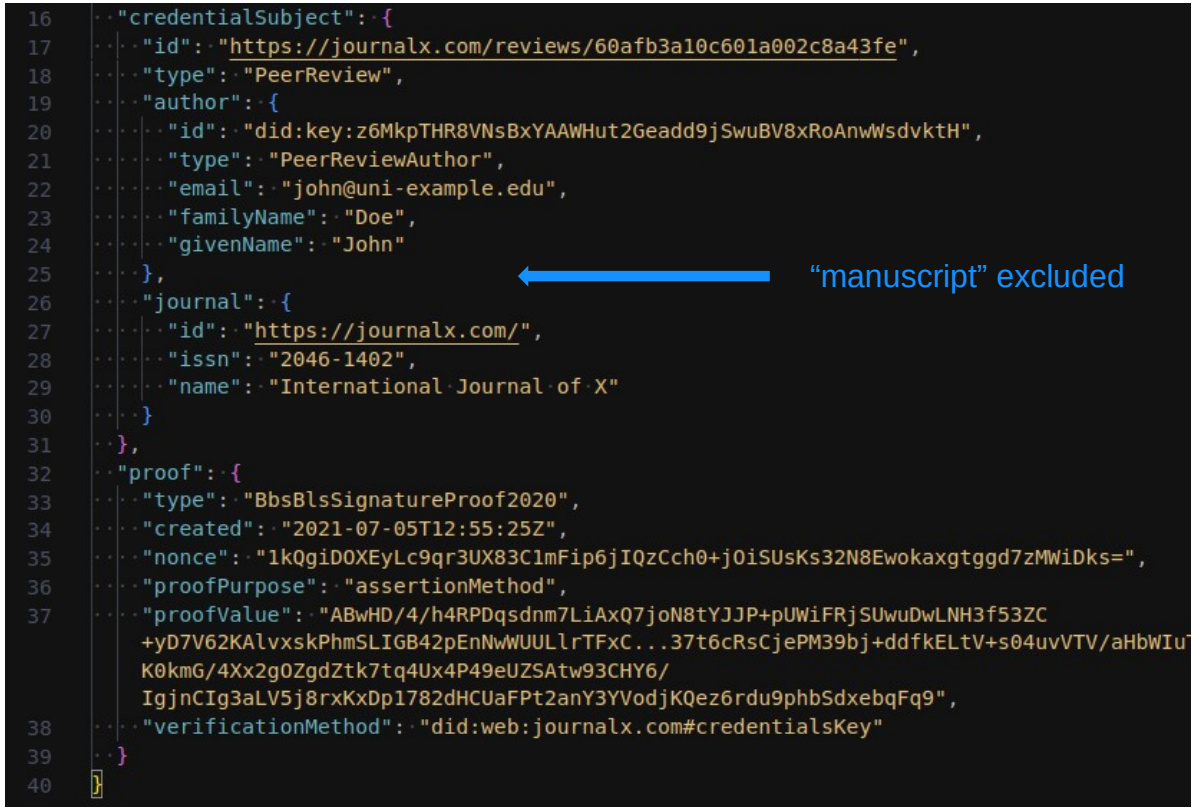
Decentralized Identifiers (DIDs)

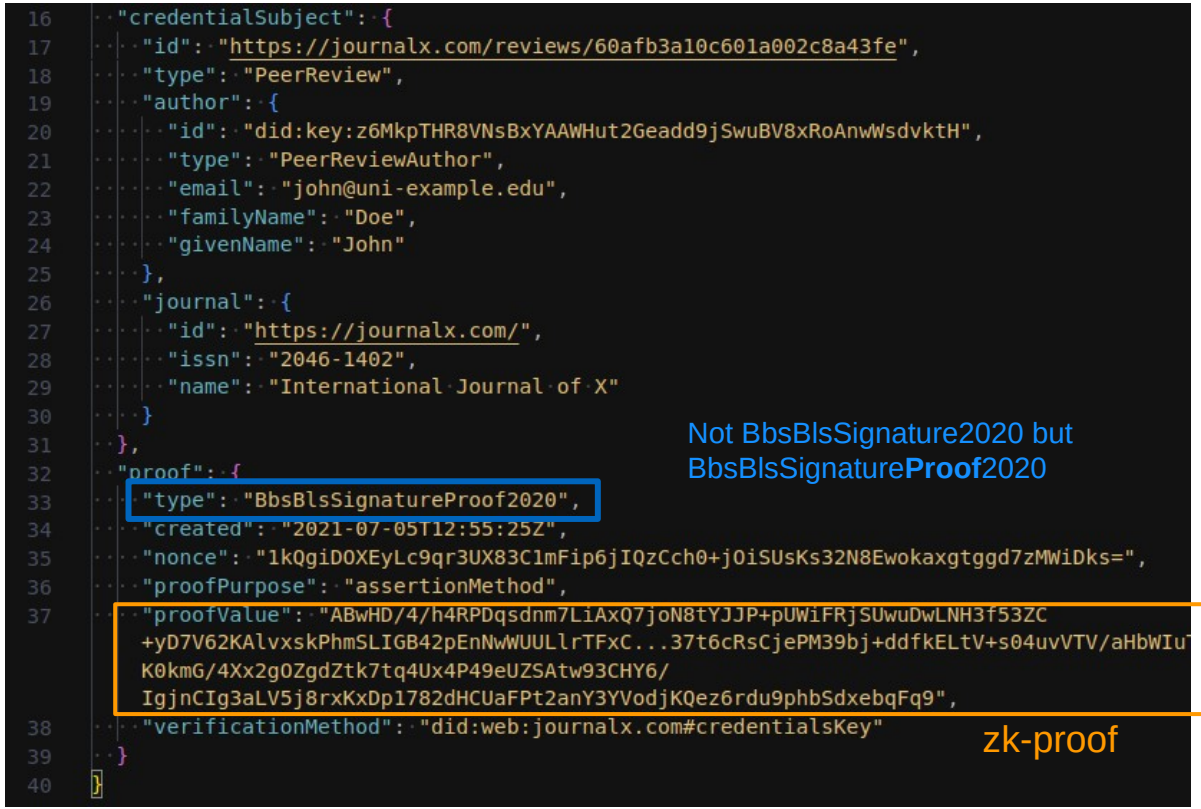
- Used for the verification of the document
- Distributed ledger based:
  - Highly available
  - Tamperproof
  - e.g. `did:sov:mnjkl98uipsndg2hdjdjuf7`
  - Requires binding `did:sov:mnjkl98uipsndg2hdjdjuf7` with the real world identity
- Non-distributed ledger based:
  - `did:web:journalx.com` stored under `https://<HOST-NAME>/.well-known/did.json`
  - Requires server to be online
  - Can be changed without noticing
  - Uses the domain name and the TLS identity of journal website.

Example DID document

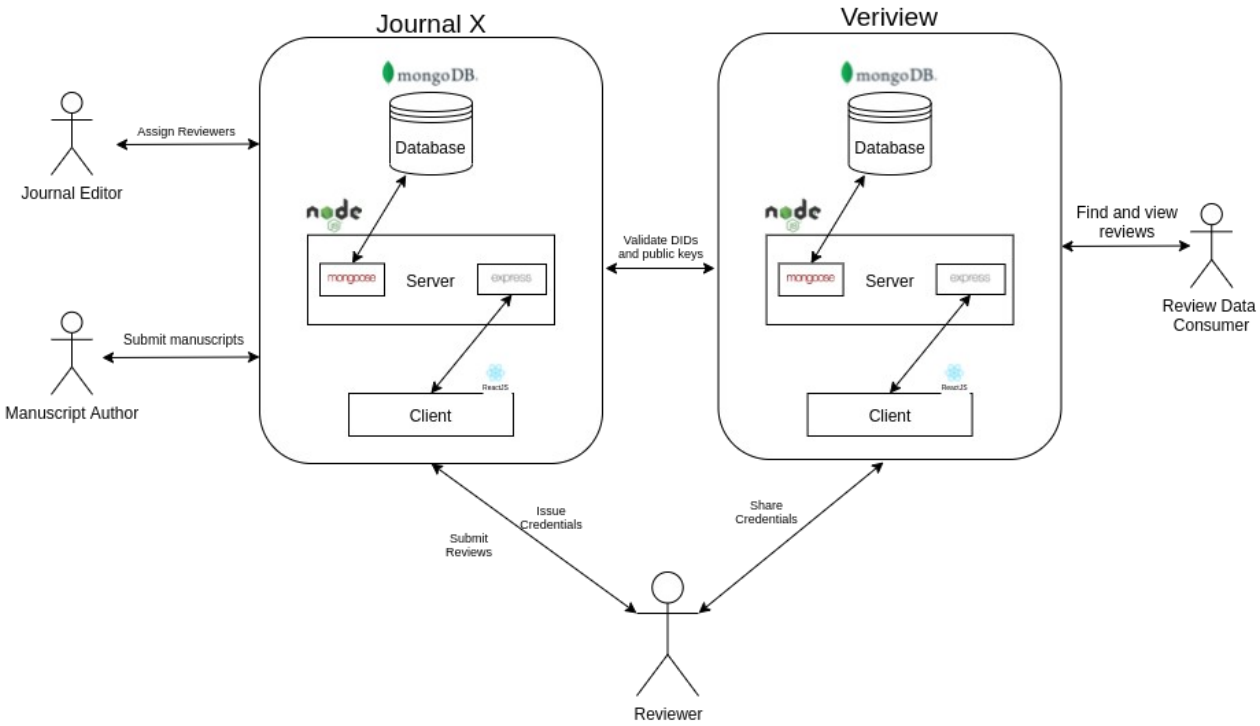


Selectively Disclosed Credential



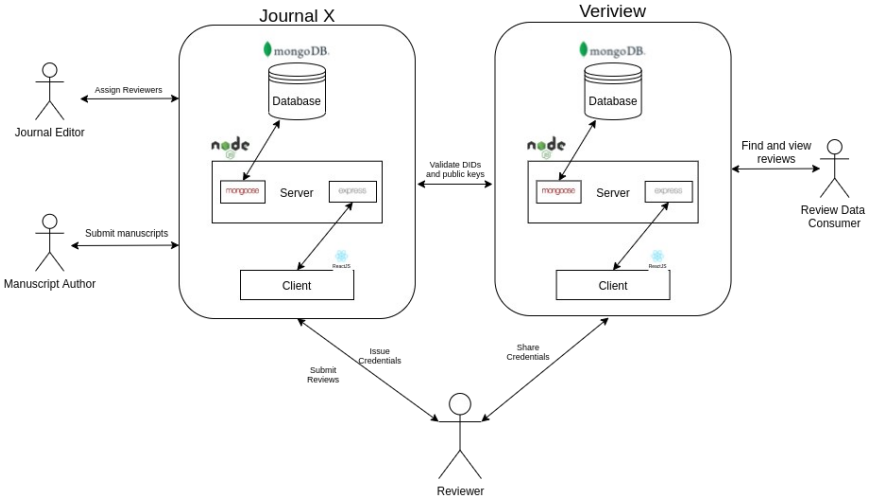

"credentialSubject": {
"id": "https://journalx.com/reviews/60afb3a10c601a002c8a43fe",
"type": "PeerReview",
"author": {
"id": "did:key:z6MkpTHR8VNsBxYAAWHut2Geadd9jSwuBV8xRoAnwWsdvktH",
"type": "PeerReviewAuthor",
"email": "john@uni-example.edu",
"familyName": "Doe",
"givenName": "John"
},
"journal": {
"id": "https://journalx.com/",
"issn": "2046-1402",
"name": "International Journal of X"
}
},
"proof": {
"type": "BbsBlsSignatureProof2020",
"created": "2021-07-05T12:55:25Z",
"nonce": "1kQgiDOXEyLc9qr3UX83C1mFip6jIQzCch0+jOiSUsKs32N8Ewokaxgtggd7zMWiDks=",
"proofPurpose": "assertionMethod",
"proofValue": "ABwHD/4/h4RPDqsdnm7LiAxQ7joN8tYJJP+pUWiFRjSUwuDwLNH3f53ZC
+yD7V62KAlvxskPhmSLIGB42pEnNwWUULlrTFxC...37t6cRsCjePM39bj+ddfkELtV+s04uvVTV/aHbWIu
K0kmG/4Xx2gOZgdZtk7tq4Ux4P49eUZSAtw93CHY6/
IgjnCIg3aLV5j8rxKxDp1782dHCUaFPt2anY3YVodjKQez6rdu9phbSdxebqFq9",
"verificationMethod": "did:web:journalx.com#credentialsKey"
}
}
Not BbsBlsSignature2020 but
BbsBlsSignatureProof2020
zk-proof

Prototype

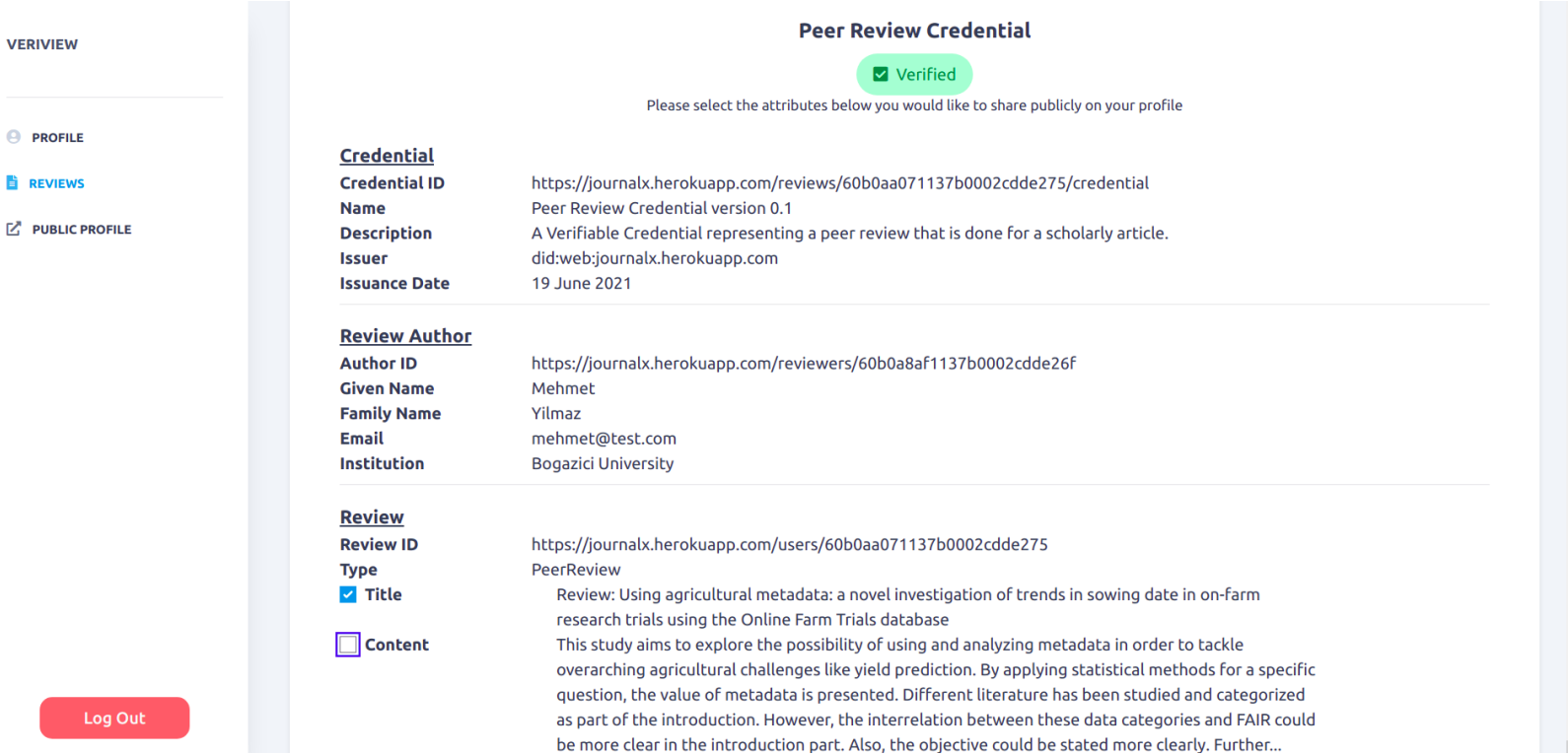

Journal X
Veriview
mongoDB.
Database
node
mongoose
Server
express
ReactJS
Client
Journal Editor
Assign Reviewers
Manuscript Author
Submit manuscripts
Validate DIDs and public keys
Find and view reviews
Review Data Consumer
Issue Credentials
Submit Reviews
Share Credentials
Reviewer

Prototype

- A working prototype is built to demonstrate the feasibility of the designed system
- Code and instructions: https://github.com/kuzdogan/peer-review-verifiable-credentials-thesis
- Deployed to
  - https://journalx.herokuapp.com/
  - https://veriview.herokuapp.com/



Reviewers can issue verifiable credentials for the reviews done



Selecting attributes to disclose on Veriview prototype

VERIVIEW

PROFILE

REVIEWS

PUBLIC PROFILE

Log Out

Selectively Disclosed Peer Review Credential

SHOW CODE DOWNLOAD CREDENTIAL

SUBMIT

Credential

Credential ID	https://journalx.herokuapp.com/reviews/60b0aa071137b0002cdde275/credential
Name	Peer Review Credential version 0.1
Description	A Verifiable Credential representing a peer review that is done for a scholarly article.
Issuer	did:web:journalx.herokuapp.com
Issuance Date	19 June 2021

Review Author

Author ID	https://journalx.herokuapp.com/reviewers/60b0a8af1137b0002cdde26f
Given Name	Mehmet
Family Name	mehmet@test.com
Email	mehmet@test.com
Institution	Bogazici University

Review

Review ID	https://journalx.herokuapp.com/users/60b0aa071137b0002cdde275
Type	PeerReview

Journal

ID	https://journalx.herokuapp.com/

Evaluation

Fulfilling self imposed requirements:

- RQ1: Verifiability
- RQ2: Transparency
- RQ3: Open Data
- RQ4: Open Standards
- RQ5: Direct Trust
- RQ6: Selective Disclosure
- RQ7: Compatibility

60 min interviews with 5 different researchers with peer review experience

- Personal connections and contacting Publons users directly.
- Finding out their perspectives on peer review
- Their insights into showcase platforms
- Introduce conceptual design
- User test prototype
- Solicited review on the system

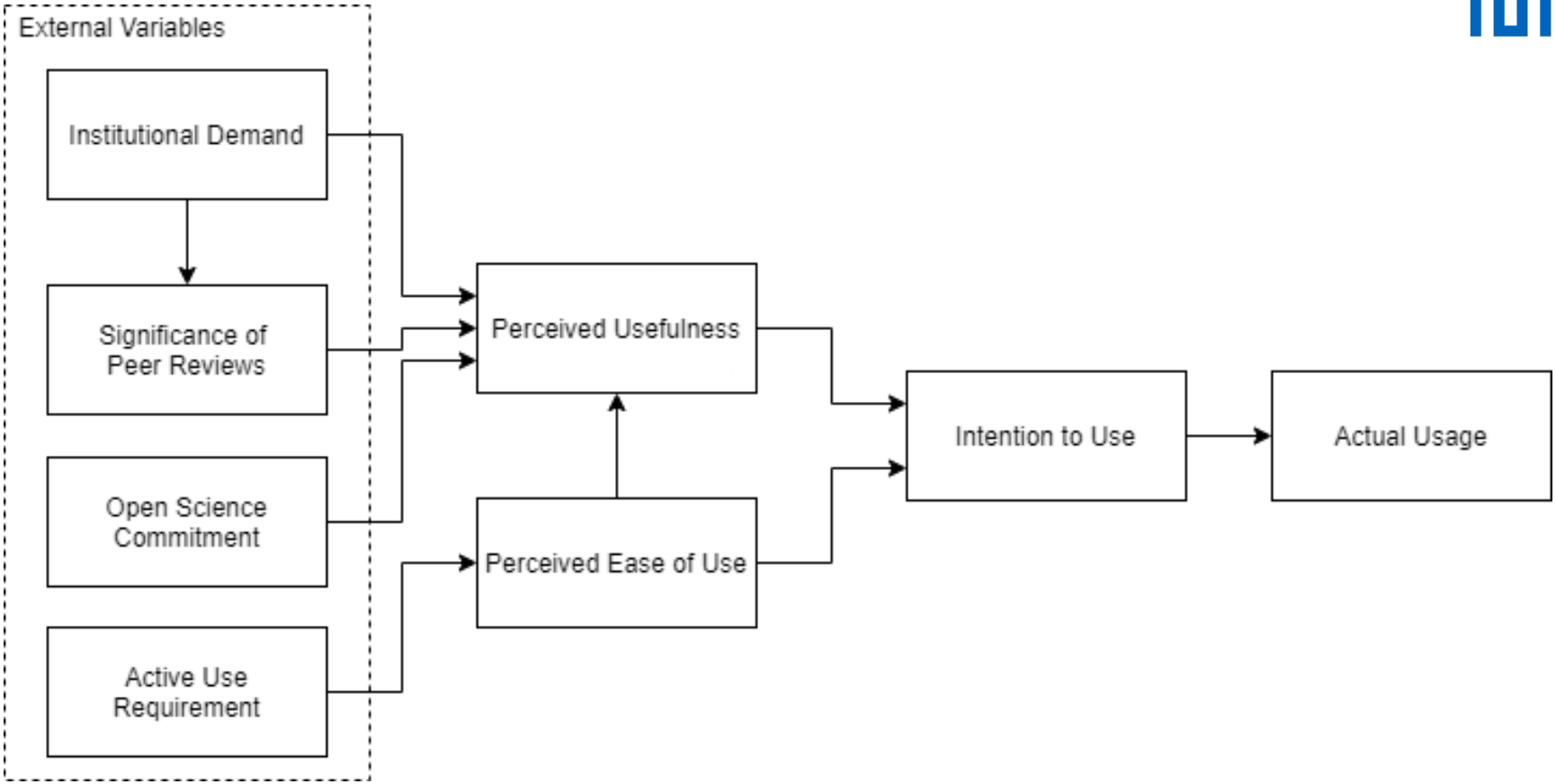
Evaluation

Interview Findings:

- **Too much user involvement:** download-upload and manage credentials Should be automated as possible.
- **Worth the effort?** Some researchers don’t seem to attribute much importance to reviews. For them the system would not provide much value.
- **Need to verify reviews?** Is peer review verification really a necessity? There is not much incentive to lie on reviews
  - Needs validation but there are already widely used platforms so this is also a validation.
- **Open science aspect:** Mostly yes, existing platforms can be a concern in terms of open science

Acceptance Model

- A model for the adoption of the proposed system
- Based on the Technology Acceptance Model (TAM) by Davis, Bagozzi, & Warshaw, 1989
  - Usage of a system depends on the **perceived usefulness** and **perceived ease of use** of the system
- Variables influencing perceived usefulness and perceived ease of use were extracted from interviews.


External Variables
Institutional Demand
Significance of Peer Reviews
Open Science Commitment
Active Use Requirement
Perceived Usefulness
Perceived Ease of Use
Intention to Use
Actual Usage

Discussion and Conclusion

- The dilemma of peer review: Transparency & accountability vs. need for anonymity
- The designed system does not solve the dilemma but attempts to bring two ends closer by leveraging emerging standards and cryptography.
- Learnings from scientific publishing: adherence to open science principles

Shortcomings:

- Journals need to issue credentials
- Consensus on peer review vocabularies
- Peer review habits difficult to change
- Cold start problem
- Inherits the UX problems of Self-Sovereign Identity and Verifiable Credential systems.
  - Credential and key management
- Who will run the system?

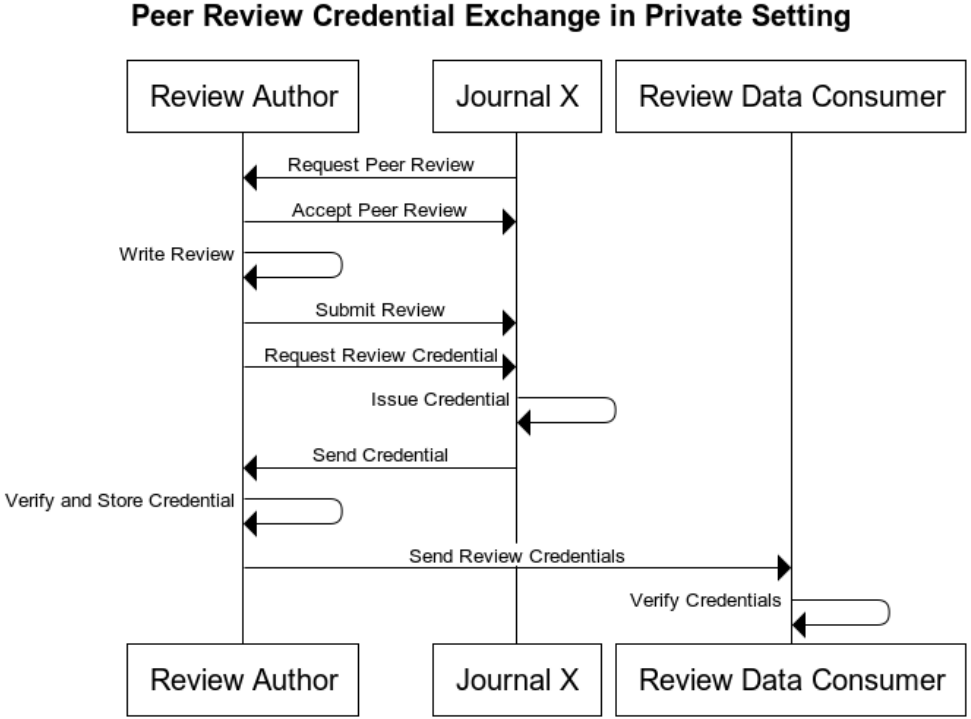
- Wider fundamental problems:
  - What is a good review?
  - How can it be measured?
  - How should it be incentivized or should it be?

- As with anything involving humans, peer review is not perfect either. But it can only improve by research and experimentation.

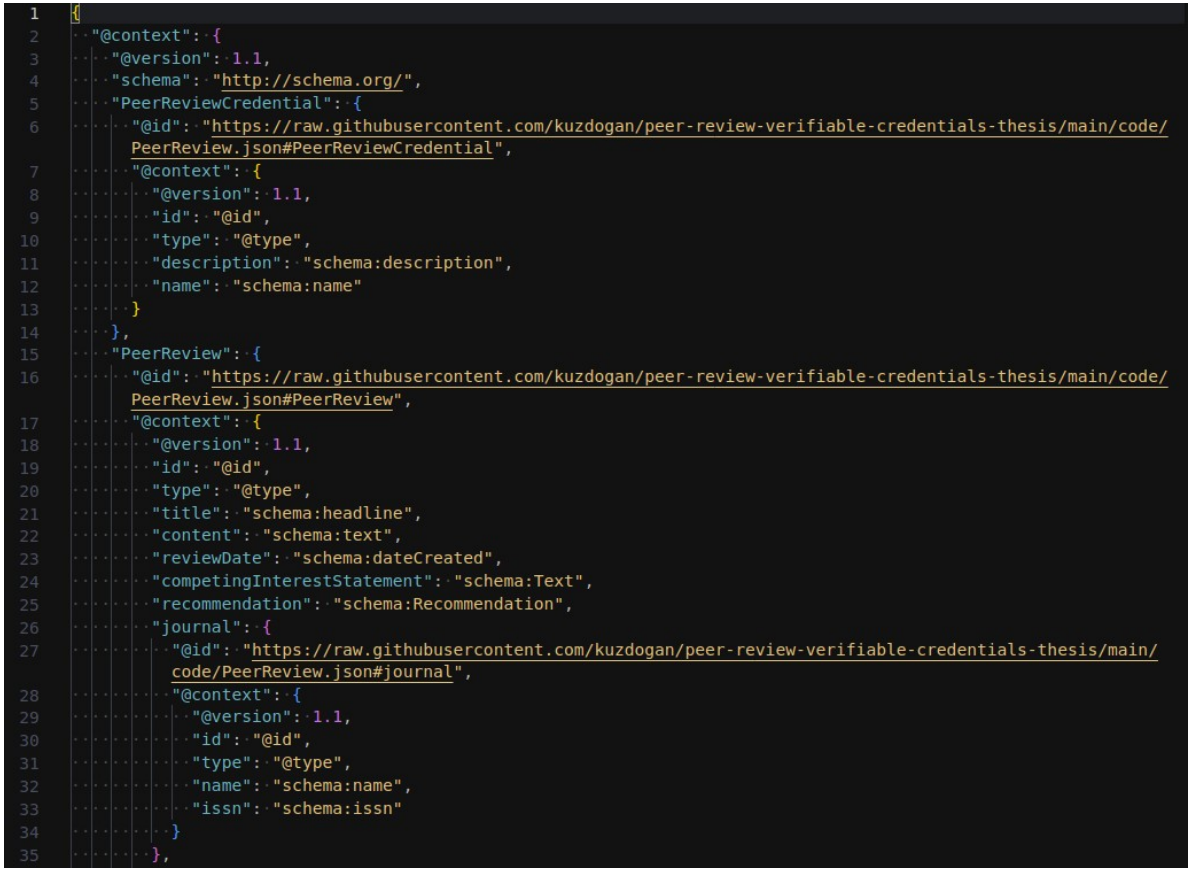
Thank you for your attention

Questions?

Additional Resources

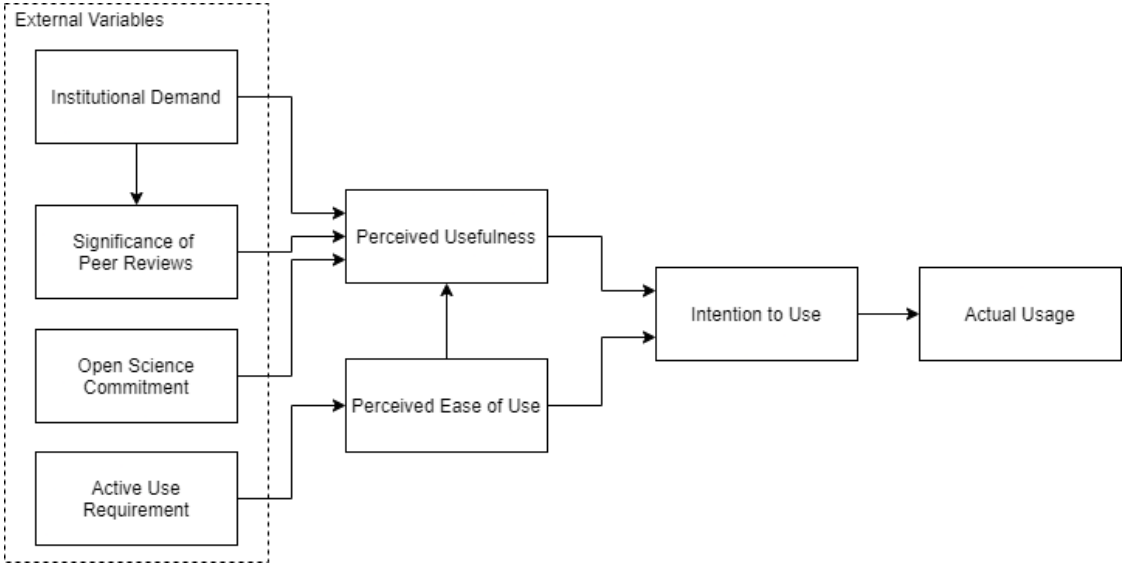


- In a private setting there’s no need for attribute selection and the Veriview platform.

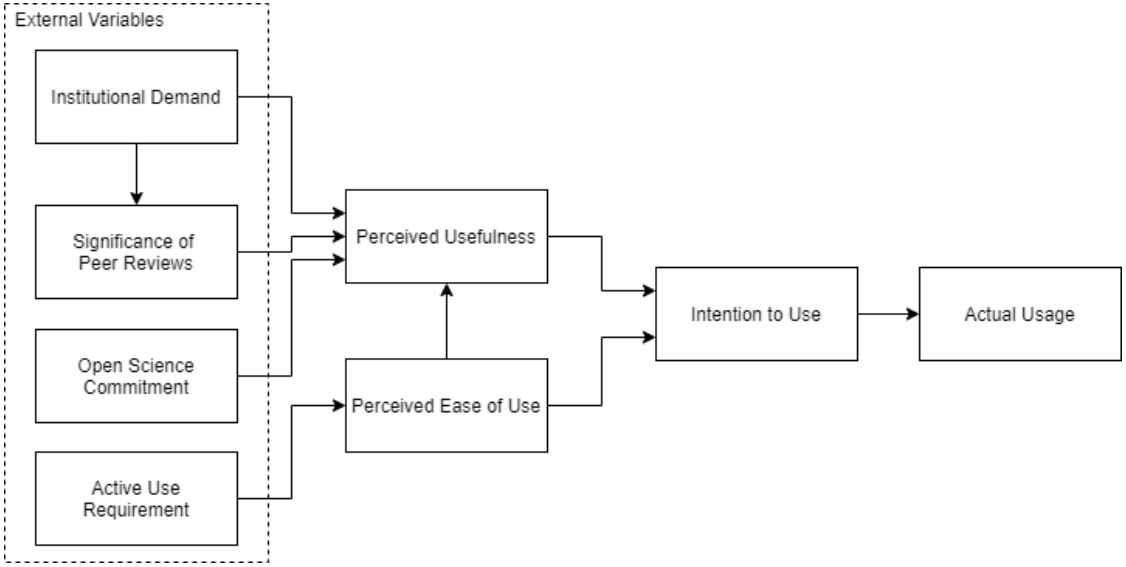


```
{
  "@context": {
    "@version": 1.1,
    "schema": "http://schema.org/",
    "PeerReviewCredential": {
      "@id": "https://raw.githubusercontent.com/kuzdogan/peer-review-verifiable-credentials-thesis/main/code/PeerReview.json#PeerReviewCredential",
      "@context": {
        "@version": 1.1,
        "id": "@id",
        "type": "@type",
        "description": "schema:description",
        "name": "schema:name"
      }
    },
    "PeerReview": {
      "@id": "https://raw.githubusercontent.com/kuzdogan/peer-review-verifiable-credentials-thesis/main/code/PeerReview.json#PeerReview",
      "@context": {
        "@version": 1.1,
        "id": "@id",
        "type": "@type",
        "title": "schema:headline",
        "content": "schema:text",
        "reviewDate": "schema:dateCreated",
        "competingInterestStatement": "schema:Text",
        "recommendation": "schema:Recommendation",
        "journal": {
          "@id": "https://raw.githubusercontent.com/kuzdogan/peer-review-verifiable-credentials-thesis/main/code/PeerReview.json#journal",
          "@context": {
            "@version": 1.1,
            "id": "@id",
            "type": "@type",
            "name": "schema:name",
            "issn": "schema:issn"
          }
        },
```

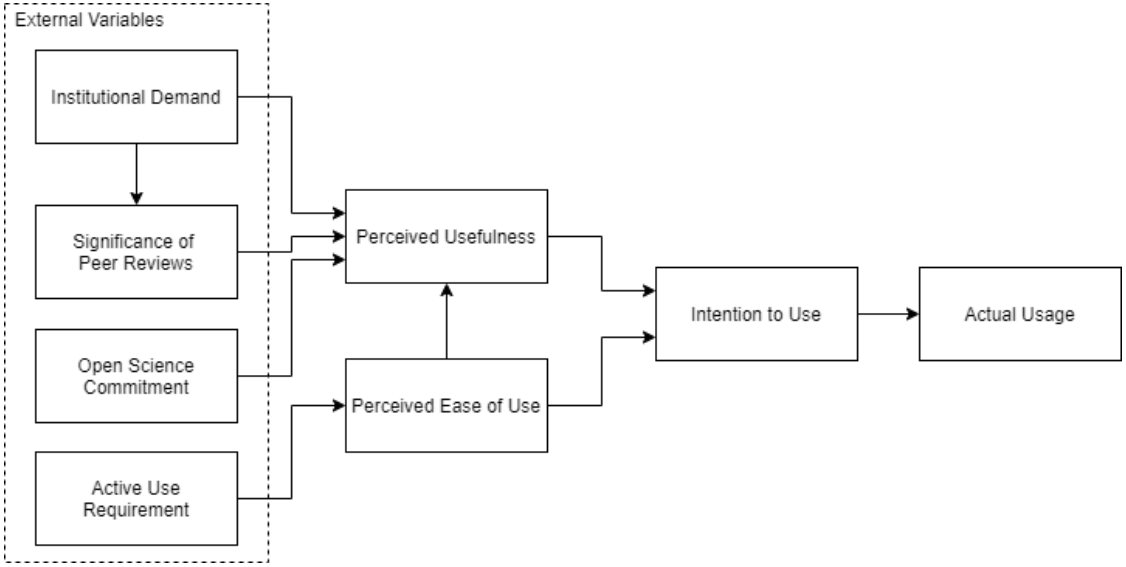
Defined peer review context



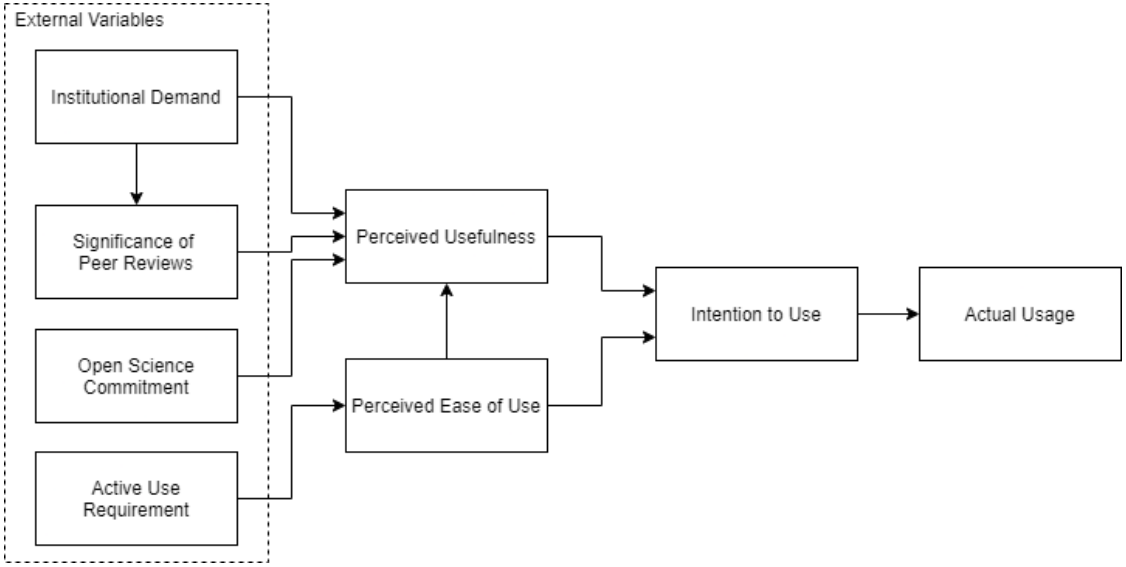
- Institutional Demand: Peer reviews are not an important part of the resume as it is not considered by institutions.
- Higher institutional demand will have a positive impact on the perceived usefulness of the system



- Significance of Reviews: Value given by researchers, recognition from institutions, perception of the broader scientific community
- Higher significance of reviews will have a positive impact on the perceived usefulness
- Chicken and egg problem here: Showcasing platforms not used because peer review is not recognized, and peer review is not recognized because peer review can't be "showcased".
- Room for positive feedback: more visibility → more recognition → more showcasing → more visibility



- Open Science Commitment of a researcher has a positive impact on the perceived usefulness.



- Active Use: As researchers don’t spend much time on peer review, more active use would influence the ease of use negatively.
  - This variable is the only variable of the designed system. The rest are outside the system design.