Assignment, WASP Software Engineering

Carl Magnus Bruhner

August 2025

1 Introduction

Certificates are the foundation for secure communication over the Internet, but not all certificates are created and managed consistently. Certificate authorities (CAs) issuing certificates achieve different levels of trust, and user trust in public keys, certificates, and CAs can quickly change. This has made careful certificate management an important and challenging problem.

My research focuses on certificate management in the Web Public Key Infrastructure (Web PKI). More broadly, I study Internet security and privacy, but primarily the Web PKI with X.509 certificate usage and management, root store and CA/Browser policies and practices, trust, and related areas. I use empirical research methods, including passive and active internet measurements to study and understand behaviors “in the wild”. I identify trends and challenges and contrast findings to practices, standards, and requirements to find non-compliance, pain points, and bottlenecks, and potential improvements both technical and administrative.

The measurement studies are made possible through certificate transparency data and datasets based on scanning/crawling (existing and/or own). Research opportunities include (but are not limited to) usage characterization, management practices, requirements, revocations, transparency, etc. The aim is to propose enhancements to certificate practices/PKI that help strengthen the security of data communication protected by certificates.

Our work thus far presents novel characterization of certificate replacement (CR) relationships in the wild including reuse of public keys, head-to-head comparison of CRs where the replaced certificate was revoked vs. not revoked, and longitudinal analysis of certificate chains for popular domains, examining their evolution over time and across categories. Based on the results, we have highlighted shortcomings in existing revocation protocols and practices, emphasizing the need for improvements, proposing efficient solutions, and discussing other ongoing efforts aimed at addressing these problems.

The current research direction is to investigate challenges with certificate revocation that previous studies have proven broken, especially on mobile devices. Broken revocations mean that certificates can continue to be trusted even after explicitly revoked, allowing malicious actors to impersonate the webpage leaving everyday users of the Internet unknowingly vulnerable.

2 Lecture principles

2.1 SE = CS + Reality

One idea that really stuck with me was this concept of ‘Software Engineering’ being ‘Computer Science’ plus *‘Reality’.* Reality can be (as exemplified) people, process, business, modeling, QA, maintenance, scale, ethics, and many other things. This made me scratch down some notes on parallels to my own research area, namely certificates and how everything seem so figured out and guided by clear policies (root store policies, CA/Browser Forum baseline requirements, RFCs, etc.) but still there are so many examples of certificate misissuance, revocation mechanisms not working, etc. It is simply because of Reality! It also matches (and motivates) the research that I do in terms of active and passive measurements, looking to see how things work “in the wild” versus by standards. Reality messes up a lot, and just as we study nature, we can study the Internet and practices online to find out how things really work in order to improve it going forward. This also connects to behavioral software engineering, that is, Reality = People are not rational.

2.2 SE principle: Social interactions

One of the mentioned software engineering “principles” was ‘Social interactions’, described as between individuals and within/between teams. Expanding on this was the idea to reach out to companies to seek understanding of what the real drivers and motivations really were (e.g. understanding internal politics of how to estimate software engineering costs: keep personnel or compete against other sites). This is also of interest to my research, as I have heard from people in industry (and successful researchers) that many of the “academic” projects tend to be too theoretical for practical impact. Even if there is value in purely theoretical work, there might be even more value in practical impact. Especially in empirical studies, it might be more rewarding to see how suggested improvements can lead to real-world impact that can be measured in future studies. One way of complementing empirical studies is to reach out to industry in an aim to identify challenges that they face and use that as input when generating ideas of what to study.

3 Guest-Lecture Principles

3.1 Understand the problem before you build the solution

The name of the whole first guest lecture, but still an overarching idea that really resonated with me, was (stylized): “Understand the *problem* **before** you build the *solution*”. This was elaborated on in terms of problem-space versus solution-space. When receiving a solution-space statement, first determine the problem you are trying to solve with this solution. It is closely related to the above discussion on social interactions and understanding real-world challenges.

3.2 Stakeholder → Goal → System → Requirements

This too from the first guest lecture, presented as a technique for identifying usable/pragmatic and relevant research, the presented model looked promising: stakeholder elicitation, goal modeling, system vision, and requirements elicitation. Even though it was presented in a software engineering context, I thought of it as a relevant way to identify usable research—again connecting the discussion to the previous (social interaction and understanding the problem before building the solution). Stakeholder elicitation could be used, in my case, to visualize the certificate/web PKI ecosystem. Goal modeling can be used to see what my solution should satisfy given the stakeholders and the relationship (with potential conflicts) between these goals. Based on this, I can create a system vision followed by requirements elicitation to determine something measurable and specific that my research and solutions should satisfy (or strive for).

4 Data Scientists versus Software Engineers

Do you agree on the essential differences between data scientists and software engineers put forward in these chapters? Why or why not?

I agree, or at least I think it is a good basis for discussion and reflection. The book also notes that the distinctions are oversimplified and overgeneralized, even if they describe differences seen in practice. Personally, I think software engineers especially are more of a broader role that should be further categorized before presenting characteristics. I think some of the valuable distinction lies in the second words of the titles: data *scientist* and software *engineer*. A scientist can focus more on the craft, research, and technical depth, whereas the engineer might be at a more high-level or more of a holistic level in terms of keeping the product together. It all depends on specializations though, and as such describing the two entirely different or saying that any two data scientists are alike is, of course, very simplified. I like the description with I-shaped, generalist, and T-shaped, and by using them I would describe the data scientist as generally more I-shaped, and software engineers as generally more T-shaped.

Do you think these roles will evolve and specialize further or that "both sides" will need to learn many of the skills of "the other side" and that the roles somehow will merge? Explain your reasoning.

I think specialization is the way forward and also that there are rather more subcategories of the two (as I mention above regarding software engineers) which we will only see more of in the future. To some extent, you will need basic knowledge of the other side in order to get a better understanding (or at least some kind of system perspective). Also, deeper specialization might prove valuable when dealing with AI in order to value the output of AI systems, do quality assurance, etc. which is much more challenging if you have a broader but more shallow skill set.

5 Paper analysis

What About the Data? A Mapping Study on Data Engineering for AI Systems[1]

Core ideas and their SE importance

The core idea of the paper is to do an analysis of 25 previous papers in the domain of AI data engineering, the data engineering part of AI on how to prepare data for AI systems. It looks at the data engineering life cycle and relates it to use in AI settings and studies what parts have been covered in previous work. The focus is on the data and AI engineering lifecycle phases, the proposed technical solutions and architectures, and the lessons learned on AI data engineering. This is of course very important in the engineering of AI systems as there might be new ways to think of data engineering when working with AI systems rather than older/traditional systems.

Relation to your research

This does not at all relate to my research as it does not contain AI. Though AI data engineering will cover more—if not all—domains in the future and as such it might/will become relevant for me in the future. One part of the paper that resonated with me was on "knowledge engineering", presented as understanding and representing human knowledge. Semantic models and similar systems might have an increased value in the future to create structures for AI data engineering and for AI systems to be guided in their reasoning.

Integration into a larger AI-intensive project

There are several parts of this paper that are of relevance to large AI-intensive software projects. To begin with, the paper gives an overview of previous work and can serve as a guide to relevant knowledge. The section with implications for practitioners might be the most relevant, where ideas of big data, data quality, open source tooling, and domain-specific data engineering (to name some) are summarized and referenced, and usable in (large) AI-intensive projects.

Adaptation of your research

As mentioned in relation to my research, I guess I could tweak my research projects towards presenting some kind of semantic model that could help future use of AI in the domain of certificates and web security. This can be some kind of ontology or taxonomy of the relations and components of the web public key infrastructure, including both network of participants and (sub)components of relevant artifacts. This would allow for better AI data engineering in the future.

[1] P. Heck, "What About the Data? A Mapping Study on Data Engineering for AI Systems," *2024 IEEE/ACM 3rd International Conference on AI Engineering – Software Engineering for AI (CAIN)*, Lisbon, Portugal, 2024, pp. 43-52. DOI: 10.1145/3644815.3644954

Towards a Responsible AI Metrics Catalogue: A Collection of Metrics for AI Accountability[2]

Core ideas and their SE importance

This paper is also a systematization of knowledge based on previous work (academic and gray literature), with the aim of presenting a catalog of metrics that are used for AI accountability with special focus on generative AI. The dimensions of metrics presented are within Responsibility, Auditability, and Redressability in AI systems, with sub-criteria within oversight, competence, compliance, and others. Each process metric is paired with several key considerations and linked to product metrics, and each main criterion/dimension is linked to various resource metrics. Needless to say, these factors are highly relevant and usable in the engineering of AI systems in the future.

Relation to your research

This paper relates even less to my research (which did not relate at all). What it could do is relate to the way I conduct my research, given that it specifically targets generative AI that is increasingly being used in research. The results of using generative AI can, as mentioned in the discussion part of the paper, have implications on legal responsibilities and risks, reputational risks, and many other factors. Given this, the use of the presented metrics might prove valuable when including generative AI in both research and any resulting software.

Integration into a larger AI-intensive project

In a large AI-intensive software project, the use of the presented metrics can be highly relevant. Especially as accountability in AI grows in importance, using well-established metrics to prove certain accountability dimensions might be useful in order to communicate to stakeholders how accountability factors are being considered in the project or software. In terms of responsibility, it might be to describe the establishment and role of an AI Governance Committee (process metric 1.1.2) and in auditability it might be Model Provenance (process metric 2.1.2) to help maintain ethical integrity and transparency.

Adaptation of your research

Again, this paper and topic is far off from my research project, but imagining a related research topic with the aim to produce software using AI to investigate and judge compliance and similar factors in the web public key infrastructure, the use of the presented metrics can be useful for users and auditors of the system to help get an understanding of accountability factors in the AI software.

[2] B. Xia, Q. Lu, L. Zhu, S. U. Lee, Y. Liu and Z. Xing, "Towards a Responsible AI Metrics Catalogue: A Collection of Metrics for AI Accountability," *2024 IEEE/ACM 3rd International Conference on AI Engineering – Software Engineering for AI (CAIN)*, Lisbon, Portugal, 2024, pp. 100-111. DOI: 10.1145/3644815.3644959

6 Research Ethics & Synthesis Reflection

Search and screening process

I aimed to find something as recent as possible, starting from the latest iteration (2025) and working my way backward. Preferably, I wanted something related to my research, but even going as broad as looking for *security* as a keyword was unsuccessful. Instead, given the few accepted papers, I manually scrolled through them to find (somewhat) interesting titles that I then screened by reading the abstract and in the end settled for a pair of papers that both were systematization of knowledge/literature review papers distilling previous work to give overview of their topics.

Pitfalls and mitigations

I did not detect any misleading titles/abstracts, even though that might be due to not reading and comparing all of them. It was a very limited number of accepted papers per year, allowing to manually scroll through and read the names instead of being more systematic (as one would have had to be had the number of papers been larger). Given that I found two suitable papers with this method, I did not have to adjust.

Ethical considerations

I have not encountered any ethical issues, neither in how I conducted this work nor in the papers I read as part of this. The first paper did not have/present ethical issues (it was merely a systematization of knowledge based on previous work and clearly referenced/credited). The second paper was of the same style and thus did not present any ethical issues. However, the last paper discussed ethical factors in terms of presenting metrics to help with accountability in AI.