MERCAT: A Metric for the Evaluation and Reconsideration of Certificate Authority Trustworthiness

Michael P. Heinl
University of Ulm
Fraunhofer AISEC

Alexander Giehl
Norbert Wiedermann
Sven Plaga
Fraunhofer AISEC

Frank Kargl
University of Ulm

ABSTRACT

Public key infrastructures (PKIs) build the foundation for secure communication of a vast majority of cloud services. In the recent past, there has been a series of security incidents leading to increasing concern regarding the trust model currently employed by PKIs. One of the key criticisms is the architecture's implicit assumption that certificate authorities (CAs) are trustworthy a priori.

This work proposes a holistic metric to compensate this assumption by a differentiating assessment of a CA's individual trustworthiness based on objective criteria. The metric utilizes a wide range of technical and non-technical factors derived from existing policies, technical guidelines, and research. It consists of self-contained submetrics allowing the simple extension of the existing set of criteria. The focus is thereby on aspects which can be assessed by employing practically applicable methods of independent data collection.

The metric is meant to help organizations, individuals, and service providers deciding which CAs to trust or distrust. For this, the modularized submetrics are clustered into coherent submetric groups covering a CA's different properties and responsibilities. By applying individually chosen weightings to these submetric groups, the metric's outcomes can be adapted to tailored protection requirements according to an exemplifying attacker model.

CCS CONCEPTS

• **General and reference** → **Metrics**; • **Security and privacy** → **Authentication**; *Trust frameworks*; *Security protocols*; • **Social and professional topics** → *Computer crime*; *Surveillance*.

KEYWORDS

Cloud Security, Metric, CA, Trustworthiness Assessment, PKI, Digital Certificate, X.509

ACM Reference Format:
Michael P. Heinl, Alexander Giehl, Norbert Wiedermann, Sven Plaga, and Frank Kargl. 2019. MERCAT: A Metric for the Evaluation and Reconsideration of Certificate Authority Trustworthiness. In *2019 Cloud Computing Security Workshop (CCSW '19), November 11, 2019, London, UK*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3338466.3358917

CCSW '19, November 11, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6826-1/19/11…$15.00
https://doi.org/10.1145/3338466.3358917

1 INTRODUCTION

A series of incidents related to certificate authorities (CAs) casted doubt over the current process of trust establishment within the public key infrastructure (PKI) used for the World Wide Web (WebPKI) which secures communication of almost all modern-day cloud environments.

One of the most prominent of those cases was the breach of the Dutch CA *DigiNotar* [38] which had been taking place in 2011. The first publicly noticeable consequence was the issuance of a wildcard certificate for the domain name (DN) `google.com` which was then used for a man-in-the-middle (MITM) attack intercepting the communication with Google services [1]. Eventually, a list of over 500 DNs for which certificates illegitimately had been issued was compiled, including DNs of Skype, Yahoo!, and The Onion Router (TOR). Consequently, DigiNotar's root certificate was completely removed from all major browsers' trust stores [1, 47, 51].

Other examples of notable incidents were related to computer manufacturers Dell [33] and Lenovo [34, 36] shipping systems with preinstalled root certificates and the corresponding private key; the Turkish CA *Türktrust* [41] issuing unrestricted intermediate CA certificates to government-related CAs which were in turn used to issue a certificate for `*.google.com`; and Symantec [10, 24, 52, 76] regularly violating standards by malpractices such as issuing certificates using 1,024-bit keys or SHA-1 signatures expiring after their actual end-of-life dates, issuing extended validation (EV) certificates without undergoing the corresponding audit procedure, and issuing illegitimate test certificates. These incidents raise the question about the fundamental problems of the WebPKI in order to be able to systematically develop solutions.

First of all, there is the problem that the concept of PKI heavily relies on *trust*, which is *"bad for security"* [32]. The dependence on trust is caused by the necessity of a third party confirming an entity's identity linked to a specific public key. This confirmation of identities is the very task of a CA. However, the basic architecture of the WebPKI lacks both transparency and properly integrated mechanisms to control that the CA has really implemented all measures necessary to accomplish this task reliably. Even if relevant information about the CA's behaviour is publicly available, it is not trivial to understand and assess it.

The outcome of this situation is an *information asymmetry* which means that the user is not able to distinguish between CAs with good and the ones with bad practices. According to Backhouse et al. [8], this situation leads to a market failure because one of the key attributes of a properly functioning market is transparency. This market failure allows certain CAs to dump prices at the expense of security.

Due to the *single point of failure* originating in the fact that every trusted CA can issue certificates for any website, these low-security CAs do not only harm their own customers and their corresponding users but the WebPKI as a whole. This property is also referred to as *weakest link problem* [6].

Although actually supposed to provide an additional layer of assurance, the effectiveness of audits should also be taken with a grain of salt. Auditors are supposed to be independent but are chosen and hired by the CA and have a genuine interest of being hired again. Therefore, the auditing process establishes a sense of certainty by partially transferring trust from the CA to the auditor who, at least to some degree, financially relies on the CA.

In order to mitigate the discussed problems, this paper presents the conceptual development and evaluation of a metric to assess the trustworthiness of CAs.

2 RELATED WORK

Such a metric is especially valuable in the context of former research done in the area of user-centric trust store minimization [14, 53] which revealed that only a small number of the root certificates included in major web browsers' trust stores are actually used.

There are approaches to assess certificates by making use of information regarding the issuing CA such as results of a distributed reputation mechanism [20]; propositional logic and probability theory using so-called *trustworthiness terms* [61]; as well as certificate policies [67]. Kumar et al. [40] monitored certificates issued by a wide range of CAs and analyzed their compliance with RFC 5280, the CA/Browser Forum's Baseline Requirements (BRs), and community best practices. Fadai et al. [30] analyze the issuing institutions as well as the countries of origin of root certificates shipped with the currently most popular operating systems and browsers and correlate these findings against different indices indicating the CAs' origin states' constitutionality [3, 31, 55, 66].

Wazan et al. [69–72] emphasize that the assessment of a CA's trustworthiness is a task too complex for the ordinary user and therefore successfully proposed to introduce a new role which they call the *technical and legal expert* [70] or *trust broker* (TB) [69] into the 2016 version of the X.509 standard [39, 72] as an optional attribute.

The *Lemons Principle* [2], a theory which says that if the quality of products (both goods and services) in a specific market cannot be similarly assessed by both buyers and sellers, the low quality products tend to supersede the high quality ones, has been applied to the CA market by Backhouse et al. [8]. They emphasize that the variety of technologies, procedures, and legal frameworks in combination with the common users' insufficient expertise leads to opportunistic behaviour of CAs who tend to prefer cost efficiency over quality in order to maximize profits. To allow users to spot CAs of high quality despite this opportunistic behaviour, Backhouse et al. suggest applying Akerlof's countermeasures to the CA market:

(1) *Guarantees:* Users should analyze the CAs' published policies in order to evaluate the liability granted by the CA.
(2) *Brand names:* The authors state that companies which are already established and are now also operating a CA under an already well-known brand are less likely to risk the reputation of this brand by embracing malpractices.
(3) *Licensing*: Obligatory licensing required by law (e.g., for qualified digital signatures) and facultative accreditation (e.g., WebTrust audit) ensure at least that minimal quality standards have been met.

Countermeasure One lacks the understanding that, taking into account existing research on the privacy policies [19, 46], it can be assumed that neither relying parties (RPs) nor many certificate applicants actually read the corresponding policies. At first glance, *Countermeasure Two* seems to be logically correct. The Symantec case initially described is empirical evidence against this argument, though. Brand names basically only provide declining public reputation as a sanction. The disclosure of malpractice in the CA market, however, is far from being breaking news. Hence, it is questionable if the necessary information about breaches would even reach an audience large enough to cause a noticeable decline of public reputation. As long as there is no dedicated reputation system, it is also questionable if this is a rational or rather an emotional mechanism and therefore even effective. Specter [63] compares the numbers of certificates sold by CAs before and after they had experienced a security breach. He concludes that the number of sold certificates indeed slightly decreases after a breach but doubts that this has a corrective effect on a CA's behaviour because the losses are negligible. *Countermeasure Three*, licensing and accreditation, on the other hand does allow targeted sanctions by corresponding bodies and is therefore a rationally controllable instrument. Coming back to the example of Symantec, it was not the public reputation which eventually persuaded Symantec to sell its complete CA business but rather Google's and Mozilla's strict line of action.

Summarized, this economical analysis confirms the need for strong, fact-based rather than trust-based assessment methods of CAs' trustworthiness. Furthermore, the results of these independent assessments should be as easy to understand as possible in order to enable both RPs and certificate applicants to make informed decisions.

3 METRIC METHODOLOGY

Each of the major web browser vendors maintains its own root inclusion program [5, 35, 48, 50] which in turn refers to policies [27–29, 73, 74] and guidelines [16–18] as illustrated in Figure 1. During the course of the project, the different documents were analyzed along with relevant research literature in order to extract possible criteria of trustworthiness (CoT) which is any information useful to independently evaluate a CA's trustworthiness.

Following the process illustrated in Figure 2, the goal of the present and the next section is to incorporate these CoT into one holistic metric which can be used to measure the trustworthiness of CAs.

3.1 Introducing Obtainability

A crucial prerequisite for the successful transfer of CoT into submetrics and their eventual applicability is the availability of information indicating a CA's performance regarding the respective submetric. In order to avoid confusion with the security objective of *availability*, this property is referred to as *obtainability* hereinafter.

The categorization of CoT into different levels of obtainability as presented in Table 1 helps in two ways. Initially, during the

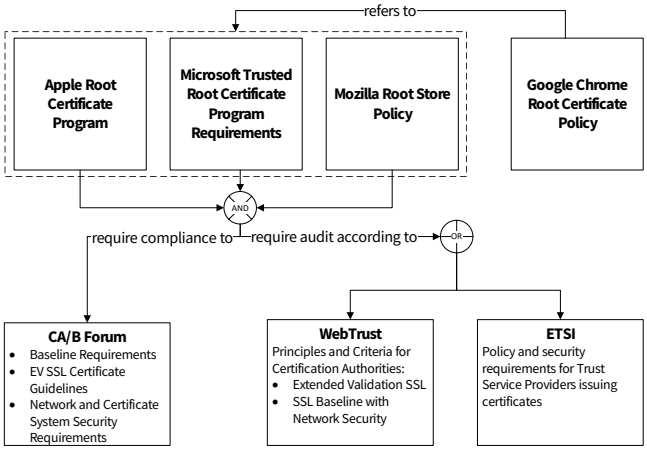


Figure 1: Relations of main policies concerning root certificate inclusion.

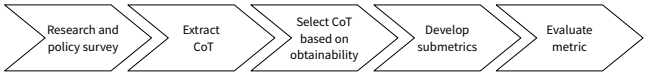


Figure 2: Overview of the metric development process.

process of metric development, it allows eliminating CoT which cannot be obtained at all. In the later stage of applying the metric's assessment methods to actual CAs, the obtainability levels allow to chose specific submetrics according to the scheduled effort in terms of provided budget and workforce.

For example, if there is only initial funding and negligible human supervision, submetrics with the obtainability level *automatical* (A) are predestinated to develop fully automated collection and assessment methods. If there is sufficient workforce for manual evaluation and funding to buy certificates, submetrics with the levels *public* (P), *requestable* (R), and *buyable* (B) can also be taken into consideration for the corresponding program. Although there emerged CAs providing certificates for free, the majority of CAs still takes money for issuing domain validation (DV) and certainly for issuing organizational (OV) and extended validation (EV) certificates. Hence, the obtainability level *B* reflects this situation despite the fact that there are exceptions such as *Let's Encrypt*.

3.2 Selection of CoT

Table 2 summarizes the CoT gathered during the policy and research survey and maps them to their corresponding level of obtainability. Based on this mapping, appropriate CoT are selected to be part of the actual metric to be developed. CoT up to the obtainability level of *B* are considered appropriate while CoT mapped to *non-verifiable* (N) or *unobtainable* (U) are neglected.

As it can be seen in Table 2, it is generally impossible to gather empirical evidence about the CA's internal security controls by an external trustworthiness assessment without trusting the CA itself or the auditor who financially depends on the CA, at least to a certain degree. As initially stated, trust is one of the key problems of the current PKI. Relying on this kind of information would therefore be rather irrational in the context of this paper, although the criteria grouped in this category are without any doubt very important for the CA's overall trustworthiness.

3.3 Submetric Development Template

In order to have an easy to understand and comparable structure of the different submetrics, they are verbalized according to the template presented in Table 3. Besides its name and submetric group (SMG), each submetric contains of a brief description of the corresponding criterion, its obtainability level, and the method to gather necessary information. The heart of each submetric is its *score values table* (SVT) which describes how the actual measurements translate into the score indicating how well the submetric's requirements are met. For each measurement, a resulting score represented as unit interval with 0 as the worst and 1 as the best score is assigned.[1] This score represents the CA's level of compliance with the corresponding submetric's requirements according to its SVT.

3.4 Attacker Model

Although highly desirable, it is hard to develop *"one-size-fits it all"* solutions serving all kinds of users equally. This principle also applies to the development of a trustworthiness metric. Hence, it is important to provide some kind of adaptability mechanism to be able to satisfy a diverse audience nevertheless. In the present metric, this is realized by applying different weightings to specific SMGs depending on their effectiveness regarding the chosen attacker model. These weightings are *low* (L), *medium* (M), and *high* (H). The corresponding numeral weighting factors are 1, 2, and 3. The employed methodologies [57, 62] are deliberately designed simple and understandable so that users can derive their individual protection profile from a generic attacker model presented in this section. The set of attackers and corresponding example scenarios related to the trustworthiness of CAs includes but is not limited to:

- **State Actor:** State actors are able to approach victims by sophisticated methods including Border Gateway Protocol (BGP) hijacking or advanced persistent threats (APTs) for purposes such as mass surveillance, industrial espionage, or repression.
- **Organized Internet Criminal:** Instead of massive technical capabilities and manpower, the toolset of Internet criminals includes attacks like medium to large-sized phishing campaigns trying to exploit human weaknesses.
- **Script Kiddie:** With today's sheer number of publicly available hacking tools, the threshold for opportunistic computer crime is rather low. Local attacks which can be performed by virtually everyone include for example Address Resolution Protocol (ARP) poisoning.

A reason for providing a flexible attacker model instead of a fixed weighting scheme is that the attributes of a real attacker are hardly generalizable. It is also impossible to take all possibly relevant circumstances into account. As a result, the following paragraphs describe the presented model's underlying assumptions. Those assumptions serve two purposes. First, they help to understand the

[1] A *unit interval* is a value $x \in \mathbb{R}$ with $0 \leq x \leq 1$.

Table 1: Levels of obtainability.

Name	Description	Example	ID
Automatical	Publicly available, fully structured and therefore automatically obtainable information.	Standardized field of X.509 certificate.	A
Public	Publicly available but in some sort unstructured information.	Information in the CP/CPS which can be empirically verified.	P
Requestable	Information not publicly available but requestable for free.	Non-public information about a CA's ownership structure which has to be requested from registers.	R
Buyable	Information gathering includes financial expenditure.	Discrepancies between CPS and effectively implemented validation process.	B
Non-verifiable	Empirically unavailable information.	CP/CPS information which cannot be verified without trusting the CA or auditors.	N
Unobtainable	Generally not available information.	Trade secrets or a single person's clandestine intentions.	U

Table 2: Mapping of CoT to obtainability level.

Category	CoT	Obtainability
Revocation	UTD of CRLs	A
	UTD of OCSP information	A
	Consistency of CRL and OCSP	A
	Response for non-existent Certificates	A
	Revocation Request Reaction Time	B
	Receptiveness of Revocation	P
CA Restrictions	Path Length Constraints	A
	Name Constraints	A
	Key Usage	A
Certificate Issuance	CAA	B
	High-Risk CSR	B
	Mixed Character Set IDN	B
	Origin Country	B
	Extended Key Usage	A
	Wildcard Certificates	A
Cryptography	Key Size	A
	Digest Algorithm	A
	Public Key Reuse	B
	Weak Key	B
Independence	Operational Independence	P/R
	Legal Independence	P
Transparency	Certificate Transparency	A
	Document Repository	P
	Legal Transparency Report	P
Internal CA Security	Annual Risk Assessment	N
	Regular Vulnerability Scanning	N
	Regular Penetration Tests	N
	Multi-Factor Authentication	N
	Use of FIPS-compliant HSM	N
	Network Segmentation	N
	Network Boundary Control	N
	Hardening of Systems	N
	Appropriate Password Policies	N
	Patch Management	N
	Personnel: Background Checks	N
	Personnel: Appropriate Training	N
	Personnel: Freedom from Pressure	U

attacker model presented in Table 4. Second, they enable a user to individually adjust the presented example model.

Despite the incidents discussed in Section 1, the discovery of rogue certificates suspected to be used by state actors is rather rare compared to the occurrence of regular Internet crime. Thus, the CAs' revocation services have to protect RPs particularly against once legitimately issued but now exposed certificates used for usual

Table 3: Information contained in the submetric development template.

Name	Description
SMG Name	Name of the SMG the submetric belongs to.
Submetric Name	Name of the respective submetric.
Description	A brief description of the submetric and why it is important.
Collection Method	Describing how performance indicators can be collected.
SVT	Describing the different possible outcomes of the assessment and corresponding values used in the scoring process.

Table 4: Attacker model mapping attackers to SMGs' effectiveness levels.

	State Actor	Internet Crime	Script Kiddie
Revocation SMG	L	H	M
Restriction SMG	H	M	M
Issuance SMG	H	H	H
Cryptography SMG	H	H	H
Independence SMG	H	M	L
Transparency SMG	H	M	L

Internet crime. In the case of disclosure of a major CA breach, the browser vendors have proven to be agile [1, 47, 51]. Opposed to states, the probability of criminals or script kiddies owning a CA is rather low so that the restriction measures cannot directly applied to them. The overall containment of unrestricted CAs generally reduces the risk for wrongly issued certificates, though. Even without an own CA, there are ways to trick a commercial CA to issue certificates which can be used for fraudulent activities. Therefore, CAs have to be prepared to handle potentially malicious certificate signing requests (CSRs) no matter who the potential attacker is that raised the request. The same applies to the cryptographic properties of issued certificates because it is the critical factor securing

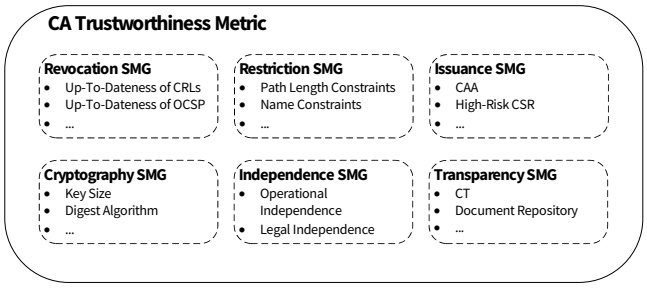


Figure 3: The metric's basic architecture.

the RP's data. The probability of script kiddies compromising or influencing a CA is rather low, the case of DigiNotar shows that this kind of attack more likely involves state actors. Hence, the independence and transparency metrics are especially important as a protection against state actors. However, the risk of criminals blackmailing CA employees is still not completely negligible.

3.5 Metric Calculation

All scores are consistently composed of unit intervals with 0 as the worst and 1 as the best score. A single SMG's averaged score a is calculated with

$$a = \frac{1}{|M|} \sum_{m \in M} m, \tag{1}$$

where M is a set of unit intervals representing the single scores of the submetrics contained in the SMG. Calculating the unweighted overall trustworthiness score t is described with

$$t = \frac{1}{|A|} \sum_{a \in A} a, \tag{2}$$

where A is a set of unit intervals representing the single SMGs' averaged scores. In order to calculate the weighted overall trustworthiness score w, the corresponding weightings have to be applied with

$$w = \frac{1}{\sum_{g \in G} g} \sum_{i=1}^{|A|} g_i a_i, \tag{3}$$

where A is a set of unit intervals representing the single SMGs' averaged scores and G is a set of unsigned integers representing the corresponding SMGs' weightings determined by applying one of the example attacker models described in Section 3.4 with ($a_i \in A$, $g_i \in G$) and $|A| = |G|$.

4 METRIC DEVELOPMENT

The next step is to apply the proposed methodology to the selected CoT. For this, the diverse factors are grouped into SMGs and formalized according to the submetric development template introduced in Section 3.3. The relation between submetrics, SMGs, and the overall metric is visualized by Figure 3. The individual SMGs are not directly weighted. Rather, they subsume closely related CoT to a scoring system designed to be as intuitive and tangible as possible. Therefore, the process of weighting is meant to be performed by individually applying the already introduced attacker model.

The metric consists of a total of six SMGs but can be easily extended due to its modular approach. The following paragraph briefly describes each SMG and the nature of the corresponding submetrics it contains. Table 6 (Appendix) provides detailed information about each submetric, e.g. obtainability level and mode of score calculation. For a comprehensive description of how to collect the needed data, mostly making use of *Censys* [25], the Appendix of this paper's underlying thesis can be referred to [37].

The *Revocation SMG* assesses how a CA handles different aspects of certificate revocation. It takes into account the up-to-dateness (UTD) and consistency of certificate revocation lists (CRLs) and Online Certificate Status Protocol (OCSP) information, how requests concerning non-existent certificates are handled, and the reaction time as well as receptiveness of revocation requests. The *Restriction SMG* evaluates which measures a CA employs to restrict certificates issued to subordinate CAs taking into account Path Length and Name Constraints as well as `keyUsage`. The *Issuance SMG* takes technical validation mechanisms into consideration which have the purpose to detect potentially fraudulent activity during the process of end-entity certificate issuance. Mechanisms to be analyzed include Certification Authority Authorization (CAA), high-risk CSRs employing DN of well-known websites, mixed character set internationalized domain names (IDNs), a CSR's origin country, as well as the handling of Extended Key Usage (EKU) and wildcard certificates. The *Cryptography SMG* evaluates cryptographic factors of the certificates issued by the assessed CA such as the size of the cryptographic key, the used digest algorithm, and whether weak (in terms of entropy) or duplicated keys are used. Apart from the aforementioned, rather technical SMGs, the *Independence SMG* relies on factors including the CA's operational and legal independence. The *Transparency SMG* assesses how much effort the CA puts into providing insight about its operations to the general public. For this, the usage of Certificate Transparency (CT) [42, 43], the maintenance of a document repository providing up-to-date versions of the certificate policy (CP), certification practice statement (CPS), and audit reports, as well as the publication of legal transparency reports [77] are taken into account.

5 EVALUATION

The metric's evaluation focuses on CAs offering free DV trial certificates in order to circumvent larger financial expenditures. All CAs are tested using a set of three second-level DNs and corresponding subdomains. For the sake of privacy, they are all substituted by the DN `uni-ulm.test` for the documentation of the evaluation. These free trial certificates come with constraints. For example, they often only allow a very small number of CSRs per second-level DN including all possible subdomains. Testing more than one submetric in one CSR is not possible because most CAs lack meaningful feedback about the reason a CSR has been rejected. Some CAs do not even clearly state that a CSR has been rejected. Instead, they just suddenly stop the process stating for example *pending validation* without any further information or progress.

Therefore, the initially planned approach of testing as many submetrics as possible with only one CSR, has to be changed in a way that each CSR enquires only one submetric at a time. This allows drawing a conclusion even if no direct feedback is given.

Table 5: Prices of DV certificates/year (in U.S. dollars).

Let's Encrypt	SSL.com	Comodo	Thawte
0.00	49.00	99.95	149.00

This new approach also requires to control a sufficient number of DNs, though. Where possible, the submetrics were applied in accordance with their respective *Data Collection Method* field. Special conditions and subsequent adaptations of the present assessment are described in the following summary.

(1) **High-Risk CSR:** In order to assess whether CSRs are checked for high-risk components, a certificate is requested for the DN `paypal.com.uni-ulm.test` . That is because the term *paypal* in the DN of a DV certificate is likely to indicate a malicious intention.[2]
(2) **Mixed Character Set DN:** The above mentioned DN can be visually imitated by employing the following Unicode characters of the Cyrillic alphabet:
   - a: U+0430
   - y: U+0443
   - c: U+0441
   - o: U+043E

   This results in the following Punycode [21] representation: `xn--yl-6kcb1fc.xn--m-0tbi.uni-ulm.test`.
(3) **Origin Country:** The CSR is sent using a VPN tunnel exiting in the United States.
(4) **Weak Key:** A weak key generated with Debian 4.0r2 is used to generate the CSR.
(5) **Revocation Request Reaction Time:** According to online investigations, some CAs offer a control center which allows the user to revoke her certificates on her own. Most trial certificates tested during this evaluation lack this functionality, though. Thus, the corresponding CAs are approached using their favourite mode of communication including an online ticketing system, the tool *certbot*, and regular email.[3]
(6) **Consistency of CRL & OCSP:** Some CAs do not offer either CRL or OCSP. Hence, this submetric is not applicable to them.

Submetrics not mentioned in the above list are rather uncritical since they employ methods of data collection not requiring direct interaction with the CA which means making use of *Censys* in the majority of cases [25].

The CAs assessed during the evaluation are *Comodo*, *SSL.com*, *Let's Encrypt*, and *Thawte*. Furthermore, *PositiveSSL*, *GeoTrust*, and *RapidSSL* were part of the initial set of CAs to be assessed because they also offer free trial certificates. However, certificates requested from PositiveSSL are in fact issued by Comodo and also bear Comodo's name.[4] In contrast to the other CAs, GeoTrust and RapidSSL allow exactly one CSR per second-level DN no matter if the CSR is rejected or not. This fact makes it impossible to properly assess those CAs with the given set of only three DNs. Hence, the mentioned CAs are neglected during the further evaluation.

[2]https://www.thesslstore.com/blog/lets-encrypt-phishing/
[3]https://certbot.eff.org/
[4]PositiveSSL's certificates' organization attribute is set to `O=COMODO CA Limited`.

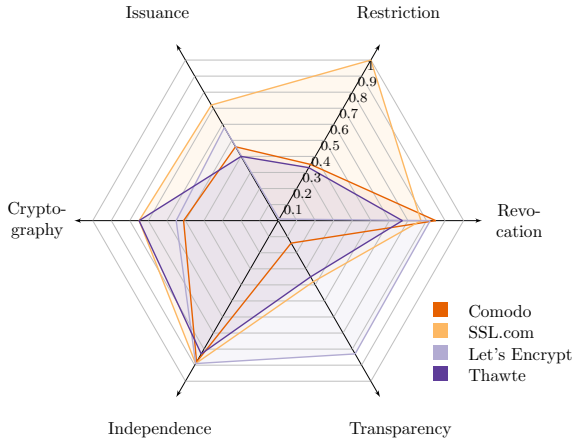


Figure 4: Comparison of the evaluation's SMG results.

5.1 Hypotheses

As shown in Table 5, there is a difference of approximately 50 U.S. dollars between each of the DV certificates offered by the different assessed CAs for a validity period of one year as of October 27, 2018.[5] This results in an overall span of 149 U.S. dollars between the cheapest and the most expensive CA. It was already discussed that customers of CAs cannot rely on the usual market mechanisms [8]. This assumption can be confirmed by falsifying the subsequent hypothesis which follows the classical market theory:

Hypothesis 1. *The higher the price of a certificate, the trustworthier the corresponding CA.*

A different aspect directs to another interesting hypothesis concerning CAs and their individual compliance with prevalent regulations presented by Kumar et al. [40]. The authors provide a comprehensive overview of CAs listing their respective certificates' error rates. All the CAs chosen to participate in the present evaluation are listed in this paper as well.

Hypothesis 2. *The assessed CAs score according to their adherence of regulations which is diametrically opposed to the error rates presented by Kumar et al. [40].*

5.2 Results

Table 7 (Appendix) shows the results of the measurements undertaken in order to calculate scores for each submetric. Those scores can be found in Table 8 (Appendix). Utilizing Figure 4, the Restriction and Transparency SMGs can be identified as the main problem areas. The *Restriction SMG's* scores especially suffer from the very low to zero percentage of subordinate CA certificates being limited by path length and name constraints. SSL.com's excellent score in this SMG is due to the fact that it does not issue CA certificates to third parties at all. On the one hand, this distorts the scoring because one could argue that SSL.com's handling of subordinate CA certificates would not be better than the others' if it issued any.

[5]https://www.ssl.com/certificates/basicssl/buy
https://www.comodoca.com/en-us/solutions/tls-ssl-certificates/domain-validated-(dv)-ssl/
https://www.thawte.com/ssl/ssl123-ssl-certificates/

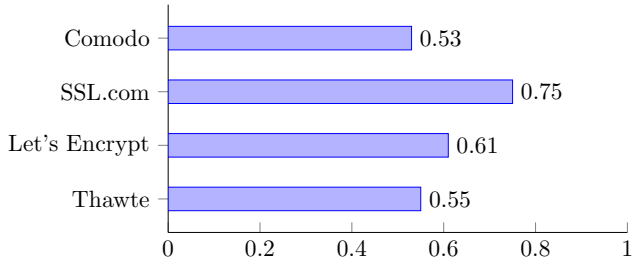


Figure 5: Overview of the evaluation's overall trustworthiness score results.

On the other hand, this practice also proves that it is possible to successfully operate a commercial CA without issuing subordinate CA certificates to third parties which means a plus of security for the whole WebPKI. The *Transparency SMG's* weak points are the limited deployment of CT and the absence of legal transparency reports. The only CA providing legal transparency reports is Let's Encrypt. It also has the highest percentage of CT deployment with only one certificate not being logged on CT servers.

There are also results on the submetric level of other SMGs which are worth discussing. The CAA Submetric's results, for example, confirm an observation already reported by Scheitle et al. [60]. Although consistently rejecting CSRs with a conflicting `issue` property tag in the CAA RR, none of the assessed CAs makes use of `iodef` notifications in the case of rejection. Another finding is that none of the assessed CAs proactively classifies the keyword *paypal* as an indicator for phishing sites, neither in ASCII nor in IDN format. While most CAs offer guidance to report certificates which are in fact used for fraudulent websites at least reactively, Let's Encrypt states that such certificates should be reported to Google's Safe Browsing list rather than to CAs. This argument makes perfect sense because Safe Browsing is not only utilized by multiple CAs but also by Google's search engine and is therefore a much more central instance than single CAs can be. However, shifting from *high-risk CSRs*, as expressed in the BRs, to *high-value domains* [45] is not only a change in phrasing. It also implies that the CAs' responsibility is not to care about potentially fraudulent terminology in DNs but only to look after the highly coveted DNs such as `google.com` or `paypal.com`. Let's Encrypt's rationale for this shift is that malicious actors would just look for the weakest CA to get their certificate.[6] The focus should thus be on malware and phishing protection features instead of DN sanitization during the process of issuance. Although this is also a good point, there are two arguments against it. First, it is never bad to have an additional line of defense, something a caring CA would definitely constitute. Second, instead of adapting to the standards of less trustworthy CAs, the community should continue to raise the overall level of security by defining stricter requirements and consequently eliminating CAs not adhering to it.

[6]Malicious actors could obviously also refrain from using certificates. However, users tend to feel safer when a website is using HTTPS despite the fact that this doesn't reveal anything about the website's actual content [4]. Furthermore, Google started to mark websites which only offer the unencrypted Hypertext Transfer Protocol (HTTP) as *not secure* with version 68 of Chrome [59].

Comparing the prices listed in Table 5 with the overall trustworthiness scores presented in Figure 5 reveals that *Hypothesis 1* which assumes a positive correlation of price and trustworthiness is indeed falsified by the results of the evaluation. Although the order is not perfectly reversed, the two more expensive CAs, namely Comodo and Thawte, constitute the lower scored half. On the contrary, SSL.com and Let's Encrypt which represent a low-cost or even free alternative are placed at the higher end of the rating scale.

Hypothesis 2 assumes a negative correlation of error rates and trustworthiness. Figure 6a shows the evaluation's overall trustworthiness scores in relation to the error rates of the certificates issued by the respective CA according to Kumar et al. [40]. Indeed, a negative correlation between those values can be identified for the assessed CAs. Again, it is not a linear but a clustered distribution. One explanation for the interrelationship between error rate and trustworthiness is of course that both measures are based at least partially on the BRs.

A surprising finding is, that the identified clusters have exactly the same order as the one observed for Hypothesis 1. That is, there is a direct, strictly monotonically increasing alignment of price and error rate for the set of assessed CAs as visualized in Figure 6b.

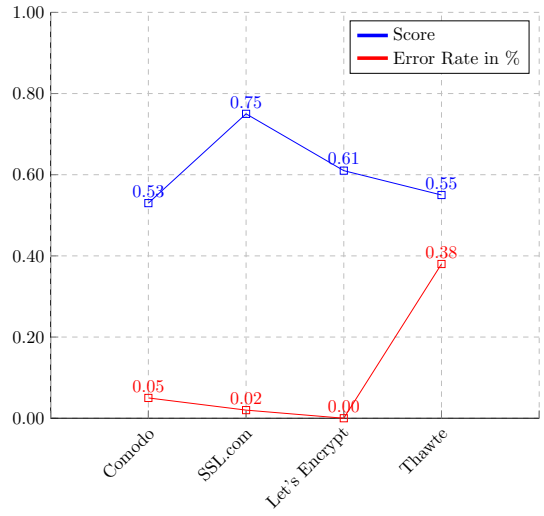
Although there are still a lot of potential improvements, the evaluation and the subsequent analysis of the results indicate that the metric can be practically applied as initially intended.

6 CONCLUSION

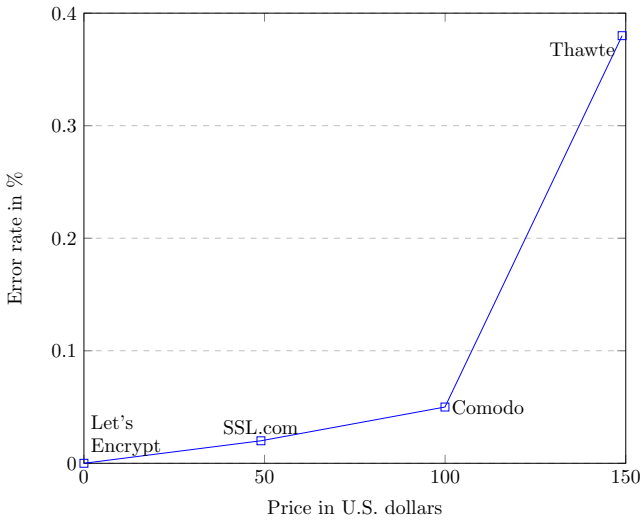
Since its existence, the Internet has been an enabler for many disruptive technologies and business ideas. It is, however, not known for rapidly deploying new technologies as part of its own infrastructure. Examples are IP version 6 or Domain Name System Security Extensions (DNSSEC). Although proposed since over 20 years [26], the deployment of DNSSEC can still be considered to be in its early stages [65, 68]. Bearing this and efforts to further deploy HTTPS in mind [59], it is questionable if the WebPKI is going to be replaced by more promising, clean-slate architectures such as SCION in the near future [11, 54]. This assumption leads to the question how to handle the risks which come along with the flawed WebPKI architecture as is?

A mid-term solution which can be employed by organizations or particularly exposed divisions could consist of two steps. First, quantitatively reducing the root certificates shipped with operating systems and browsers according to a user's or a group's web surfing behaviour [14, 53]. This should lead to a drastic minimization of the attack surface with only minor losses of comfort. As a second step, the trustworthiness of the relatively small number of remaining CAs can be checked according to the methodology presented in this paper. CAs falling below a certain level of trustworthiness can be eliminated as well. After the initial assessment of the remaining CAs, the definition of the required level of trustworthiness can thereby be readjusted based on a reasonable trade-off between the security requirements of affected systems and user comfort. To ensure the best possible backing, all relevant stakeholder groups should be involved in this discussion.

One reasonable long-term solution is to eliminate certificates issued by untrustworthy CAs not only from trust stores locally but



(a) trustworthiness scores.



(b) prices.

Figure 6: Relation of certificates' error rates [40] and their respective...

globally. This requires the cooperation of browser vendors, website operators, and RPs. While browser vendors can decide whether to retain certificates or dump them from their trust stores, they also have to keep the balance between security and their users' browsing experience. This means that CAs serving many frequented websites are virtually *"too big to fail"* [6] even if their level of trustworthiness is questionable. The higher the relying parties' awareness for the importance of trustworthy CAs, the greater the genuine interest of website operators to use a certificate issued by one of these CAs. Hence, the competency of the TB should be extended to not only serve relying parties but also to provide holistic advice to website operators to determine which CA to commission.

6.1 Contributions

This paper presents a holistic metric to assess the trustworthiness of CAs taking into account different technical, economical, political, and legal considerations. The proposed approach is modular, easily extensible, and adaptable in two ways. First, regarding a potential attacker's capabilities by weighting factors according to the attacker model. Second, in terms of own resources by adjusting the used submetrics according to the affordable level of obtainability. In contrast to other proposals, the present metric is completely independent from unverifiable information published in the CP/CPS or the audit report. This is especially important because of the auditor's conflict of interest described in Section 1. Eventually, a conclusive evaluation assessing a set of CAs offering free DV certificates ensures the practical applicability of the metric.

Apart from organizations and individuals with special protection requirements, the metric can also be used as a tool for the TB. Complementing the methods proposed by Wazan et al. [72] which rely on fragmented information provided by third parties such as the CA's competitors or users, this metric provides a tool to directly assess a wide range of factors influencing trustworthiness of a specific CA without establishing additional trust dependencies.

6.2 Future Work

In order to further increase the metric's adaptability, the system of weighting can be expanded from submetric groups to the level of individual submetrics. Additionally, submetrics can be made more granular and accurate. For example, the *Legal Independence* Submetric could be enhanced by further investigating whether there is relevant legislation allowing lawful interception in specific countries and how well the established legal mechanisms protect citizens from deliberate abuse. Another illustrative improvement regarding the *Name Constraints* Submetric could be the deeper inspection of the name constraints actually set. Meaning not only assessing if and how often name constraints are applied but also taking into account the dimension of the imposed restrictions. It is, for example, a huge difference whether the `permittedSubtrees` field is set to `.de` or to `.uni-ulm.de`. Based on new research, there is also the possibility to create new submetrics complementing the already existing ones [13]. Furthermore, methods utilizing natural language processing and machine learning would be beneficial to gather and analyze CoT which are publicly available but not machine-readable yet.

Besides the mentioned directions how to further develop the metric, there are also other findings not directly affecting the metric but the overall topic of *CA trustworthiness*. For example, none of the CAs assessed during the course of the evaluation checked the public key submitted in the CSR against the public key of certificates previously issued by the same CA although they could even utilize publicly available sources such as CT log servers for global visibility. Further research regarding the trustworthiness of CAs would also benefit from the standardization of machine-readable CP/CPS. According to other proposals [67, 72], this standardization can make use of XML or JSON. Another proposed idea, the TB, has already made the transition from academic research into one of the most important standards in the field of information technology, namely X.509 [39]. However, there is little information about its

practical relevance yet. Hence, it is time for a study on the current state of deployment and how this role is perceived by different groups of stakeholder including RPs, CAs, industry, and security researchers.

REFERENCES

[1] Heather Adkins. 2011. An update on attempted man-in-the-middle attacks. https://security.googleblog.com/2011/08/update-on-attempted-man-in-middle.html

[2] George A. Akerlof. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics* 84, 3 (Aug 1970), 488–500.

[3] Amnesty International. 2014. Death Sentences and Executions 2013. https://www.amnestyusa.org/wp-content/uploads/2017/04/act500012014en.pdf

[4] Anti-Phishing Working Group. 2018. *Phishing Activity Trends Report - 3rd Quarter 2017.* Technical Report.

[5] Apple Inc. 2016. Apple Root Certificate Program. http://www.apple.com/certificateauthority/ca_program.html

[6] Axel Arnbak, Hadi Asghari, Michel Van Eeten, and Nico Van Eijk. 2014. Security Collapse in the HTTPS Market. *Commun. ACM* 57, 10 (Sept. 2014), 47–55. https://doi.org/10.1145/2660574

[7] Nimrod Aviram, Sebastian Schinzel, Juraj Somorovsky, Nadia Heninger, Maik Dankel, Jens Steube, Luke Valenta, David Adrian, J. Alex Halderman, Viktor Dukhovni, Emilia Käsper, Shaanan Cohney, Susanne Engels, Christof Paar, and Yuval Shavitt. 2016. DROWN: Breaking TLS Using SSLv2. In *25th USENIX Security Symposium (USENIX Security 16)*. USENIX Association, Austin, TX, 689–706. https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/aviram

[8] James Backhouse, Carol Hsu, John Baptista, and Jimmy C. Tseng. 2003. The Key to Trust? Signalling Quality in the PKI Market. In *Proceedings of the 11th European Conference on Information Systems, ECIS (2003)*.

[9] Elaine Barker. 2016. *Recommendation for Key Management Part 1: General.* Draft NIST Special Publication 800-57 Part 1 Revision 4. National Institute of Standards and Technology (NIST).

[10] Elaine Barker and Allen Roginsky. 2015. *Transitions: Recommendation for Transitioning the Use of Cryptographic Algorithms and Key Lengths.* NIST Special Publication 800-131A Revision 1. National Institute of Standards and Technology (NIST).

[11] David Barrera, Laurent Chuat, Adrian Perrig, Raphael M. Reischuk, and Pawel Szalachowski. 2017. The SCION Internet Architecture. *Commun. ACM* 60, 6 (May 2017), 56–65. https://doi.org/10.1145/3085591

[12] Colin Boyd, Cas Cremers, Michèle Feltz, Kenneth G. Paterson, Bertram Poettering, and Douglas Stebila. 2013. ASICS: Authenticated Key Exchange Security Incorporating Certification Systems. In *Computer Security – ESORICS 2013*, Jason Crampton, Sushil Jajodia, and Keith Mayes (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 381–399.

[13] Markus Brandt, Tianxiang Dai, Amit Klein, Haya Shulman, and Michael Waidner. 2018. Domain Validation ++ For MitM-Resilient PKI. In *Proceedings of the ACM Conference on Computer and Communications Security (CCS)*.

[14] Johannes Braun and Gregor Rynkowski. 2013. The Potential of an Individualized Set of Trusted CAs: Defending Against CA Failures in the Web PKI. In *Proceedings of the 2013 International Conference on Social Computing (SOCIALCOM '13)*. IEEE Computer Society, Washington, DC, USA, 600–605. https://doi.org/10.1109/SocialCom.2013.90

[15] Bundesverfassungsgericht. 2006. Authorisation to shoot down aircraft in the Aviation Security Act void. https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/2006/02/rs20060215_1bvr035705en.html.

[16] CA/Browser Forum. 2018. Baseline Requirements for the Issuance and Management of Publicly-Trusted Certificates. https://cabforum.org/wp-content/uploads/CA-Browser-Forum-BR-1.6.0.pdf

[17] CA/Browser Forum. 2018. Guidelines For The Issuance And Management Of Extended Validation Certificates Version 1.6.8. https://cabforum.org/wp-content/uploads/CA-Browser-Forum-EV-Guidelines-v1.6.8.pdf

[18] CA/Browser Forum. 2018. Network and Certificate System Security Requirements Version 1.1. https://cabforum.org/wp-content/uploads/CABForum_Network_Security_Controls_v.1.1-corrected.pdf

[19] F. H. Cate. 2010. The Limits of Notice and Choice. *IEEE Security & Privacy* 8, 2 (March 2010), 59–62.

[20] J. Classen, J. Braun, F. Volk, M. Hollick, J. Buchmann, and M. Mühlhäuser. 2015. A Distributed Reputation System for Certification Authority Trust Management. In *2015 IEEE Trustcom/BigDataSE/ISPA*, Vol. 1. 1349–1356. https://doi.org/10.1109/Trustcom.2015.529

[21] A. Costello. 2003. *Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA).* RFC 3492. Internet Engineering Task Force (IETF). https://tools.ietf.org/html/rfc3492

[22] A. Deacon and R. Hurst. 2007. *The Lightweight Online Certificate Status Protocol (OCSP) Profile for High-Volume Environments.* RFC 5019. Internet Engineering Task Force (IETF). https://tools.ietf.org/html/rfc5019

[23] Debian Security Team. 2008. CVE-2008-0166. https://security-tracker.debian.org/tracker/CVE-2008-0166

[24] DigiCert. 2018. Getting Ahead of Chrome 70 Distrust of Symantec-Issued Certificates. https://www.digicert.com/blog/getting-ahead-chrome-70-distrust-symantec-issued-certificates/

[25] Zakir Durumeric, David Adrian, Ariana Mirian, Michael Bailey, and J. Alex Halderman. 2015. A Search Engine Backed by Internet-Wide Scanning. In *22nd ACM Conference on Computer and Communications Security.*

[26] Donald E. Eastlake and Charles W. Kaufman. 1997. *Domain Name System Security Extensions.* RFC 2065. Internet Engineering Task Force (IETF). https://tools.ietf.org/html/rfc2065

[27] European Telecommunications Standards Institute (ETSI). 2015. *Electronic Signatures and Infrastructures (ESI); Trust Service Provider Conformity Assessment - Requirements for conformity assessment bodies assessing Trust Service Providers.* ETSI EN 319 403 V2.2.2. https://www.etsi.org/deliver/etsi_en/319400_319499/319403/02.02.02_60/en_319403v020202p.pdf

[28] European Telecommunications Standards Institute (ETSI). 2018. *Electronic Signatures and Infrastructures (ESI); General Policy Requirements for Trust Service Providers.* ETSI EN 319 401 V2.2.1. https://www.etsi.org/deliver/etsi_en/319400_319499/319401/02.02.01_60/en_319401v020201p.pdf

[29] European Telecommunications Standards Institute (ETSI). 2018. *Electronic Signatures and Infrastructures (ESI); Policy and security requirements for Trust Service Providers issuing certificates; Part 1: General requirements.* ETSI EN 319 411-1 V1.2.2. https://www.etsi.org/deliver/etsi_en/319400_319499/31941101/01.02.02_60/en_31941101v010202p.pdf

[30] T. Fadai, S. Schrittwieser, P. Kieseberg, and M. Mulazzani. 2015. Trust me, I'm a Root CA! Analyzing SSL Root CAs in Modern Browsers and Operating Systems. In *2015 10th International Conference on Availability, Reliability and Security.* 174–179. https://doi.org/10.1109/ARES.2015.93

[31] Freedom House. 2014. Freedom on the Net 2014. https://freedomhouse.org/sites/default/files/FOTN_2014_Full_Report_compressedv2_0.pdf

[32] Dieter Gollmann. 2006. Why Trust is Bad for Security. *Electronic Notes in Theoretical Computer Science (ENTCS)* 157, 3 (May 2006), 3–9. https://doi.org/10.1016/j.entcs.2005.09.044

[33] Dan Goodin. 2015. Dell does a Superfish, ships PCs with easily cloneable root certificates. https://arstechnica.com/information-technology/2015/11/dell-does-superfish-ships-pcs-with-self-signed-root-certificates/

[34] Dan Goodin. 2015. Lenovo PCs ship with man-in-the-middle adware that breaks HTTPS connections [Updated]. https://arstechnica.com/information-technology/2015/02/lenovo-pcs-ship-with-man-in-the-middle-adware-that-breaks-https-connections/

[35] Google LLC. 2018. Root Certificate Policy. https://www.chromium.org/Home/chromium-security/root-ca-policy

[36] Robert Graham. 2015. Extracting the SuperFish certificate. https://blog.erratasec.com/2015/02/extracting-superfish-certificate.html

[37] Michael P. Heinl. 2019. *A metric to assess the trustworthiness of certificate authorities.* Master's Thesis. University of Ulm, Ulm, Germany. https://doi.org/10.18725/OPARU-12173

[38] Hans Hoogstraaten, Ronald Prins, Daniël Niggebrugge, Danny Heppener, Frank Groenewegen, Janna Wettinck, Kevin Strooy, Pascal Arends, Paul Pols, Robbert Kouprie, Steffen Moorrees, Xander van Pelt, and Yun Zheng Hu. 2012. *Black Tulip: Report of the investigation into the DigiNotar Certificate Authority breach.* Technical Report. Fox-IT BV.

[39] International Telecommunication Union (ITU). 2016. *Information technology – Open Systems Interconnection – The Directory: Public-key and attribute certificate frameworks.* ITU-T Recommendation X.509, ISO/IEC 9594-8:2017. https://www.itu.int/rec/T-REC-X.509-201610-I/en

[40] D. Kumar, Z. Wang, M. Hyder, J. Dickinson, G. Beck, D. Adrian, J. Mason, Z. Durumeric, J. A. Halderman, and M. Bailey. 2018. Tracking Certificate Misissuance in the Wild. In *2018 IEEE Symposium on Security and Privacy (SP).* 785–798. https://doi.org/10.1109/SP.2018.00015

[41] Adam Langley. 2013. Enhancing digital certificate security. https://security.googleblog.com/2013/01/enhancing-digital-certificate-security.html

[42] Ben Laurie. 2014. Certificate Transparency. *Queue* 12, 8, Article 10 (Aug. 2014), 10 pages. https://doi.org/10.1145/2668152.2668154

[43] B. Laurie, A. Langley, and E. Kasper. 2013. *Certificate Transparency.* RFC 6962. Internet Engineering Task Force (IETF). https://tools.ietf.org/html/rfc6962

[44] Arjen Lenstra and Benne de Weger. 2005. On the Possibility of Constructing Meaningful Hash Collisions for Public Keys. In *Information Security and Privacy*, Colin Boyd and Juan Manuel González Nieto (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 267–279.

[45] Let's Encrypt. 2015. The CA's Role in Fighting Phishing and Malware. https://letsencrypt.org/2015/10/29/phishing-and-malware.html

[46] A. M. McDonald and L. F. Cranor. 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society* 4, 3 (2008).

[47] Microsoft. 2011. Microsoft Security Advisory 2607712. https://docs.microsoft.com/en-us/security-updates/SecurityAdvisories/2011/2607712

[48] Microsoft Inc. 2018. Microsoft Trusted Root Program Requirements. http://aka.ms/RootCert

[49] MITRE. 2018. CWE-338: Use of Cryptographically Weak Pseudo-Random Number Generator (PRNG). https://cwe.mitre.org/data/definitions/338.html

[50] Mozilla Foundation. 2018. Mozilla Root Store Policy, Version 2.5. https://www.mozilla.org/en-US/about/governance/policies/security-group/certs/policy/

[51] Johnathan Nightingale. 2011. DigiNotar Removal Follow Up. https://blog.mozilla.org/security/2011/09/02/diginotar-removal-follow-up/

[52] Devon O'Brien, Ryan Sleevi, and Andrew Whalley. 2018. Chrome's Plan to Distrust Symantec Certificates. https://security.googleblog.com/2017/09/chromes-plan-to-distrust-symantec.html

[53] Henning Perl, Sascha Fahl, and Matthew Smith. 2014. You Won't Be Needing These Any More: On Removing Unused Certificates from Trust Stores. In *Financial Cryptography and Data Security*, Nicolas Christin and Reihaneh Safavi-Naini (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 307–315.

[54] Adrian Perrig, Pawel Szalachowski, Raphael M. Reischuk, and Laurent Chuat. 2017. *SCION: A Secure Internet Architecture*. Springer International Publishing AG. https://doi.org/10.1007/978-3-319-67080-5

[55] Reporters without Borders. 2014. 2014 World Press Freedom Index. https://rsf.org/sites/default/files/index2014_en.pdf

[56] Andrew Rukhin, Juan Soto, James Nechvatal, Miles Smid, Elaine Barker, Stefan Leigh, Mark Levenson, Mark Vangel, David Banks, Alan Heckert, James Dray, and San Vo. 2010. *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*. NIST Special Publication 800-22 Revision 1a. National Institute of Standards and Technology (NIST).

[57] Chris Salter, O. Sami Saydjari, Bruce Schneier, and Jim Wallner. 1998. Toward a Secure System Engineering Methodolgy. In *Proceedings of the 1998 Workshop on New Security Paradigms (NSPW '98)*. ACM, New York, NY, USA, 2–10.

[58] S. Santesson, M. Myers, R. Ankney, A. Malpani, S. Galperin, and C. Adams. 2013. *X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP*. RFC 6960. Internet Engineering Task Force (IETF). https://tools.ietf.org/html/rfc6960

[59] Emily Schechter. 2018. A milestone for Chrome security: marking HTTP as "not secure". https://www.blog.google/products/chrome/milestone-chrome-security-marking-http-not-secure/

[60] Quirin Scheitle, Taejoong Chung, Jens Hiller, Oliver Gasser, Johannes Naab, Roland van Rijswijk-Deij, Oliver Hohlfeld, Ralph Holz, Dave Choffnes, Alan Mislove, and Georg Carle. 2018. A First Look at Certification Authority Authorization (CAA). *SIGCOMM Comput. Commun. Rev.* 48, 2 (May 2018), 10–23. https://doi.org/10.1145/3213232.3213235

[61] Guido Schryen, Melanie Volkamer, Sebastian Ries, and Sheikh Mahbub Habib. 2011. A Formal Approach Towards Measuring Trust in Distributed Systems. In *Proceedings of the 2011 ACM Symposium on Applied Computing (SAC '11)*. ACM, New York, NY, USA, 1739–1745. https://doi.org/10.1145/1982185.1982548

[62] Murugiah Souppaya and Karen Scarfone. 2016. *Guide to Data-Centric System Threat Modeling*. Draft NIST Special Publication 800-154. National Institute of Standards and Technology (NIST).

[63] Michael Alan Specter. 2016. *The Economics of Cryptographic Trust: Understanding Certificate Authorities*. Master's Thesis. Massachusetts Institute of Technology, Cambridge, MA, USA.

[64] Marc Stevens, Elie Bursztein, Pierre Karpman, Ange Albertini, and Yarik Markov. 2017. The First Collision for Full SHA-1. In *Advances in Cryptology – CRYPTO 2017*, Jonathan Katz and Hovav Shacham (Eds.). Springer International Publishing, Cham.

[65] The Internet Corporation for Assigned Names and Numbers (ICANN). 2018. TLD DNSSEC Report (2018-07-19 00:02:15). http://stats.research.icann.org/dns/tld_report/

[66] Transparency International. 2014. Corruption Perceptions Index 2014. https://www.transparency.org/cpi2014

[67] Z.E. Uahhabi and H.E. Bakkali. 2016. An approach for evaluating trust in X.509 certificates. In *2016 11th International Conference for Internet Technology and Secured Transactions, ICITST 2016*. 196–203. https://doi.org/10.1109/ICITST.2016.7856696 cited By 0.

[68] Roland van Rijswijk-Deij. 2017. *Improving DNS security: a measurement-based approach*. Ph.D. Dissertation. University of Twente. https://doi.org/10.3990/1.9789036543293 CTIT Ph.D. thesis series no. 17-430.

[69] Ahmad Samer Wazan, Romain Laborde, François Barrere, Abdelmalek Benzekri, and David W. Chadwick. 2013. PKI Interoperability: Still an Issue? A Solution in the X.509 Realm. In *Information Assurance and Security Education and Training*, Ronald C. Dodge and Lynn Futcher (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 68–82.

[70] A. S. Wazan, R. Laborde, F. Barrère, and A. Benzekri. 2012. The X.509 trust model needs a and legal expert. In *2012 IEEE International Conference on Communications (ICC)*. 6895–6900. https://doi.org/10.1109/ICC.2012.6364860

[71] A. S. Wazan, R. Laborde, D. W. Chadwick, F. Barrere, and A. Benzekri. 2016. How Can I Trust an X.509 Certificate? An Analysis of the Existing Trust Approaches. In *2016 IEEE 41st Conference on Local Computer Networks (LCN)*. 531–534. https://doi.org/10.1109/LCN.2016.85

[72] Ahmad Samer Wazan, Romain Laborde, David W. Chadwick, Francois Barrere, Abdelmalek Benzekri, Mustafa Kaiiali, and Adib Habbal. 2017. Trust Management for Public Key Infrastructures: Implementing the X.509 Trust Broker. In *Security and Communication Networks*, Vol. 2017. Hindawi. https://doi.org/10.1155/2017/69071465

[73] WebTrust. 2017. *WebTrust Principles and Criteria for Certification Authorities*. Version 2.1. http://www.webtrust.org/principles-and-criteria/docs/item85228.pdf

[74] WebTrust. 2018. *WebTrust Principles and Criteria for Certification Authorities - SSL Baseline with Network Security*. Version 2.3. http://www.webtrust.org/principles-and-criteria/docs/item85437.PDF

[75] Florian Weimer. 2008. New openssl packages fix predictable random number generator. https://lists.debian.org/debian-security-announce/2008/msg00152.html

[76] Kathleen Wilson and Gervase Markham. 2017. MozillaWiki - CA:Symantec_Issues. https://wiki.mozilla.org/CA:Symantec_Issues

[77] Liz Woolery, Ryan Budish, and Kevin Bankston. 2016. *The Transparency Reporting Toolkit*. New America and The Berkman Center for Internet & Society at Harvard University.

A APPENDIX

Table 6: Submetric details.

SMG	Submetric	Obtainability	Description	Variable Declaration	Score Calculcation
Revocation	UTD of CRLs	A	The BRs require that CAs have to publish an updated CRL at least every seven days. However, the more often a CA updates its CRL, the lower the risk for users to erroneously accept an actually revoked certificate.	t: the (average) difference between the current time and the CRL's `Last Update` field in days.	$(1 - \frac{\lfloor t \rfloor}{7+1})^+$
	UTD of OCSP	A	The BRs require that CAs have to update their OCSP information at least every four days. However, the more often a CA updates its OCSP information, the lower the risk for users to erroneously accept an actually revoked certificate.	t: the (average) difference between the current time and the OCSP response's `This Update` field in days.	$(1 - \frac{\lfloor t \rfloor}{4+1})^+$
	Consistency of CRL & OCSP	A	CRLs and OCSP fulfill the same task and must therefore be consistent if both are provided.	S: the set of random samples chosen from the CRL; n: the number of corresponding OCSP responses returning `Cert Status: revoked`.	$\frac{n}{\lvert S \rvert}$
	Non-existent Certificates	A	An OCSP request answered with a good response signals that the corresponding certificate has not been revoked. However, for a non-existent certificate which indeed cannot be revoked because it has not been issued, this kind of response also implicitly confirms its existence. Hence, before responding to an OCSP request, a CA has to check whether the certificate has been issued after all and if not, it has to reply with `Cert Status: unknown` or with a `unauthorized` response depending on the CA's ability to access authoritative records for the corresponding certificate [22].	s: `Cert Status`	if s == `good` 0 else 1
	Revocation Request Reaction Time	B	The BRs define that a CA has to react within 24 hours after receipt of a revocation request filed by the certificate holder. It is crucial that the corresponding certificate is revoked as soon as possible because in the meantime certificate abuse could take place depending on the reason of the revocation request. The important question is therefore not how long it takes until the CA responds to the certificate holder but how long it takes until measures are taken which reflect the incident in CRLs and OCSP responses. This could either happen temporarily by employing the revocation reason `certificateHold` or permanently [58].	t: the timespan in hours from the moment the revocation request is filed to the moment the certificate is revoked.	$(1 - \frac{t}{24+1})^+$
	Receptiveness of Revocation	P	A CA should not only react to revocation requests filed by the holder of the certificate but also to CPRs filed by any third party detecting misuse or fraud related to the certificate. In order to reduce inhibitions and facilitate the process, CAs should provide instructions and forms supporting the creation of CPRs.	i: the existence of information concerning revocation; f: the existence of a distinct, public form to request revocation (both boolean).	$(i + f) * 0.5$
Restriction	Path Length Constraints	A	Every CA issuing a certificate for a subordinate CA has to set the field `cA` to `true` in order to enable the corresponding subordinate CA issuing certificates on its own. Hence, the `cA` field, which is part of the basic constraints extension, is a powerful tool used quite frequently. It does, however, not only empower the holding CA to issue end-entity but also further CA certificates. Hence, the basic constraints extension also contains a mechanism to restrict a subordinate CA's ability to further issue CA certificates by defining a maximum path length which is crucial in order to minimize the risk of certificate misissuance and abuse. Therefore, this submetric assesses to which extent a CA utilizes path length constraints.	C: the set of subordinate CA certificates issued by the assessed CA; $R \subseteq C$: only certificates with `pathLenConstraint` set to `0`.	$1 - \frac{\lvert C \rvert - \lvert R \rvert}{\lvert C \rvert}$
	Name Constraints	A	A CA can hinder subordinate CAs from issuing certificates for other DNs or IP addresses than the ones it is legitimately entitled to by including name constraints. The name constraints extension comprises the two fields `permittedSubtrees` and `excludedSubtrees`. The use of `permittedSubtrees` is thereby seen as a more effective way because it implicitly excludes any names except for the ones explicitly permitted. `excludedSubtrees`, however, only prohibits issuing certificates with DNs explicitly listed. Hence, the use of `excludedSubtrees` is recommended for further narrowing down the scope of a certificate but should not be used as a stand-alone solution.	C: the set of subordinate CA certificates issued by the assessed CA; $R \subseteq C$: only restricted certificates with `permittedSubtrees` field set.	$1 - \frac{\lvert C \rvert - \lvert R \rvert}{\lvert C \rvert}$

Table 6: Submetric details (continued).

SMG	Submetric	Obtainability	Description	Variable Declaration	Score Calculcation
	Key Usage	A	In order to properly restrict the usage of a CA certificate issued for the WebPKI and contain potential damage in case of compromise, `keyUsage` should be set to `keyCertSign`.	C: the set of subordinate CA certificates having set the `cA` field to `true`; $R \subseteq C$: only subordinate CA certificates with `keyUsage` set to `keyCertSign`.	$1 - \frac{\lvert C\rvert - \lvert R\rvert}{\lvert C\rvert}$
Issuance	CAA	B	In order to provide an additional line of defense against illegitimate CSRs, the CA should check the CAA RR's `issue` property tag and reject the request if it does not point on the CA's DN. Furthermore, the subject should be warned about illegitimate requests by reaching out to the contact information defined in the `iodef` property tag.	p: the CA reports a violation; r: the CA rejects CSR because of violation (both boolean).	$(p + r) * 0.5$
	High-Risk CSR	B	Even legitimately requested certificates can be used for fraudulent activity by containing components confusing the user. In order to detect and reject CSRs with a high risk of being used as part of fraudulent activity such as phishing or scam, corresponding mechanisms have to be employed, including well-known services such as the MillerSmiles phishing list and Google's Safe Browsing list.[7,8] Furthermore, the CA's own historical data can be used as an indicator for a high-risk CSR, for example the similarity to already revoked certificates or previously rejected CSRs.	C: the set of all CSRs including high-risk CSRs; $D \subseteq C$: all detected CSRs.	$\frac{\lvert D\rvert}{\lvert C\rvert}$
	Mixed Character Set IDN	B	Using other character sets than ASCII, IDNs with identically looking symbols can be used to issue certificates for DNs resembling high-risk DNs without triggering the corresponding security controls.	C: the set of all CSRs including high-risk mixed character set IDN; $D \subseteq C$: all detected CSRs.	$\frac{\lvert D\rvert}{\lvert C\rvert}$

	Origin Country	B	In order to prevent fraud, the CA is supposed to properly check the origin of a CSR before including the `CountryName` field in a certificate.	s: status of CSR; c certificate's country code field (`C=`).	if (s == `rejected`) OR (c == `NULL`)
					1
					else
					0

	Extended Key Usage	A	In order to properly restrict the usage of an end-entity certificate issued for the WebPKI and contain potential damage in case of compromise, EKU should be set to `id-kp-serverAuth`.	C: the set of end-entity certificates having set the `cA` field to `false` or not set at all; $R \subseteq C$: only certificates with EKU set to `id-kp-serverAuth`.	$1 - \frac{\lvert C\rvert - \lvert R\rvert}{\lvert C\rvert}$
	Wildcard Certificates	A	Wildcard certificates considerably ease the process of certificate issuance for both the CA and its customers who potentially have to manage a huge and rapidly changing number of subdomains. However, every issued wildcard certificate also comes with the risk of abuse because it is also valid for subdomains containing deliberately deceptive elements.[9]	C: the set of end-entity certificates having set the `cA` field to `false` or not set at all; $W \subseteq C$: all wildcard certificates.	$1 - \frac{\lvert W\rvert}{\lvert C\rvert}$

Cryptography	Key Size	A	The BRs define the minimum requirements for public keys according to NIST recommendations [9]. The BRs allow weaker parameters for legacy certificates. For the sake of consistent security, these certificates could be proactively replaced by CAs, though. Hence, these exceptions are not taken into consideration. With each certificate employing a key length shorter than defined by the BRs, the score decreases exponentially and independently from the total number of issued certificates due to the *weakest link problem*.	C: the set of all certificates issued by the CA; $F \subseteq C$: active certificates failing to accomplish the minimum requirements.	if ($\lvert F\rvert == \lvert C\rvert$)
					0
					else
					$\frac{1}{1+\lvert F\rvert}$

	Digest Algorithm	A	According to the BRs, only SHA-2 with a minimum digest length of 256 bits is sufficient as the digest algorithm for signature generation. SHA-1 and even MD5 are allowed for legacy certificates, although not recommended. Because of their known weaknesses [44, 64] they are both considered to be insufficient for the purpose of this paper.	C: the set of all certificates issued by the CA; $F \subseteq C$: certificates which do not employ the digest algorithms defined in the BRs.	$1 - \frac{\lvert F\rvert}{\lvert C\rvert}$

[7] http://www.millersmiles.co.uk/

[8] https://developers.google.com/safe-browsing/v4/lists

[9] The wildcard certificate for `*.uni-ulm.de`, for example, could be used to legitimately operate the domain `thisIsYourFavouriteBank.uni-ulm.de` which could be used for fraudulent purposes.

Table 6: Submetric details (continued).

SMG	Submetric	Obtainability	Description	Variable Declaration	Score Calculcation
	Public Key Reuse	B	A certificate enables its owner to identify herself by proving to possess the corresponding private key. This statement assumes that solely the certificate's owner is in possession of the private key. Hence, each certificate's key pair has to be unique. In order to prevent attacks such as DROWN [7] and impersonation, the CA should check whether the public key contained in a CSR is already in use with another certificate [12].	r: CA checks for key reuse.	if $r == true$ 1 else 0
	Weak Key	B	Randomness is crucial to generate strong cryptographic keys [49, 56]. In the past, however, there have been implementation flaws drastically reducing entropy. For example, an instance of the OpenSSL suite included in some versions of the operating system *Debian* takes its process ID as seed [23].[10] Since the maximum possible process ID is `32,768`, the outcome of the PRNG is predictable. Hence, the CA has to check whether the public key provided by the certificate applicant has been generated employing flawed methods such as predictable pseudo-random number generators (PRNG).	r: CA rejects weak key.	if $r == true$ 1 else 0
Independence	Operational Independence	P/R	In order to avoid conflicts of interest between state actors and CAs, it is desirable that CAs are completely independent from any operational influence of state actors. An obvious indicator for influence is that the CA is directly operated by the state. Other, more subtle, indicators are governmental agencies holding shares or funding the CA.	n: the smallest number of edges between the CA and any governmental agency having operational, financial, or coordinating influence over the CA.	if $n == NULL$ 1 else $1 - \frac{1}{n}$
	Legal Independence	P	This submetric indicates how likely it is that in the countries where the CA is located, legislation exists which is allowing/not preventing government agencies from forcing CAs to issue rogue certificates. For this purpose, the indexes proposed by Fadai et al. [30] are first normalized to match the unit interval format employed in this paper.[11] Then, the combined average score for the CA's place of business can be easily calculated using the formula stated below. For some CAs, such as *Thawte*, the CA's certificates' country attribute (`C=US`) and the company's legal headquarters (South Africa) are not equal. In these cases, the score is calculated for both countries whereby the worse is adopted. The same applies for foreign subsidiaries of enterprises. The legal status of capital punishment [3] is neglected since it lacks an ordinal scale because human lives must not be weighed up against each other [15].	c: the normalized score of the CPI; n: the normalized score of the FOTNR; p the normalized score of the WPFI.	$\frac{c+n+p}{3}$
Transparency	Certificate Transparency	A	In order to detect mistakenly or even maliciously issued certificates, CT requires CAs to publish any issuance on public and independent CT log servers. This metric quantifies how well a CA follows this practice. With each certificate not properly logged on CT servers, the score decreases exponentially and independently from the total number of issued certificates because of the *weakest link problem*.	C: the set of all (at least once) browser-trusted certificates issued by the CA; $L \subseteq C$: certificates contained in public CT logs, having valid SCTs attached, or including the CT Poison extension.	if $(\lvert L\rvert == 0)$ 1 else $\frac{1}{1+\lvert C\rvert-\lvert L\rvert}$
	Document Repository	P	Although it should not be solely relied on information contained in the CP, CPS, or audit report, those are important documents indicating how a CA approaches their core business. Therefore, all three of them should be publicly available. Applying Mozilla's requirement for the audit report to all of those documents, they should be provided in English language because of their global impact. Ideally, all of them are stored in one central repository in order to lower barriers for the interested public.	n_1, n_2, n_3: the availability of the CP, CPS, and audit report, respectively; n_4, n_5, n_6: if their language is English; n_7 that these documents can be found in one central repository.	$\frac{\sum_{i=1}^{7} n_i}{7}$
	Legal Transparency Report	P	For the sake of transparency, CAs should follow the example of other telecommunication companies and publicly provide *legal transparency reports* [77]. These reports contain information about disclosure requests made by authorities and help users to estimate possible risks.	n: the difference between the current date and the most recent legal transparency report in years.	if $(n == NULL)$ 0 elseif $(n \leq 1)$ 1 else 0.5

[10] Beginning with OpenSSL version 0.9.8c-1 and ended with 0.9.8c-4etch3 (stable) or 0.9.8g-9 (unstable and testing), respectively [75].

[11] The indexes are represented by scores between 0 and 100, whereby 0 is the best and 100 is the worst in the case of the FOTNR and the WPFI. For the CPI, the opposite applies. Thus, the native scores are recalculated to range between 0 (worst) and 1 (best) without losing their significance.

Table 7: Evaluation results (raw data collection details without calculation).

		Comodo	SSL.com	Let's Encrypt	Thawte
General	Organization (O=)	COMODO CA Limited	SSL.com	Let's Encrypt	thawte, Inc.
	No. of EE certs[12]	52,683,134	52,747	597,018,962	881,790
	No. of CA certs[13]	1,616	0	41	44
	No. of RSA certs[14]	15,589,322	52,734	585,744,437	881,411
	Country (C=)	GB	US	US	US
	Headquarters	CA: GB / Corp.: US	US	US	CA: ZA / Corp.: US
Revocation	UTD of CRLs	< 1d	< 1d	n/a	< 1d
	UTD of OCSP	1d 12m	1d 4h 35m	1d 5h 18m	n/a
	CRL & OCSP Consistency	5/5	5/5	n/a	n/a
	Non-existent Certs	unauthorized	unauthorized	unauthorized	n/a
	Revoc. Req. React. Time	4.68 h	2.87 h	< 1h	> 24h
	Receptiveness of Revoc.	only info[15]	not available	only info[16,17]	info & form[18]
Restriction	Path Length Constraints	125	0	0	0
	Name Constraints	4	0	0	0
	Key Usage	1,571	0	1	44
Issuance	CAA	no iodef	no iodef	no iodef	no iodef
	High-Risk CSR	issued	issued	issued	issued
	Mixed Character Set IDN	issued	not issued	issued	issued
	Origin Country	no C=	no C=	no C=	issued
	Extended Key Usage	52,678,705	52,738	597,018,962	880,103
	Wildcard Certificates	38,973,091	8,598	13,081,442	80,961
Cryptography	Key Size < 2,048	18	0	4	0
	Digest Algorithm	0	0	0	0
	Public Key Reuse	issued	issued	issued	issued
	Weak Key	not accepted	not accepted	not accepted	not accepted
Independence	Operational	private	private	non-profit	private
	CPI '17	GB: 82 / US: 75	75	75	ZA: 43 / US: 75
	FOTNR '17	GB: 24 / US: 21	21	21	ZA: 25 / US: 21
	WPFI '18	GB: 23.25 / US: 23.73	23.73	23.73	ZA: 20.39 / US: 23.73
Transparency	Non-CT	629	7	1	23
	Document Repository	no CP & Audit Report	available	available	available
	Legal Transp. Reports	not available	not available	available	not available

[12] As of October 21, 2018.
[13] See Footnote 12.
[14] See Footnote 12.
[15] https://www.comodoca.com/en-us/support/report-abuse/
[16] https://letsencrypt.org/repository/
[17] https://community.letsencrypt.org/t/how-to-report-abuse/41106/2
[18] https://www.thawte.com/about/contact/ssl-certificate-complaint.html

Table 8: Evaluation results (down to the submetrics level).

		Comodo	SSL.com	Let's Encrypt	Thawte
Revocation	UTD of CRLs	1	1	n/a	1
	UTD of OCSP	0.8	0.76	0.76	n/a
	Consistency of CRL & OCSP	1	1	n/a	n/a
	Non-existent Certificates	1	1	1	n/a
	Revoc. Req. Reaction Time	0.81	0.86	1	0
	Receptiveness of Revoc.	0.5	0	0.5	1
	Recovation SMG score	**0.85**	**0.77**	**0.82**	**0.67**
Restriction	Path Length Constraints	0.08	1	0	0
	Name Constraints	0	1	0	0
	Key Usage	0.97	1	0.02	1
	Restriction SMG score	**0.35**	**1.00**	**0.01**	**0.33**
Issuance	CAA	0.5	0.5	0.5	0.5
	High-Risk CSR	0	0	0	0
	Mixed Character Set IDN	0	1	0	0
	Origin Country	1	1	1	0
	Extended Key Usage	1	1	1	1
	Wildcard Certificates	0.26	0.84	0.98	0.91
	Issuance SMG score	**0.46**	**0.72**	**0.58**	**0.40**
Cryptography	Key Size	0.05	1	0.2	1
	Digest Algorithm	1	1	1	1
	Public Key Reuse	0	0	0	0
	Weak Key	1	1	1	1
	Cryptography SMG score	**0.51**	**0.75**	**0.55**	**0.75**
Independence	Operational Independence	1	1	1	1
	Legal Independence	0.76	0.77	0.77	0.65
	Independence SMG score	**0.88**	**0.89**	**0.89**	**0.83**
Transparency	Certificate Transparency	0	0.13	0.5	0.04
	Document Repository	0.43	1	1	1
	Legal Transp. Reports	0	0	1	0
	Transparency SMG score	**0.14**	**0.38**	**0.83**	**0.35**
Unweighted overall trustworthiness score		**0.53**	**0.75**	**0.61**	**0.55**