Building trust through risk management in computer science

Maryline Laurent, Armen Khatchatourov

Laurent, M., Khatchatourov, A.

«Building trust through risk management in computer science»

in **Signs of trust – The impact of seals on personal data management** (Chapter 4, pages 48 to 60). Coordinated by Claire Levallois-Barth, Chair Values and Policies of Personal Information (France), January 2018.

Handbook available in electronic version: http://www.personal-information.org/
Also available in paper format: ISBN 978-2-9557308-6-7

Building trust through risk management in computer science

Trust in computer science mostly relies on evaluating the risks of using a tool (software or hardware) or, more generally, any other form of digital service (i.e. a website). This evaluation and its reliability are all the more critical as stakes get higher: they are most important when dealing with an organisation's Information Systems Security (ISS).[1]

There are two main approaches to risk qualification. The older method regards products delivering ISS (software and hardware) and implementing security functions, and trust service providers (e.g. providing timestamps, signatures, electronic certificates). Often, public authorities are involved in the process of qualifying the level of risk. Here, trust is assumed to be transitive: if users trust the qualifying entity or the electronic certificate, they will also trust the object that is qualified. Qualifying products or providers is not always mandatory, but it is unavoidable when designing critical security solutions or competing for public procurement, among other cases. Therefore, such risk management approach is, together with the reliability level it is associated with, an external sign aiming to reinforce the trust of individuals and companies (4.1.).

The second and more recent approach relies on the large and growing number of data points available in IT system. It works by scoring the security performance of individuals and services. This score, used as a risk indicator, is based on a behavioural analysis that

1 In this chapter, "Information Systems Security" refers to all the technical, organisational, legal and human processes in place to ensure the protection of an organisation's IT system.

benchmarks one behaviour with a reference behaviour. It is likely to have a direct influence on trust (4.2.).

These approaches – product and provider qualification and behavioural analysis — can be used jointly, for instance in order to authenticate a user based both on an electronic certificate and on their behaviour.

4.1. Risk evaluation of ISS products and services

Security is fundamental for States and companies

The evaluation of the risk associated with using ISS products and services has historically been tied with the strong need for companies and States to keep providing trusted and available infrastructure and services and fight cybersecurity threats. Designed in a top-down way, risk evaluations take place in a strict framework laid down by national and/or European authorities. This framework regulates the reliability level expected from services, hardware and software contributing to the security of information systems — each level is associated with a level of qualification. The goal here is to maintain a high level of vigilance, the stakes being all at once economic, political and strategic. Therefore, in order to ensure national sovereignty, States qualify ISS products and trust services likely to be used by their administration, critical infrastructure providers or otherwise sensitive companies. The highest level of qualification corresponds to low risk-taking and is therefore adapted to critical infrastructures.

None of these qualifications are mandatory, yet they are difficult to avoid in practice. In particular, they make it easier, through a sort of nested doll effect, to obtain data protection seals, as they guarantee that confidentiality and security requirements are taken into account. Besides, certain regulations are compulsory, notably those regarding the provision, import, export, or transfer of cryptological tools associated with a product or service towards another EU country.

In this context, public authorities — in France, the National Agency for the Security of Information Systems (ANSSI in French) — publish a catalogue of qualified products that includes the level of qualification obtained and the list of qualified trust service providers. This does not provide absolute guarantee — indeed, recent events have shown that certain security products included backdoors or purposefully deteriorated security functions so that data flows could be decrypted with no prior knowledge of secrets. In 2013, Reuters therefore revealed[2] that the National Security Agency (NSA) had paid a $10 million bribe to RSA so that it would implement by default a weak random number generator called Dual EC DRBG (Dual Elliptic Curve Deterministic Random Bit Generator) in their security product BSAFE, in order to enable rapid decryption of the data of millions of users. Besides, it seems the NSA also originated a modification of the Dual EC DRBG algorithm officially meant to enhance the security of the encrypted data; yet, as researchers have shown, the modification actually reinforced vulnerabilities.

ANSSI-issued qualifications for products

In France, ANSSI, within the Secretariat-General for National Defence and Security (SGDSN in French) under the Prime Minister's Office, developed its own certification scheme for information systems security products on the basis of a co-regulating scheme (see Chapter 5, "Numerous and heterogeneous seals…", page 64): the qualification is issued by ANSSI while the evaluation is carried out by private evaluation centres accredited by ANSSI. Depending on the products and levels of reliability, qualifications are issued based on audit or technical test results.

Three levels of qualification are issued[3] (see Table 1):

2 https://www.schneier.com/blog/archives/2007/11/the_strange_sto.html

3 Chochois, M., Magnin, N., (2015). *Qualité des produits de SSI, les labels français, Techniques de l'ingénieur,* H5825 v2, October 2015.

Object		Title	Benchmark	Number of qualified solutions	Duration of qualification
Products		Elementary qualification	ANSSI	70+	Unlimited for a given version
		Standard qualification	Common criteria EAL3+	30+	6 months
		Strong qualification	Common criteria EAL4+	70+	
Trust service providers	SecNum Cloud	Simple, advanced or qualified qualification depending on the type of service, *see "Identités numériques", Cahier n°1, Chair Values and Policies of Personal Information*	ANSSI	0	Up to 3 years
	PSCE			240+	
	PRIS			0	
	PDIS			0	
	PASSI			26	
	PSHE			240+	

Table 1. Security qualifications issued by ANSSI for products and trust service providers

SecNumCloud: Cloud Service Provider; PSCE: Electronic Certification Service Provider; PRIS: Security Incident Response Service Provider; PDIS: Security Incident Detection Service Provider; PASSI: Information Systems Security Audit Provider; PSHE: Timestamping Service Provider.

- **Elementary qualification** corresponds to a first-level seal for the ISS product, issued with limited time and resources. After ANSSI studies the file, an evaluation centre that has been accredited by ANSSI for First Level Security Certificates (CSPN in French) implements the CSPN certification scheme. Verifications include compliance of the product with its security specifications and the threats it protects against.

- **Standard qualification** requires more time and resources and guarantees the product for the treatment of sensitive unclassified information. The product is evaluated by the Centre for Evaluation of the Security of Information Technology (CESTI in French), also accredited by ANSSI. The evaluation relies on a benchmark with common criteria (see next section) under control of ANSSI. Standard qualification is granted for six months and requires the product to obtain at least the EAL3+ level determined by the common criteria. To this end, the manufacturer needs to provide several inputs, including cryptographic mechanisms (protection of private keys, random number management, etc.).

- **Strong qualification** also lasts six months and relies on obtaining an EAL4+ level of the common criteria. French products with this level of qualification are granted "*Confidentiel Défense*" and/or "*Secret Défense*" clearance, which enables them to deal with classified information.

International mutual recognition

Two different types of international mutual recognition agreements enable State A to accept a qualification issued by State B.

The first relies on the Common Criteria Recognition Arrangement (CCRA), the most recent update of which was signed in 2014. 28 countries currently recognise as valid the qualification of a given ISS product issued by one of their certification authorities, in accordance with the common criteria framework: Australia, Austria, Canada, the Czech Republic, Denmark, Ethiopia, Finland, France, Germany, Greece, Hungary, India, Israel, Italy, Japan, Malaysia, the Netherlands, New Zealand, Norway, Pakistan, Qatar, South Korea, Singapore, Spain, Sweden, Turkey, the United Kingdom, and the United States.

The common criteria allow to certify a product through a certification level called Evaluation Assurance Level (EAL); EAL1 being the lowest score and EAL7 the highest. They are often used to mandate certification levels according to uses. For instance, a smart card used for interbank transactions needs to be certified with at least EAL4+.

Mutual recognition agreements include certain limits depending on the type of evaluation scheme implemented. For evaluations under the generic common criteria, mutual recognition used to apply up to EAL2. In 2014, CCRA relaxed this rule and defined collaborative Protection Profiles (cPP) with a specific evaluation scheme on top of common criteria. For evaluations carried out according to cPP, mutual recognition now stands up to EAL4.

A second type of agreement was signed in 1999 and updated in 2010: the European Mutual Recognition Agreement of the Senior Officials Group Information Systems Security (SOG-IS).1 This agreement established mutual recognition of the validity of certificates in several technical domains. By default, the recognition applied up to EAL4 as with the common criteria arrangement — certain domains such as “smartcards and similar devices” and “hardware devices with security boxes” can benefit from a mutual recognition up to EAL7. 11 countries are part of this agreement: Austria, Finland, France, Germany, Italy, the Netherlands, Norway, Poland, Spain, Sweden, and the United Kingdom. For each technical domain, the agreement specifies which countries are qualified participants and can issue high-level qualifications.

The qualification of trust service providers

While ANSSI’s interventions may sometimes seem to disregard end users’ daily issues, the situation is changing with the implementation on July 1, 2016 of EU Regulation 910/2014 on electronic identification and trust services for electronic transactions in the internal market (eIDAS).[4]

The Regulation introduces a legal framework common to all EU Member States for electronic identification means and trust services: electronic signatures, electronic seals,

4 Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC (eIDAS Regulation), EUOJ L 257, 28.8.2014. Readers are encouraged to consult Levallois, C. (2016). *La réglementation mise en place par l’Union européenne en matière d’identification électronique et des services de confiance (règlement eIDAS). in « Identités numériques », Cahier n°1 de la Chaire Valeurs et Politiques des Informations Personnelles,* coordinated by Claire Levallois-Barth.

electronic timestamps, electronic documents, electronic registered delivery services and certificate services for website authentication. It requires transposition at the national level; in France, ANSSI is the responsible agency. ANSSI is currently establishing the eIDAS requirements and issuing accreditations to organisations responsible for evaluating compliance.[5] As anticipated in the eIDAS Regulation, ANSSI has defined 4 types of services it deems useful: cloud service providers, incident response service providers, incident detection service providers, and ISS audit providers.

However, although the eIDAS Regulation has established a certain level of harmonisation, including a common terminology for trust services, it also comes with some shortcomings and ambiguities relating to data protection and user privacy, specifically regarding tracking and surveillance abilities. On this topic, we refer the reader to Chapters 7, 8 and 9 of the first volume published by the **Chair Values and Policies of Personal Information on Digital Identities**.

4.2. New forms of risk analysis associated with services and users

Behavioural analysis

In computer science, behavioural analysis primarily aims at detecting intrusions in IT systems and risky behaviours. Initially, it relied on the creation of a "normal" behaviour model for the information system and required a long training period. Since then, technology and its use cases have evolved to focus on User Behaviour Analytics (UBA) and incorporate the latest advances in Big Data and Machine Learning.

Table 2 presents a snapshot of current trends in risk evaluation: individuals, services and platforms can all be the target of behavioural analyses, carried out either by a team of individuals (II.) or through automated algorithmic methods (III.). In order to identify the true purpose behind using behavioural analysis, it is necessary to know both the organisation

5 https://www.ssi.gouv.fr/administration/reglementation/confiance-numerique/le-reglement-eidas/documents-publies-par-lanssi/

setting the norm and the type of criteria that will characterise this norm. Such information allows to better identify the nature of IT risks, the type of trust, and to discuss potential abuses.

For instance, an analysis conducted by a set of individuals (II.) allows to score a service provided by individuals or companies, be it by Amazon or eBay for products, or by TripAdvisor for restaurants and hotels. The reliability associated with reputation gets higher with every score and comment left on such websites, since it becomes all the more difficult to compromise the scoring system by leaving either positive reviews on your own page or negative reviews on the pages of competitor services. These scoring systems have already had a major impact on consumer behaviours. According to a PhoCusWright study,[1] 83% of respondents state that reviews on TripAdvisor help them pick the "right" hotel. Even though the technical infrastructure is not sophisticated, designers intend to build a trust relationship between service providers and consumers by drawing up a risk indicator; however, this only ensures a level of trust that some would call "weak."

Behavioural analyses can also be automated by algorithms for better efficiency and accuracy (III.). They can target one individual in particular (III.1.). In such case, the analysis can be used to reinforce the authentication mechanism between this individual and the information system — in addition to password or hardware-based authentication, the distance between the individual's usual behaviour and their current behaviour is taken into account in the authentication process in order to limit risks. The usual behaviour is therefore taken as a benchmark, while the nature and size of the acceptable differences are set by the system administrator. The reliability associated with the behavioural analysis is based on the quality and size of the available data on the individual's behaviour within the system, and therefore on how precisely their behaviour can be quantified (geolocation, which applications are used when, from which terminals, …). It also depends on the algorithm's ability to detect any unusual behaviour. The tool should thus include personalised thresholds to avoid both wrongly accusing individuals (false positives) and not detecting identity fraud (false negatives).

The main purpose of automated individualised analysis can however be abused, especially to generalise control over people's behaviours (III.1.). Each individual could receive a score depending on their behaviour and from there advantages or penalties. For instance, China is working on a new "social credit scoring" system which is announced for 2020.

	Risk analysis associated with digital services	***Behavioural analysis: establishing a score for services/users***		
External sign	Digital certificates, qualified services	Scoring		
Evaluator	Certification organisation (I.)	Human (II.)	Algorithm (III.)	
Subject of evaluation	Hardware / Software / Digital services	Service provided	Individuals (III.1.)	Individuals / websites (III.2.)
Norm designer	European Commission / Institutions	Set of individuals	Government / Services	Institutions / Platforms
Volume of Data		Large datasets	Large datasets on individuals	Large set of individuals / websites
Forms of trust	EAL/eIDAS certification	Scoring	Profiling / Scoring	Profiling / Ranking
Trust in…	Certification organisation / Service provider	Operator / Platform / Government		
Proof of trust	ANSSI-issued list of qualified service providers and products	Number of evaluations	Algorithm and number of profiles	Algorithm
Purposes	Trust ++	Evaluation of a service	Authentication ++ / Surveillance	Cyber surveillance / Website ranking

Table 2. *Two approaches to risk management in computer science*

Chinese citizens would be “ranked” according to their actions, and the “riskiness” of their behaviours would be measured.

Finally, automated analysis can be used on large groups of individuals, platforms or websites (III.2.), with either commercial or political purposes here as well.

Scoring systems, most notably Google’s, rank popular websites according to keywords; Apple’s rank popular apps in the App Store. However, algorithms supposed to rank products, apps or websites according to their popularity are still very opaque in the way they work, which can make it difficult to prevent abuse. For instance, in exchange for $11,000, Taobao was able to consolidate its ranking in the top 10 mobile apps in the App Store.[6]

Automated analysis can also be used to support implementing legislation, such as the HADOPI2 law (Creation and Internet law) or the Intelligence Act in France. The Intelligence Act, passed in 2015 in the wake of the January 2015 terrorist attacks, entitles authorities to collect and process data related to internet connections (metadata) and defines the cases where such measures are allowed. This detection, which mainly aims at ensuring national security, preventing terrorist actions and defending France’s economic interests, may be automated by an algorithm that benchmarks user behaviours against pre-set “normal” behaviours.

Classically, we observe that behavioural analysis techniques are a double-edged sword. They can contribute to laudable objectives such as the overall security of the digital environment, but also to more problematic commercial or institutional ambitions.

6 https://recombu.com/mobile/article/manipulate-apple-app-store-rankings-for-money-in-china

4.3. Towards hybrid, distributed and privacy-preserving trust systems

In computer science, three solutions are currently being studied to limit security risks and data leaks and to increase trust in digital products and services.

- ***Hybrid approaches*** to reinforce the security of classic ISS solutions by using behaviour analysis methods. Improvements in Machine Learning and Big Data, together with the collection of data on a massive scale, have led to the increasing reliance of security services on behavioural analyses in order for them to define the behaviour of an individual or an information system and to be able to measure deviations. Banks, for instance, are implementing strong authentication mechanisms relying on usual strong cryptographic tools together with behavioural authentication including contextual data (geolocation, time of connections, IP addresses of the terminal, terminal fingerprinting)[7] and data about user-terminal interactions (browsing habits on a website, mouse movements, typing patterns). Future trends will dive deeper into these behaviours and be more specific about the risk levels incurred.

- More ***transparency*** and ***decentralised governance***. Blockchain-related solutions are heading in this direction. The first goal of blockchain is to provide a service administered by multiple authorities, instead of being centralised in the hands of a single one. The algorithm implementing the service is publicly accessible and readable, and can thus be interpreted by anybody; therefore, any change in the way the service functions or is governed needs to be approved by consensus of the participating authorities before it is implemented. The results are increased transparency, seemingly more stability, and the impression for users that have control over the service and actors, which results in higher levels of trust (see Chapter 11).

- A better ***protection of user privacy***. Technological solutions are being developed to guarantee both security and data protection. Among these solutions is anony-

7 The digital signature of a terminal (terminal fingerprinting) contains multiple pieces of information (OS version, screen resolution) which are meaningless on their own but the combination of which identifies a specific terminal amongst millions of others.

mous certification,[8] which aims to minimise the quantity of personal data collected by service providers while guaranteeing them strict access controls (that the users are not minors, that they are geographically located in a certain region, …). One can also mention ***homomorphic encryption***, which aims to delegate part of data processing to a third party without revealing unencrypted data, and ***secure multi-party computation***, which enables a group of participants to contribute to computing operations while hiding which operations are being carried out and the data on which the computation is being done. However, these solutions are still slow to develop in practice. They face technical obstacles, with high energy costs, and economic ones, with the lack of incentives to adopt other models than the exploitation of personal data.

If blockchain technologies, decentralised governance systems, and the work towards a better protection of privacy are indeed factors of trust, one interesting avenue for research would be to identify more precisely the technical solutions underpinning confidence-friendly environments. This topic of research is undoubtedly necessary, but also questions the intervention and the role of public authorities in this area.

8 Laurent, M., et Kaâniche, N. (2016). *Les preuves d'identités ou d'attributs préservant le pseudonymat ; in « Identités numériques », Cahier n°1 de la Chaire Valeurs et Politiques des Informations Personnelles,* coordinated by Claire Levallois-Barth.

How to cite this chapter: Laurent M., Khatchatourov A. “Building trust through risk management in computer science”, *in Signs of trust – The impact of seals on personal data management*, Paris, Handbook 2 Chair Values and Policies of Personal Information, Coordinated by Claire Levallois-Barth, January, 2018, chapter 4, pages 47–59.

http://www.personal-information.org/