Magdeburger Journal zur Sicherheitsforschung

Gegründet 2011 | ISSN: 2192-4260
Herausgegeben von Stefan Schumacher
Erschienen im Magdeburger Institut für Sicherheitsforschung
http://www.sicherheitsforschung-magdeburg.de/publikationen/journal.html

This article appears in the special edition »In Depth Security – Proceedings of the DeepSec Conferences«.
Edited by Stefan Schumacher and René Pfeiffer

Cryptographic Enforcement of Segregation of Duty

Thomas Maus

Workflows with Segregation-of-Duty requirements or involving multiple parties with non-aligned interests (typically mutually distrustful) pose interesting challenges in often neglected security dimensions.

Cryptographic approaches are presented to technically enforce strict auditability, traceability and multi-party-authorized access control, and thus also enable exoneration from allegations.

These ideas are illustrated by challenging examples - constructing various checks and balances for Telecommunications data retention, a vividly discussed and widely known issue.

Citation: Maus, T. (2017). Cryptographic Enforcement of Segregation of Duty. (Volume 14, Pages 858–870). Retrieved September 28, 2017, from http://www.sicherheitsforschung-magdeburg.de/uploads/journal/MJS_057_Maus_Segregation.pdf

1 Introduction

Cryptography can do much more than just the common HTTPS or the not so common e-mail or disk encryption. It can do even more than just keeping confidentiality and proving authenticity of origin and content.

Neglected in the shadows cast by the ubiquitous »confidentiality, integrity, availability« lies a wealth of security dimensions well worth considering:

- Multilateral Security – multiple parties with conflicting interests might need to establish »checks and balances« or explicitly shared control over critical objects or processes. Imagine some competing companies, sharing a common online platform – e. g. for regulatory reasons – where they offer their services.
- Segregation of Duties – certain transactions might be so critical, that even within one party they should depend on the joint decisions of several different persons. The authorization of large loans or investments are typical examples.
- Verifiability and Auditability – it might be necessary to prove the correctness of the work-flow and decision-making.
- Exoneration Capability – independent of the question of correctness, anyone wielding significant powers should be interested in being able to prove, which actions were done with these powers, and – equally important – which not. A classical example are system and DB administrators, which want to prevent both false accusations as well as leading the list of the usual suspects, whenever some kind of cyber-crime happens in their area of responsibility.
- Accountability – essentially the other side of the exoneration capability coin: the attribution of actions to specific actors and ideally the proof of an act of volition (as opposed to error), providing a base for (legal) non-repudiation and responsibility.
- Privacy – in many democracies a fundamental civil right, considered central for the functioning of democracies.[1] It has several independent aspects:
  - Transparency and control over data usage and processing for the subject.
  - Data minimization – more or less an application of the security strategies »need to know« and »minimal privileges« to personal data.
  - Non-traceability of activities of the subject, including robustness against inference and extrapolation – essentially the other side of the accountability coin, with obvious potential for conflict ...

Cryptography can go a long way in providing these security dimensions by enforcing properly defined work-flows, for example under a Clark-Wilson security model. This paper will illustrate some possibilities by example.

2 Audience and Preconditions

This paper is intended as an introductory reading. Besides interest in the topic no previous knowledge of cryptography is necessary (but not spoiling the fun either ;-).

The discussion will only make use of the most fundamental and generic cryptographic primitives. Therefore we need not to dig into the mathematical features of specific crypto-systems but can graphically illustrate the cryptography used. The following primitives are used – see figure 1 for their graphical representations (letters or symbols indicate different personas and their corresponding keys):

- Symmetric cryptography, also known as »shared secret«: encryption and decryption of data is performed with the same secret key which is shared between all participants. Everyone holding this key can encrypt and decrypt, so controlling the key is crucial and authenticity of messages can not be verified. Consequently history and historic novels are full of stratagems and disasters possible around these fundamental weaknesses of symmetric cryptography (besides deciphering the often weak historic ciphers directly).
- Asymmetric cryptography, better known as »public-key cryptography«: each participant has two personal keys. Each of these key enables a mathematical transformation of data, which the other key can reverse.
  - The »private key« must – as the name implies – be kept secret. The following operations are only possible for somebody possessing the private key:
    - construct the »public key« (The »private key« *must not* be derivable from the »public key«)
    - »sign« or better »seal« data[2] – »signing« is the term commonly used, but an unfortunate choice, in my humble opinion. It evokes wrong associations and the resulting misjudgements are at the core of many security incidents involving asymmetric cryptography: »signing« always evokes the image of a manual,

1 Look up the »Panopticon Effect« and the read-worthy Grounds of the Judgement of the German Constitutional Court (»Bundesverfassungsgericht«) from 15th December 1983, concerning a total census.

2 Actually not the data is »signed« or »sealed«, but a cryptographic fingerprint of the data, i.e. a function which computes some constant width number from documents in a manner that the slightest change in the document produces a different result and that it is not feasible to construct two similar documents with the same result. Constructing these functions is an art by itself.

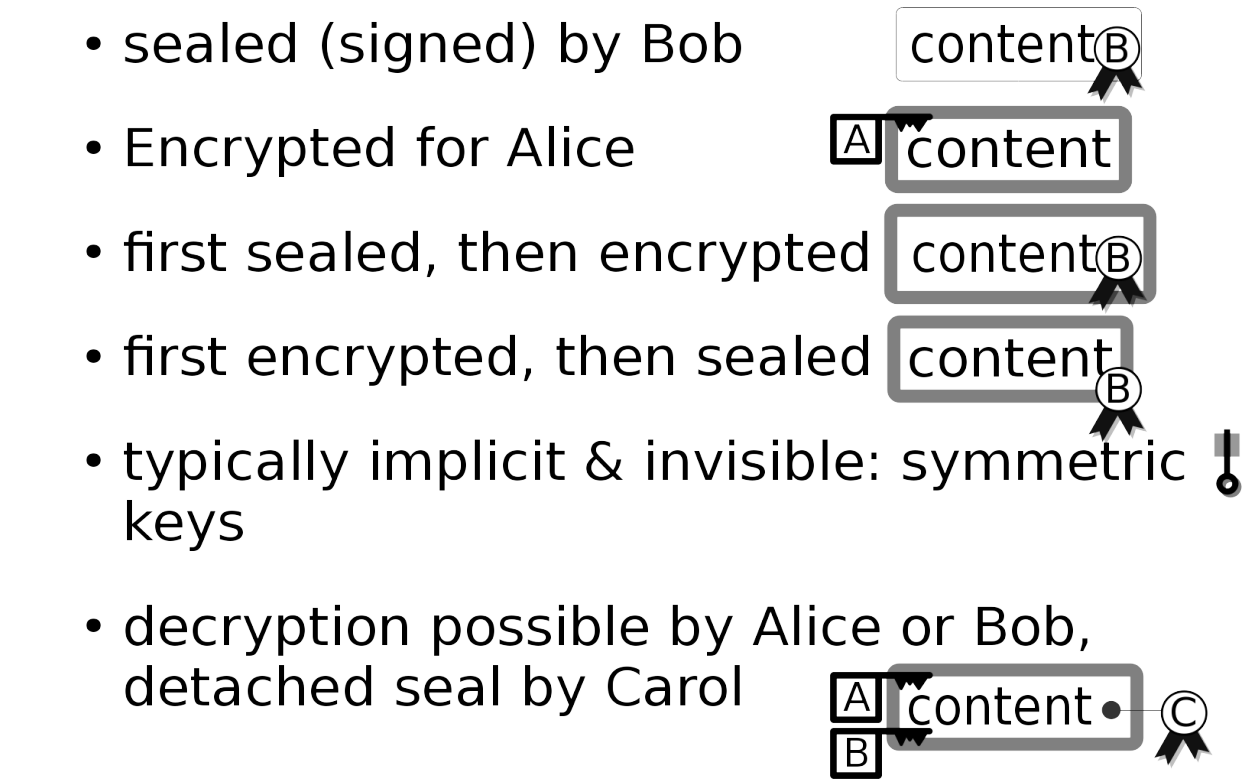


Figure 1: Visualized Cryptographic Primitives

immediate act of volition. But nobody actually can sign digitally by herself. Always some kind of computer system is acting as an intermediary. So the operation is much more comparable to the medieval »sealing« of documents, with a seal wielded by a seal-keeper, which might be misused, copied or stolen …

  - »decrypt« data – i.e. reverse the data transformation via the public key, which is called »encryption"

– The »public key« which – as the name suggests – can be publicly known. It allows for the following operations:

  - »encrypt« data[3] for the holder of the »private key« – aptly named, as everyone can »encrypt« but only the holder of the »private key« can revert the operation, i.e. recover the encrypted data
  - »verify« seals attached by the »private key« holder – again aptly named, as everybody can check the correctness of an operation, feasible only with access to the »private key« (seal)

3 Example Scenario: Data Retention

Our working example is essentially a science-fiction setting:

- The fictional part – with all artistic liberties – are the roles, stances, choices and requirements of a fictitious society, which decides to implement data retention with a wide range of possible privacy-enhancements, institutional checks and balances, as well as civil-society audit and potentially even veto rights.
- The scientific – and non-fictional – part is the use of established cryptography to explore the technical possibilities to achieve the issues raised by the fictional part.

This should provide a complex and challenging working example, in a scenario not needing much introduction, and hopefully of interest.[4] Further it motivates the exploration of multiple mechanisms and techniques: In the given setting, technology must not unduly limit the creative leeway and options of demo-

3 Again the truth of public-key cryptography is somewhat more complex: In practice all systems are hybrid, using a symmetric key to do the actual encryption of data, and only encrypting this symmetric key with asymmetric cryptography. (There are many good reasons for this approach, not the least: better security if done correctly!)

4 With the added bonus of avoiding the risk of disclosing secrets of my customers by trying (and failing) to disguise an actual use case beyond recognition …

cratic decision-making. The methods presented in this paper thus should accommodate manifold imaginable socio-political decisions. A wide range of reasonable work-flows should be possible with only minor adaptions owed to the inevitable limits of technology.

3.1 Here be Dragons ...

Before we set out, here is a little road-map for our expedition into the scenario.

The sequence of events is split into two distinct phases: In the *investigation phase* (see section 5.1), the investigators choose and narrow down the privacy-relevant records for which they seek access. This access is granted or denied according the defined procedures and applicable legal regulations, producing an audit trail.

After some time – in the *verification phase* (see section 5.2) – these audit trails are then verified. This should prevent misuse of the special powers, and provide exoneration for those wielding these special powers responsible and correctly.

Along this road the acting personas will encounter data in various »aggregate states«:

- *Clear* – the content is in plain-text, i.e. unencrypted, and directly usable.
- *Diluted* – while the data is in plain-text and usable, it has been preprocessed in such a way, that it will not point to identifiable persons, i.e. is not impairing fundamental rights. The diluted data is the instrument by which the investigators can do useful analysis without breaching into privacy, and minimize the number of requests for actual disclosure of personal data. Obviously a subtle balance needs to be found between accuracy and obscuring, to minimize the privacy disclosure requests as well as the privacy risks by inference and extrapolation. This will be explored in section 4.
- *Opaque* – the content is encrypted and unusable for personas without access to the key, but as an unique identifier and an object of power of disposition. This type of data is typically used for information pertaining to fundamental rights, which is made accessible only via safeguarded procedures.

3.2 Dramatis Personae

The fictitious society of our working example is a constitutional democracy with politically participating citizens (»citoyens«) They see the protection of fundamental civil rights as foundation of democracy and are vigilant about the panopticon effect as well as crime prevention and prosecution. The obvious conflict is resolved by bestowing special powers under special precautions. The correct exercise of office by representatives and officials is generally monitored to preserve confidence of the society in its institutions. But the conflicting needs of criminal investigation and fundamental civil rights are considered especially sensitive and the special precautions should equally prevent the misuse of these powers as well as false accusations against the wielders of these powers.

The following personas – their respective symbols for the illustrations are given in the descriptions – are actively involved in the special procedures and thus our example work-flow:

3.2.1 Telecommunication Service Providers »✆«

The telcos have the obligation to prepare and provide the legally required data structures for the procedure.

Otherwise they are interested in minimal involvement and in holding only the minimal set of data needed for billing for the shortest possible time. The main motivation for this data minimization is to reduce the risk of any breeches of laws on privacy of correspondence, posts and telecommunications, and especially to avoid any allegations of short-cutting the checks and balances of the legal procedure by directly giving data to governmental authorities.

As they are competing in a regulated liberal market, the citizens as consumers wield immediate and existence-threatening power over the telcos: the citizens can shun those with a suspicious privacy or security track record, and generally prefer those with a security and privacy stance in line with their own.

3.2.2 Investigative Authorities »👁«

The investigative authorities are tasked with crime investigation for prevention and prosecution of crimes. For this purpose they may invade the privacy of individuals in accordance with legal procedures.

Their main concerns are:

- Tactical secrecy of the investigation
- Earning and keeping public confidence
- Auditability and lawfulness of all investigative activities invading fundamental rights of individuals.

3.2.3 Examining Magistrate »⚖«

It is the responsibility of the Examining Magistrate to decide if individual disclosure requests are granted within the legal framework and the context of the investigation in question and its results so far. The magistrate has to make an assessment and a fully informed decision if the legal priority of fundamental rights is guaranteed in this criminal investigation.

The main concerns coincide with those of the investigative authorities:

- Tactical secrecy of the investigation

- Earning and keeping public confidence for the office and the officers holding it
- Auditability and lawfulness of all decisions, either supporting or declining disclosure requests.

3.2.4 Federal Privacy Commissioner »☂«

Besides the other responsibilities of a federal privacy commissioner within the context of criminal investigations the following obligations arise:

- Formal control of disclosure requests – i.e. the protection of fundamental privacy rights within statutes and without knowledge of investigative results
- Official auditing, statistics and reporting
- Special checks, verification, and information in special cases, e.g. medical doctors, lawyers, priests, ...
- Official investigation of complaints
- Destruction of the private key of an office in certain cases as a constitutional safeguard, for example upon changes in certain constitutional clauses or in laws applying to the office or a coup d'état or ...

Earning and keeping the public confidence in the office and the officer is a major concern for the Federal Privacy Commissioner.

3.2.5 Representatives of the Civil Society »👥«

The civil society represents itself in this procedure by specifically elected representatives. Many forms of democratic participation are conceivable – and this is the area where the flexibility of the demonstrated approach has to prove itself.

We will start with simple observer roles, and then explore various scenarios of direct policing or juror/assessor roles, including quorum decisions, ranging from hard to soft decisions with various forms of graduated denial of disclosures.

4 Dilutions and Pseudonyms

Many personae in our example will work with diluted data, providing information necessary for a specific step in the work-flow, but not yet disclosing an individual. The most pronounced usage of diluted data occurs through investigative authorities – so this will be the best example for discussing the topic.

The data prepared by the telcos according to the legal requirements is immediately transferred to the investigation authorities (and erased at the telcos). Thus all processing can be done within the investigative authorities, which provides several advantages:

- Protection of the tactical secrecy of investigations should be simpler.
- Advances in search technology can be introduced with less need for coordination.
- The retention periods of data records can – within the limits of statutes[5] – be handled more flexible and intelligently: While records pertaining to suspects or crime hot-spots could be hold for the maximum time, all other data might be erased after a few weeks – a cost-driven, but probably innoxious optimization.

The records provided by telcos handed over to the investigators have the general form »handle → opaque data«, where »handle« is diluted data enabling useful investigative activity but not yet giving away information to identify individuals hidden in the »opaque data«, which is only accessible via disclosure work-flow. So, based on the standard investigative approaches, the »handles« should allow a – ideally quite narrow – preselection of records for disclosure requests in an indiscriminative manner, which could be considered a value in itself.

4.1 Diluted Data

A thin line must be walked between providing enough details for meaningful selection but not yet disclosing the protected information within the »opaque data«. To illustrate this we construct a too specific set of handles, enabling the identification of the speaking parties of a specific conversation without the need to use an extra disclosure procedure. Imagine the following records:

- **handle**(calling id, precise start time, precise end time) → **opaque** (called id)
- **handle**(called id, precise start time, precise end time) → **opaque**(calling id)

By correlating time stamps between these two sets of records, it would be easy to infer the pair of calling and called IDs, and thus identify the speaking parties, bypassing the disclosure work-flow.

Now consider instead the following records:

- **handle**(pseudonym(caller), diluted start time, diluted duration) → **opaque**(callee)
- **handle**(pseudonym(callee), diluted start time, diluted duration) → **opaque**(caller)
- **handle**(diluted location, diluted start time, diluted duration) → **opaque**(subscriber)

The purpose of dilution is to avoid correlation of records and inference. So it should provide that each »bucket« contains enough[6] records, but no meaningful inference is possible.

An effective dilution of start time will depend on the time-of-day: while around midday rounding to

5 The safeguarded workflow prevents disclosure requests for too old records, and a re-pseudonymisation process (discussed later) obsoletes them ...

6 »enough« is a political decision in striking the balance between investigation effectiveness, as well as privacy both on the level of diluted records and the number of disclosure requests to be granted.

minutes might be sufficient, at other times a granularity of 5 minutes or – in the hours of the night – even 15 or 30 minutes might be necessary.

Diluted durations might be by the minute or an enumeration like {»not answered«, ≤1 minute, ≤2 min., ≤3 min., ≤5 min., ≤10 min., ≤15 ...}.

Location dilution would depend on the area and the time-of-day: In a thinly populated rural area identifying a cell base station might be to concise. The same might be true of downtown in the dead of night or a dormitory town at day. During office hours on the other hand, a cell base station in downtown would create very full buckets, and a much finer granularity coordinates in the order of tens of meters might be adequate.

4.2 Pseudonyms

While diluted data aims at clustering individuals in groups precluding individual inference, pseudonyms should be opaque identifiers of individuals. They provide accountability but the identity of the individual should only be disclosed by our safeguarded disclosure work-flow.

Now pseudonyms are a quite ticklish topic: studies showing how pseudonyms are broken abound. Essentially there are two main pitfalls:

- Reversible or enumerable pseudonyms – you wouldn't believe how often I have encountered in the field (and in critical areas) »pseudonyms« constructed from the date of birth, the postal code of current or birth residence, an indicator of sex and a collision resolution counter, or something similar. Very obviously you can guess the »pseudonym« for a specific person and with high probability you can identify any person from their »pseudonym«. A common »countermeasure« to this problem is, to use a hash value or cryptographic fingerprint derived from the badly chosen »pseudonym" instead. As this functions are (mathematically) not reversible, the »hashed pseudonym« is – falsely – deemed safe. The error in reasoning is, that it is still possible to guess the »pseudonym« of a specific person and compute the corresponding »hashed pseudonym«. Worse: it is typically trivial and well within computational reach, to construct all conceivable »pseudonyms« of real persons and construct a dictionary, which allows to revert all »hashed pseudonyms«, thus providing the identifying data from the underlying weak clear-text »pseudonym«. We see: effective pseudonymisation is no trivial topic!
- But even perfect pseudonyms might be revealed by accompanying data: If a pseudonym gets regularly used in a certain area at night, this probably might be its owner's area of residence. If two pseudonyms are regularly used at night either at one place or another, that probably indicates a pair of lovers. So our dilution of data has to take into account which pseudonyms are used and how, to still provide the intended level of privacy and investigatory capacity.

Further the investigators might arrive at additional data through their legal investigations in the »real world«, which could expose pseudonyms, too. And this effect might actually be either intended or unwanted by our fictitious society.

In response to this various requirements the pseudonyms could be varied in the following dimensions:

- Granularity – the calculation of pseudonyms might be parameter-based. Pseudonyms might be only constant within specific conversation pairs, or they might depend on call direction or on location areas of different sizes: country, state, district, postal code or base station. To illustrate the range of possible granularities, here are some examples:
  - the pseudonyms of caller and callee might be specific for a conversation pair A ↔ B independent of call direction: if A calls B or B calls A, both will have constant pseudonyms. But if A calls C or is called by C, A will be listed under a different pseudonym.
  - the pseudonym of a mobile subscriber might change with the location of the call with some granularity like country, state, district, postal code, base station, ...
  - pseudonyms could depend on time-of-day with some granularity or day-of-week ...
  - or combinations thereof, for example if somebody from Vienna calls somebody in Graz, caller and callee will have constant pseudonyms. In the opposite direction different pseudonyms will be assigned, as well as A's pseudonym in Vienna changes if, instead of somebody in Graz, somebody in Salzburg is called.
- Durability – pseudonyms might change at intervals or event-driven. The change could be triggered for example by the Federal Privacy Commissioner with a specific delay when a pseudonym is officially disclosed and this delay may depend on the individual identified, e. g. faster for journalists or a lawywer's or doctor's office. The periods of re-pseudonymisation might depend on the pseudonyms owner and might be strictly regular like »at the start of every quarter« or slightly randomized as in »every 8–10 weeks«.
- Scope – we do not need to use the same pseudonym type all over. Each type of record might have it's specific pseudonym type. On the other hand, we could purposely use the same pseudonym types in different places to explicitly enable powerful investigations on the pseudonym level, if our main concern would be to prevent ill-founded surveillance. Then, the investigators might even have records of the form:

  - **handle**(**pseudonym**(caller), diluted start time, diluted duration) → **pseudonym**(callee)
  - **handle**(**pseudonym**(callee), diluted start time, diluted duration) → **pseudonym**(caller)
  - **handle**(diluted location, diluted start time, diluted duration) → **pseudonym**(subscriber)
  - **pseudonym**(subscriber) → **opaque**(subscriber)
- Visibility – we can decide at which points in our work-flow which kind of pseudonyms will be opaque or visible in the clear to which personae. So the investigators might only see short-lived pseudonyms unique per contact, while the examining magistrate might have access to more long-lived and globally constant pseudonyms. The investigators then may provide a query to the examining magistrate which selects the records for disclosure with pseudonyms visible to the magistrate ...

This should enable us to have varying degrees of pseudonymity[7] and a graduated resistance against de-pseudonimisation, which in turn enables a wide range of designs.

Again, the actual choice is a socio-political democratic decision of our fictitious society. But it would be well advised, to spell out the decision in all detail – in terms of what is intended and unwanted – and augment this decision by a funding of continuous research if the goals of the decision are actually achieved in the actual implementation!

5 Disclosure Procedure

Now all is set to explore the ideas within our scenario as an extensive, realistic[8], and non-trivial example. For sake of simplicity we start with a representative and straight-forward work-flow outline with only some small variations included. We explore extensions and fundamental variations to this scheme in the section 6.

The basic work-flow consists of two major phases:

- the investigation phase, during which investigators over time identify records of which they want the corresponding individual disclosed and these disclosure requests processed in (partially) blind trust (i.e. keeping the tactical secrecy of the investigation)
- the verification phase, where the disclosure requests are reviewed for validity with the reviewers having complete access to all relevant information

5.1 Investigation Phase

The main topic of this phase is to control who when will have access to which information. It is about minimizing information access, while at the same time providing all knowledge needed for indiscriminative decision-making, and about guaranteeing the accountability and auditability both of access and authorization of access. These topics – especially the decision-making without regard to the specific person – are relevant beyond our example in which we exercise them now in the following subsections.

The essential steps of the investigation phase are illustrated in figure 2 – clarifying the encryption and sealing details.

5.1.1 Investigative Authorities

The investigators have unlimited[9] access to all kinds of data records in the form of:

- **handle** → **pseudonym, opaque_data**

This is a slightly generalized form compared to our discussion in section 4, but either pseudonym or opaque data could be empty, resulting in the limited cases discussed there. Handles and pseudonyms are crafted to provide the level of privacy and immediate investigatory power our fictitious society believes to represent the intended balance.[10] The investigators select opaque data sets for further disclosure by using the handles and pseudonyms.

They build, sign, and forward to the Examining Magistrate – as often and whenever needed – disclosure requests stating:

- urgency
- reasons for requests
- optional further selection criteria (more about this soon)
- set of opaque records to be disclosed
- optionally further tactical considerations

7 Justifiably you might be sceptical, if all these features could be achieved in pseudonyms. Let's dive into technicalities and consider the following: $key = \mathrm{HMAC}(Nonce_{Interval,Provider}, Parameters)$, where Parameters represents all the granularity variables we want to factor into pseudonym computation, $Nonce_{Interval,Provider}$ is a secure random number, specific to the telco and changed at intervals – one mechanism limiting durability. A second durability limit is $Nonce_{Subscriber}$, a secure random number, specific to the individual subscriber. This will be renewed under certain conditions in our workflow. We then calculate the actual pseudonym as $\mathrm{Pseudonym} = \mathrm{encrypt}(key, Nonce_{Subscriber})$ I'm convinced that the construction enables all the claimed features, but you are sincerely invited to challenge this construction, and to improve it.

8 »Realistic« not in the sense of »ready for real-world application«, but in the sense of sufficiently complex to give room to all kind of different motivations, considerations, and constraints which are to be expected in any practical application of the methods presented.

9 Of course, there will be access restrictions within the investigative authorities. But from the perspective of the civil society and the Federal Privacy Commissioner these internal controls do not implement an external and verifiable control, which allows exoneration. From this external point of view the investigative authorities as a whole have unlimited access – independent of how the internal access is structured.

10 The balance perceived as adequate of course depends on the confidence of the society in its authorities. Thus, obtaining and preserving a high level of confidence is in the interest of the authorities and the intended primary benefit from our example disclosure procedure.

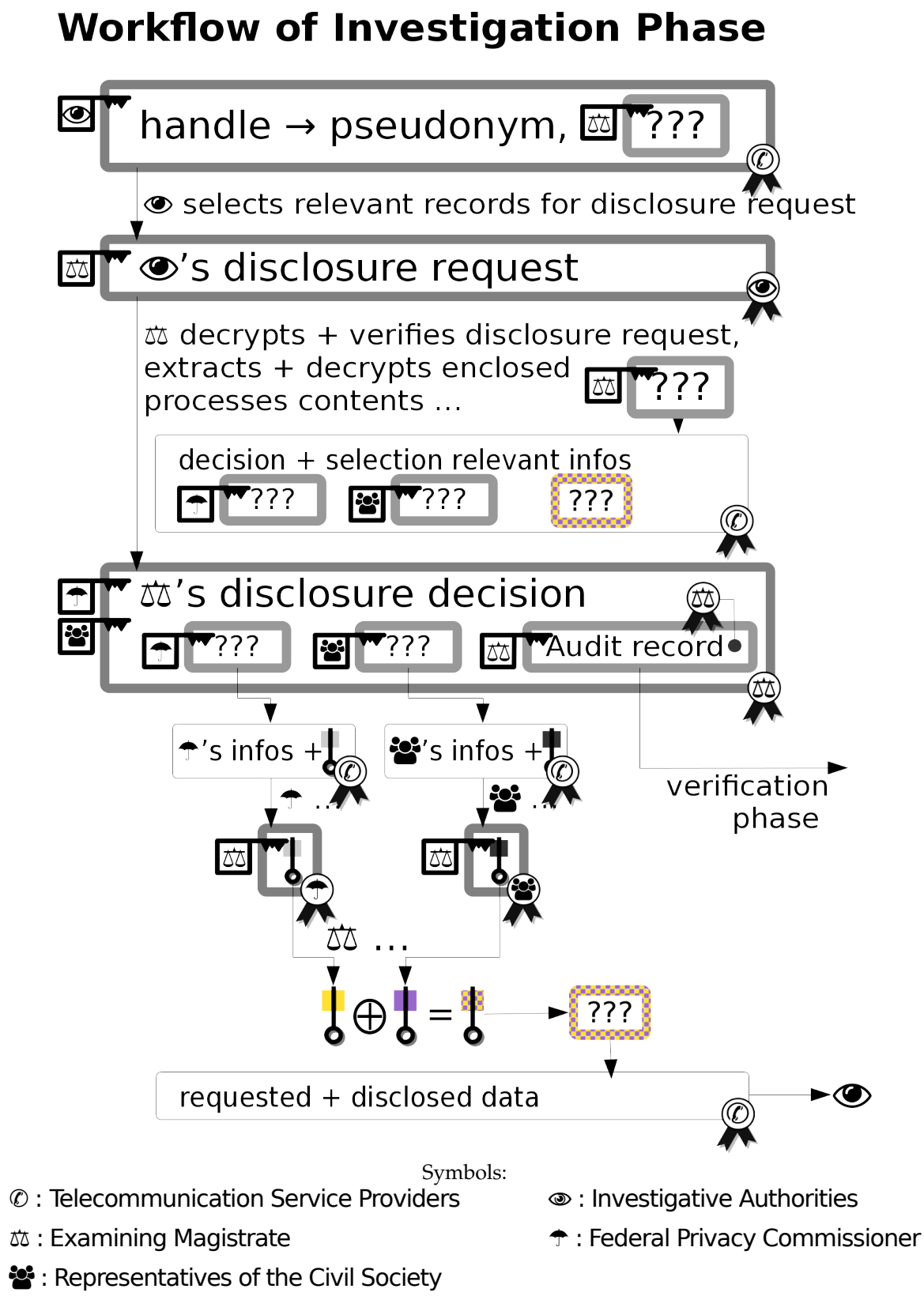


Figure 2: Workflow of Investigation Phase

5.1.2 Examining Magistrate

The Examining Magistrate decrypts and verifies the investigator's disclosure requests in full knowledge of the state of investigation. Decryption of the opaque records contained will further reveal

- decision-relevant information about the (still unidentified) subjects – e.g. medical or criminal records, records of emergency services, ...
- potentially more significant pseudonyms and less diluted data to augment the investigatory power, if it is constrained on the level of the investigative authorities – the optional further selection criteria provided by the investigators will be applied to this data to further narrow down the actual disclosure requests to the next levels.
- opaque records – decryptable by the Federal Privacy Commissioner and the delegates of the civil society, respectively – to be propagated to the next work-flow steps as part of the forwarded disclosure requests
- an opaque record, containing the actually to be disclosed data, encrypted symmetrically with keys under multi-party control, i.e. both the Federal Privacy Commissioner and the delegates of the civil society have to provide key material to arrive at the actual key and enable decryption

The Examining Magistrate may narrow down the selection of actual disclosure requests – either because additional information showed that specific records are not eligible for disclosure according to the statutes, or because of the optional selection criteria. Then the magistrate's disclosure decision is prepared and submitted, including:

- Urgency.
- The reason for the Examing Magistrat's decision in clear, but short form for tactical secrecy.
- A »decision audit record«, containing the complete reasons for its decision (including all facts) as well as the investigator's disclosure requests, is prepared, self-encrypted, and sealed with a detached seal.
- An indexed list of all opaque records selected for actual disclosure.

Further, the Examining Magistrate keeps an indexed list – corresponding to the indexes in the disclosure decision – of the opaque records under multi-party control for later decryption if the disclosure decision is granted.

To avoid any unnecessary delays, the Federal Privacy Commissioner and the delegates of civil society can work in parallel on these requests.

5.1.3 Federal Privacy Commissioner

The Federal Privacy Commissioner decrypts the disclosure decision, as well as the opaque records within – as far as decryptable with the commissioner's key –, which each contain

- decision-relevant information about the (still unidentified) subjects – e.g. medical, criminal or emergency records, ... (possibly in a different information granularity from e.g. the Examining Magistrate's)
- potentially more significant pseudonyms and less diluted data to augment the investigatory power, if it is constrained on the level of the investigative authorities and the Examining Magistrate – the optional further selection criteria provided by the investigators will be applied to this data to further narrow down the actual disclosure requests to the next levels.
- Federal Privacy Commissioner's key material within the multi-party control scheme for the respective record in custody of the Examining Magistrate

The Federal Privacy Commissioner then makes a purely formal – probably even automated – decision about the validity of each disclosure request based on the visible data. Statutory periods and subscriber criteria might be checked, for example. The Examining Magistrate gets an encrypted and sealed indexed list of key material for granted disclosures or justifications for denials.

Further the disclosure requests and the information provided therein are used within the statistics for periodic reports of the Commissioner, and might trigger specific actions like special audit watch lists, notifications to specific institutions – medical boards, attorney bar, ... –, and re-pseudonymisation orders for subjects.

5.1.4 Delegates of Civil Society

A vast range of possible roles, functions, and powers could be imagined for the delegates of the civil society – and we will do so in section 6 in order to explore complex decision-control schemes.

But first we explore a simple and pure – but interesting – observer role, which would be a very conservative approach to the role of the delegates of the civil society. Their sole purpose is to monitor all disclosure requests, guarantee their later review and report independent statistics from the Federal Privacy Commissioner (providing control over and confidence in this office). Each delegate can provide the same key material to the Examining Magistrate – so a single delegate granting the request is sufficient for the work-flow to continue.[11]

The delegates decrypt the disclosure decision, as well as the opaque records within, which each contain

- statistics-relevant information about the (still unidentified) subjects – e.g medical or criminal records, records of emergency services, ... (possibly again with different information granularity or even no information at all)

11 Thus vastly reducing the risk of a single or few delegates obstructing investigations for egoistical reasons ...

- their (shared) key material within the multi-party control scheme for the respective record in custody of the Examining Magistrate

The Examining Magistrate gets – more or less automatically, as there are actually no decisions to take – in return an encrypted and sealed indexed list of key material for the requests.

The delegates are free to inform the general public or professional boards (if these are not delegates themselves) of any suspicious observations, as for example a surge of disclosure requests for journalists and their contacts during some »whatever-gate« scandal. The issue could then be publicly discussed and resolved, while individual citizens could take precautions.

5.1.5 Examining Magistrate

The Examining Magistrate receives and pairs the key-material from the Federal Privacy Commissioner and the delegate of the civil society. Now all opaque records in custody of the magistrate, for which the disclosure request was granted, can be decrypted and the data pertaining to the actual subject forwarded to the investigative authorities.

With the investigation progressing the investigators can clear innocent bystanders, and the Examining Magistrate – after verifying that the subjects are not locked by other investigations – can issue re-pseudonymisation orders, protecting their privacy again.

5.2 Verification Phase

This section is concerned with the auditability and review of procedures and subsequently with their regular verification, as would be of interest in the areas of auctions, tenders, . . . , or – of course – in our working example:

The Federal Privacy Commissioner and the delegates of the civil society have the right to verify all disclosure requests and decisions either after a statutory period or if the investigation is closed or the case tried, depending on the statutes.

To that end the responsible Examining Magistrate has to decrypt each body's copy of its decision audit record – the authenticity of decrypted content can be verified via the detached seal – for review by both bodies.

After the review of the decision audit record the officials responsible for the investigation are either exonerated or impeached – depending on the results of the review.

Further the Federal Privacy Commissioner and the Delegates will verify re-pseudonymisations and have the right to initiate them on their own.

5.3 Constitutional Safeguards

Another aspect are instruments of control which enable parties in a multi-party work-flow to prevent violation of their own interests by other parties. In this section we look at our example work-flow from this perspective:

The procedure illustrates many options of intervention and constitutional safeguards, should special governmental powers be put to use in a dubious way on a small or large scale.[12] Actually there is a scale of graduated reactions:

- Individual disclosure requests can be denied (and even that can be graduated as we will see in section 6).
- Orders of re-pseudonymisation can be given frequently. So during a »whatever-gate« scandal it might be considered an useful ad-hoc protection to daily re-pseudonymise journalists, lawyers, or medical doctor offices. The power to do so might also be granted to the delegates of the civil society as well as – for example – the bar associations, the medical doctors, and the journalists for their respective profession.
- The keys of the federal privacy commissioner and the delegates of the civil society can be destroyed, effectively invalidating all available opaque records – an extreme measure and act of civil courage even if legal, reserved for extreme cases where prosecution of dissidents of a new regime has to be feared . . .

6 Complex Decision Modes

Let's explore some more complex decision modes, like quorum and majority decisions or graduated denial.

To that end within our example scenario we grant the delegates of the civil society actual statutory and political powers: denying disclosure requests and negotiating the conditions of disclosure, up to potential roles as examining jurors or assessors (obviously sworn to secrecy as they would share the point of data visibility of the Examining Magistrate). Obviously such far-reaching powers may not rest on a single person but need some peer control, i.e. some kind of quorum-based decisions (which depending on the quorum of course can implement majority decisions).

12 In my humble opinion, history is full of constitutional crises, coup d'états, game-changing landslide victories – especially German history contains some stark warnings. All of these events were catastrophic enough without modern surveillance capabilities in place. So – hoping that the necessity of use never arises – there is the necessity of provision of means to at least temporarily limit the governmental surveillance capabilities in a democratic crisis.

6.1 Quorum Decisions

For a quorum decision to pass, a minimum number of delegates have to support it.

In section 5.1.4 our decisions amounted to a quorum of 1 out of N delegates for granting, and unanimity to deny disclosure. Let's assume we have 3 delegates and any 2 out of 3 majority will decide.

A simple, straight-forward approach to implement such a 2/3-quorum is presented in figure 3. The quorum-protected data is enclosed in two layers of encryption: the inner one decryptable by two parties, the outer one by the last party. A second key-container, encrypted similar but with differently distributed parties, is needed to guarantee that if the »outer party« of the first key-container denies, but the two »inner parties« grant access to the data, this access is actually possible. This easily extends to a quorum 2 out of 4, and it is obvious that at least two parties have to grant access.

Alas this simple-minded approach does not scale well: three key-containers would already be needed for a quorum 2 out of 5 or 6 – and because of the combinatorial explosion the number of needed containers raises swiftly when the threshold of the quorum and the number of delegates increases. So while it is an useful solution for small numbers, it quickly becomes clumsy for larger quorums and numbers of parties.

Fortunately there exist sophisticated solutions for this problem – also known as »secret-sharing« or »split key« – which scale better and are proven to be correct:

- *Shamir* – based on polynomials and providing perfect secrecy of the key material as long as the quorum threshold is not reached: i.e. even with one vote missing to the quorum, absolutely no useful information about the key material is gained.
- *Blakley* – based on the intersection of hyperplanes, each »pro« vote actually reduces the search space for the correct key, producing a unique solution if the quorum threshold is met. The simplest implementation of Blakley: simply break the key in equal parts and distribute it between the parties. This clearly illustrates a weakness of the approach: It might well be feasible – depending on the dimensions of the hypercube containing the planes – to recover the key by brute-force if a few votes are missing, because the remaining search space is computationally feasible.

This imperfection of the Blakley algorithm might be considered advantageous or detrimental – in the later case: stick to Shamir's algorithm. On the other hand, it allows to deliberately construct a system, were consent can be substituted to some extent with computational effort. In other words: lacking consent produces additional effort, forcing either more compliant behavior or reducing the throughput of only partially approved work-flows. This opens up the possibility of more political »soft« decisions and negotiations, which might be a value in itself, in order to handle unforeseen situations – we will explore this further in the next section.

6.2 Graduated Denial and political Power-Play

Now imagine within our example scenario, that the delegates of the civil society should have a democratic mandate for the political assessment of investigatory needs and civil rights, and should quickly decide according to the specific situation in all those cases not adequately provided for in the statutes – and these are bound to occur!

In »real life« decisions are not always mathematically clear and hard choices. There is often the need to demonstrate disapproval and reluctance. If appeals are not sufficient to actually stimulate the intended behavior, demonstrations of power – especially to stifle the objected behavior – are needed. Such powers of course need fine-grained democratic control and precautions against a blockade by a fundamental fraction.

The Blakley algorithm could be used for this, as sketched above, but it cannot provide the complete key without any delay if a single party blocks. If we want the capability to retrieve the key without any penalty if – for example – q=25% of the delegates deny access, we need something else.

Consider the idea illustrated in figure 4.[13] We create a set of random bits r. From these – by an appropriate key derivation function (KDF), introducing a carefully chosen computationally work-load, a symmetric key can be derived. Now we feed this set of random bits r into an error correcting code (ECC), designed to exactly recover up to q missing bits. The resulting secret is sliced into segments, which are signed[14] and attributed and encrypted for each delegate, and finally committed to the delegates.

Now the delegates could individually vote on all controversial disclosure requests. If – for example – during a »whatever-gate« scandal delegates set triggers to vote manually if disclosure requests for journalists (or their contacts) occur, they can immediately politically take influence: a rising number of withhold segments indicates the loss of confidence in the investigative authorities and the Examining Magistrate, and might lead to caution. But as long, as the denials stay below q (in our example 25%), the investigation is not hampered in any way. If the threshold q is crossed, the error correction capacity of the ECC is exhausted. The exact impact of this depends on various factors:

- Procedural definitions – the key segments may be immediately forwarded to the Examing Magistrate, who then can apply any available com-

13 I've not found this idea in literature so far, but my search was not exhaustive …

14 To prevent a non-cooperative delegate to return a manipulated fragment and claim it to be authentic.

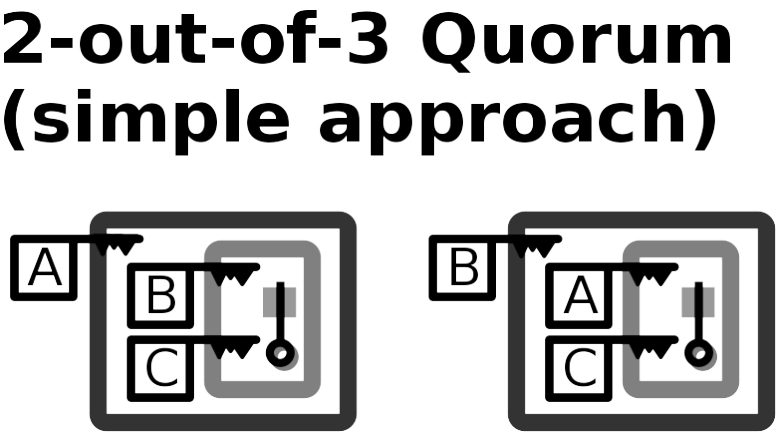


Figure 3: 2-out-of-3 Quorum (simple approach)

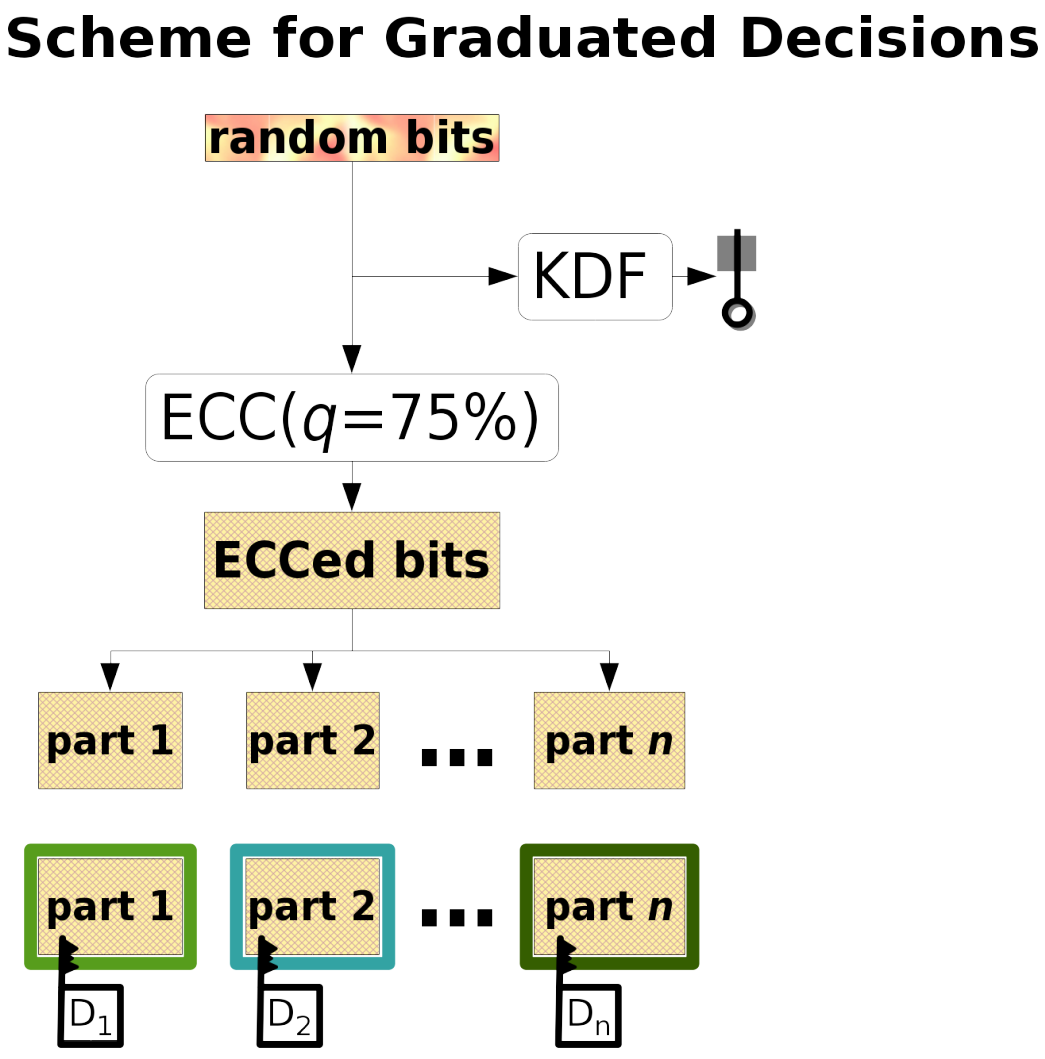


Figure 4: Scheme for Graduated Decisions

putational power to brute-force. Or they are first collected in a neutral clearing institution – e.g. chair of the delegates or Federal Privacy Commissioner –, where they are withheld if the quorum q is not met. The Examining Magistrate then may order brute-forcing, which is then done only with a well-defined computational power at the clearing institution.

- Key-stretching within the KDF – the key deriving function can afflict each attempt to guess a key with a well-defined computational effort.
- The number of missing bits m has two effects:
  - It sets an exponentially growing upper bound on the effort for naive brute-force (in the order of 2^m).
  - The deliberate choice of ECC has a marked effect, too. Some ECCs provide for »short-list decoding«, where from the inner coherence of the ECC a list of possible candidates for the missing bits can be generated. The generation of this »short-list« of course comes at a computational price which approaches brute-force effort the higher m becomes. This feature allows to adjust the impact of missing bits to some extent.

Between all these choices it should be possible to construct a system, where the rejection of a request by some fraction of the delegates has no effect at all, while a rejection by a certain quorum, e.g. majority, effectively denies access. Between these thresholds a ramp of penalty of effort exists – and can to some extent even be shaped. The result is, that the Examining Magistrate either is forced to select the few most important records for brute-forcing the missing consent, or to negotiate more consent.

There is a nice paradox here: Implementing these powers on the technical level will require quite some effort. But on the political level their pure existence will ensure, that they will practically never be exercised.

7 Conclusion

The example hopefully served its purpose as a vivid, entertaining, and instructive scenario to demonstrate a wide range of techniques.

Obviously not each and every imaginable work-flow can be mapped with these techniques. But it should be as obvious that – with some thoughtfulness and flexibility to the inherent limits of technology – most practically interesting work-flows could be easily adapted to take advantage of the demonstrated possibilities for cryptographic enhancement of their security.

8 About the Author

Thomas Maus holds a graduate in computer science. He is consulting in the areas of system security, the analysis, tuning, and prognosis of system performance, as well as the management of large, heterogenous, mission-critical installations since 1993.

Projects range from architecture, implementation and operation of large application clusters over technical project management, organisational and technical trouble-shooting, security assessments, establishing of security governance processes, security policies and analysis for trading rooms and the like to training of international police special forces for combatting cyber-crime.

He started his computing career 1979, at the age of sixteen, when winning the computing equipment for his school in a state-wide competition. Soon followed the teamworked development of a comprehensive SW for school administration on behalf of the federal state – here a long lasting affection for questions of system security, performance and architecture started. Around 1984 he fell in love with UNIX systems and IP stacks and embraced the idea of Free Software.