Privacy vs. Data: Business Models in the digital, mobile Economy

Lecture 9 + 10
Privacy & Privacy Protection

SS 2015

Dr. Andreas Albers

Agenda

- What is Privacy?
- Data Protection Directives and Laws
- Technical Data Protection
- General Requirements for Privacy
- Selected Privacy Concepts
- One more thing about Privacy

What is privacy?

- Privacy (from latin: separated from the rest, deprived of something, esp. office, participation in the government", from privo "to deprive“)

- Some definitions …
  - „… right to be left alone” [Warrend and Brandeis, 1890]
  - “… right not to be annoyed” [Varian, 1996]

- But there are many more … and privacy is a very complex, multi-disciplinary concept …
  - Technical, economic, legal, socio-economic, philosophic aspects

→ So, no working definition here …

What is privacy?

- Privacy has multiple dimensions …
  - Technological, economical, social, legal
- … and multiple stakeholder perspectives
  - Users, online businesses, regulators, public authorities, etc.
- Privacy for users is highly individual and may depend on
  - Usage context (e.g. user location, application, personal data)
  - Online experience and past privacy violations of a user
  - Cultural background and privacy attitude of users
- Privacy Protection as a challenge for individuals
  - takes effort, knowledge and often technical understanding
  - is not directly rewarding (short term) and not perceivable (Privacy Calculus)
  - is often demanded by many, but without the willingness to take the effort (Privacy Paradox)
  - can most likely never be outsourced or automated
  - but can actively be enabled and its effort minimised

Why privacy?

- Societal perspective
  - Foundation of democracy
  - Freedom of speech
- Individual perspective
  - Free personal development
  - Ownership of personal data of any kind
- However, in an information society, it always takes effort for individuals to protect their privacy.

Why privacy? Parallels to political instable regions of the world?

- Political instable regions of the world
  - Enterprises hesitate to invest and develop their business because they are afraid of losing it again soon

- Individuals without privacy
  - Individual hesitate to develop personally because they are afraid of being observed/surveilled and may experience consequences from this act

Privacy Online vs. Offline

- Offline Privacy
  - In the offline world individuals are mainly able to maintain their privacy intuitively

- Online Privacy
  - In the online world, privacy
    - has to be maintained through often complex privacy settings or identity management systems
    - often cannot be maintained at all by individuals because personal data is collected even without their knowledge or consent

Challenges for Privacy in the Online/Mobile World

- The Internet does not forget or is sometimes not allowed to do so (data retention).
- The Internet allows to easily connect social roles or partial identities, which would have been otherwise separated in the offline world.
- Profiling is easy and can be done automatically. In contrast, managing personal information is complex and has to be done manually.

Privacy Protection

- Data Protection Law (EU / Germany)
- Technical Data Protection
- Privacy by Design & more

- Identity Management (see Lecture 11)

Agenda

- What is Privacy?
- Data Protection Directives and Laws
- Technical Data Protection
- General Requirements for Privacy
- Selected Privacy Concepts
- One more thing about Privacy

Term “Data Protection”

- Definition
  Measures for the protection of stored and transferred personal data against manipulation or misuse; Federal Data Protection Act in place since 1978 (amendment in 1990).

- Originally for the protection of the citizen against governmental institutions.

- Businesses are regulated with regard to special aspects (telecommunications, medicine) of data protection.

- Increased need for regulation due to the use of information technology (data warehouses, globalisation of information processing).

Data Protection Principles in the Context of Communication Services

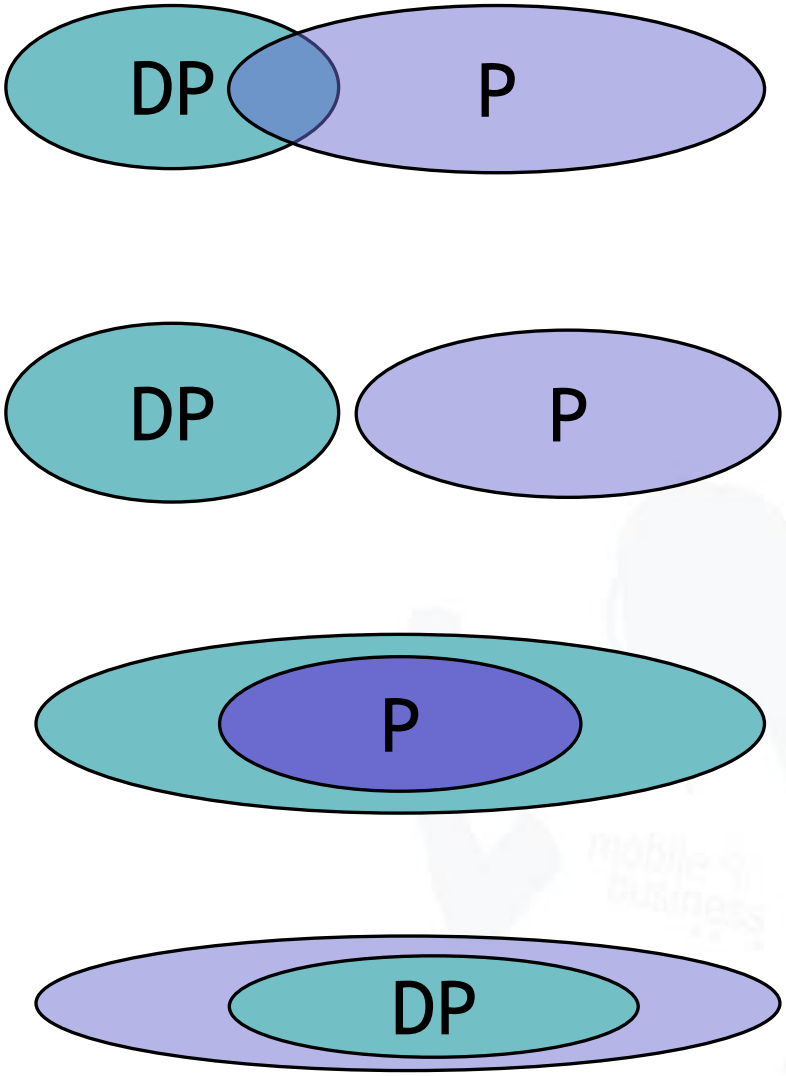
- **Data minimisation:**
  The service should be offered with a minimum of needed data.

- **Information of data subject**:
  The person, whose data is being stored, should know what has been stored.

- **Acceptance not without consent:**
  The data subject is to be asked in advance.

Privacy vs. Data Protection (Working Definition)

- Privacy (Individual)
  - Non-Absolute
  - Contextual
  - Relational
  - Opacity of the Individual
- Data Protection (General)
  - Procedural Safeguards
  - Accountability
  - Transparency
  - Personal Data



Source: Seda Gürses (2010)

The Montreux Declaration ("International" Privacy Principles)

- 27th International Conference of Data Protection and Privacy Commissioners
- 2005-09-14/16 in Montreux, Switzerland
- "The protection of personal data and privacy in a globalised world: a universal right respecting diversities" [Source: ICDPPC (2005)]
- Agreement on 11 principles by participating data protection and privacy commissioners

11 Principles of the Montreux Declaration

- Lawful and fair data collection and processing,
- Accuracy,
- Purpose-specification and –limitation,
- Proportionality,
- Transparency,
- Individual participation and in particular the guarantee of the right of access of the person concerned,
- Non-discrimination,
- Data security,
- Responsibility,
- Independent supervision and legal sanction,
- Adequate level of protection in case of transborder flows of personal data.

EU Privacy Directives related to Personal Data and Privacy

- Data Protection Directive (Directive 95/46/EC)
  - Directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data

- Directive on Privacy and Electronic Communications (Directive 2002/58)
  - Directive on Privacy and Electronic Communications with regard to data retention, spam and cookies

9 Principles of EU Privacy Law I

1. **Intention and notification:** The processing of personal data must be reported in advance to a Data Protection Authority.
2. **Transparency:** The person involved must be able to see who is processing her data for what purpose.
3. **Finality principle:** Personal data may only be collected and processed for specific, explicit and legitimate purposes.
4. **Legitimate grounds of processing:** The processing of personal data must be based on a foundation referred to in legislation, such as permission, agreement, and such.
5. **Quality:** Personal data must be as correct and as accurate as possible

Source: Blarkom and Borking (2003)

9 Principles of EU Privacy Law II

6. **Data subject's rights:** The parties involved have the right to take cognisance of and to update their data as well as the right to raise objections.
7. **Processing by a processor:** This rule states that, with the transfer of personal data to a processor, the rights of the data subject remain unaffected and that all restrictions equally apply to the processor.
8. **Security:** A controller must take all meaningful and possible measures for guarding the personal data.
9. **Transfer of personal data outside the EU:** The traffic of personal data is permitted only if that country offers adequate protection.

Source: Blarkom and Borking (2003)

The Implementation of Data Protection

Germany: Federally organised data protection

- Responsibility in Germany:
  Federal Commissioner for Data Protection and Freedom of Information (BfDI)
- Each state in Germany has its “Länder” Data Protection Commissioner.
  - Specialisation on certain fields, e.g. in Schleswig-Holstein (ICPP) on Privacy in the Internet
- **Additionally:**
  Data protection commissioners within governmental administration and within companies

The origin of data protection in Germany?

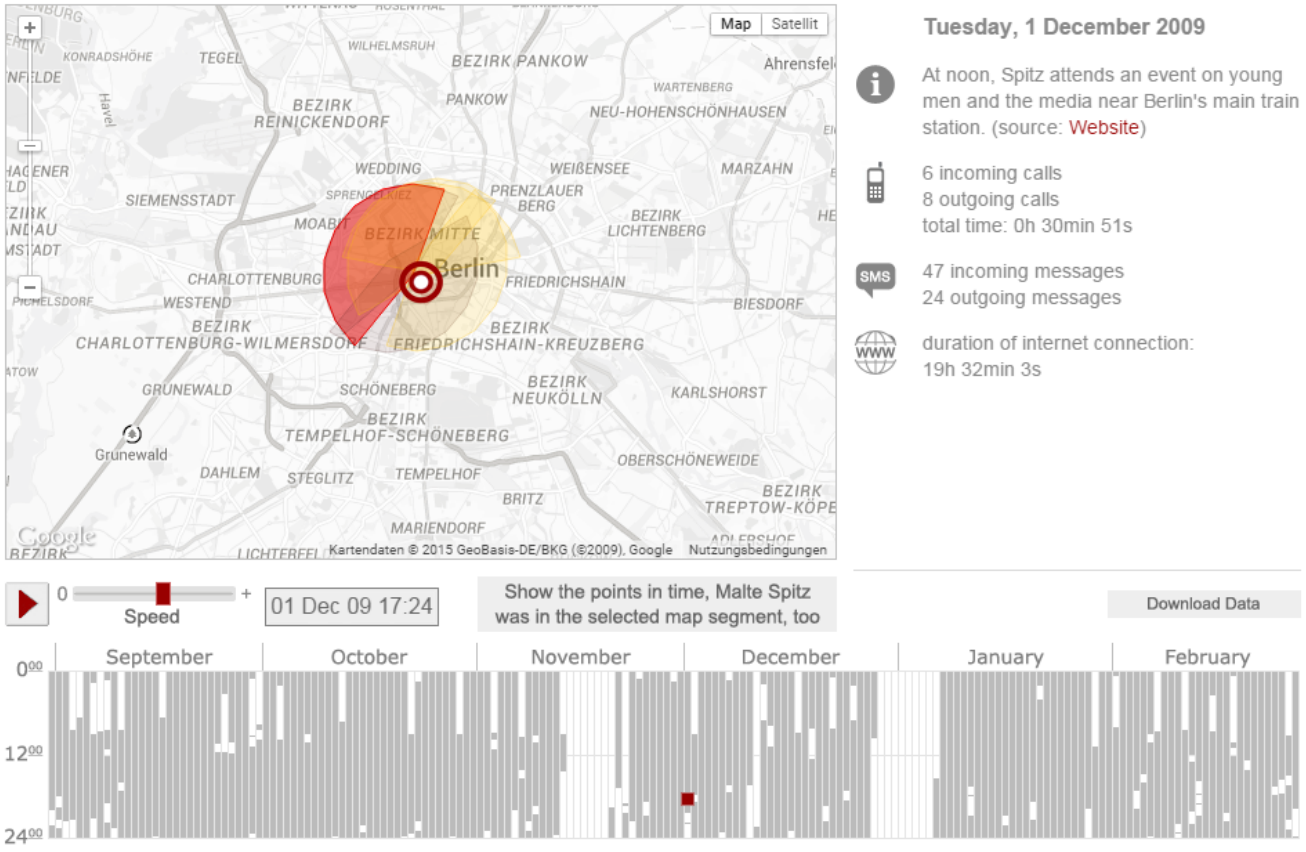
- The term “Privacy” (‘the right to be left alone’) originates from [Warren and Brandeis, 1890].
- Data protection in Germany (“Datenschutz”) originates from concerns over too much information und power in the hands of large (governmental” institutions (“Big Brother”).
- Nowadays Data protection and Privacy in Germany are based on the right of informational self determination derived from the constitution in the “Volkszählungsurteil“ [BVG 1983]).
- Germany has one of the most advanced infrastructures for Privacy but still no established German language term for Privacy beyond the (misleading “Datenschutz”).
- Some (more or less established) related terms are:
  - Privatheit
  - Privatsphäre
  - Schutz der Privatsphäre

Data Protection Interests and Conflicts

- Users want to keep their personal data under their control.
- Service providers want to use the customers' data mainly for commercial purposes (e.g. customer profiles / targeting).
- The legislator demands:
  - Data protection on the one hand
  - Surveillance and retention of data on the other hand.
  - Conflicts between expectations and regulations often arise.

Illustration of possible consequences of „Vorratsdatenspeicherung“

Green party politician Malte Spitz sued to have German telecoms giant Deutsche Telekom hand over six months of his phone data that he then made available to ZEIT ONLINE. They combined this geolocation data with information relating to his life as a politician, such as Twitter feeds, blog entries and websites, all of which is all freely available on the internet.



Source: www.zeit.de/datenschutz/malte-spitz-vorratsdaten

Law alone is not sufficient

- The increased usage of IT systems and networks leads to
  - huge amounts of data
  - easily searchable data
  - automatic analysis
  - and knowledge extraction

- Data protection / Privacy law alone not sufficient
  - Not all processing can be controlled (e.g. every network node).
  - Deliberate breaking and bending of law (different legislations on the internet)
  - Economic pressure can force customers to give consent to almost any kind of ‘privacy’ policy (e.g. selling privacy for “peanuts”).

- Slow pace of privacy self-regulation in the US, Focus on self-help
  - Self regulation by sustaining user ignorance
  - Enforcing norms may violate anti-trust.
  - Being a good actor (e.g. by exposing privacy practices) increases liability.
  - Legal compliance and related business processes (deemed) expensive

Source: Reagle (1998); SelfReg (1999); Bell (2001); Hoofnagle (2005)

⇨ Technical Privacy Protection
⇨ Standardisation

Agenda

- What is Privacy?
- Data Protection Directives and Laws
- Technical Data Protection
- General Requirements for Privacy
- Selected Privacy Concepts
- One more thing about Privacy

Technical Data Protection

- Individuals
  - want to control the amount of identity information visible from the outside.
  - consider what personal information they reveal to whom.

- Typical protection techniques are:
  - Anonymization and identity management tools
  - Data encryption
  - Spontaneous switching between different levels of anonymity and pseudonymity depending on the context

Examples of Privacy Enhancing Technologies (1)

- The Anonymizer
  www.anonymizer.com

- Mixmaster – Anonymous Remailer
  http://mixmaster.sourceforge.net

- Java Anonymous Proxy (JAP)
  http://anon.inf.tu-dresden.de

- Tor Network
  http://tor.eff.org/

Examples of Privacy Enhancing Technologies (2)

- Cookie Cooker

  www.cookiecooker.de

- P3P – Platform for Privacy Preferences

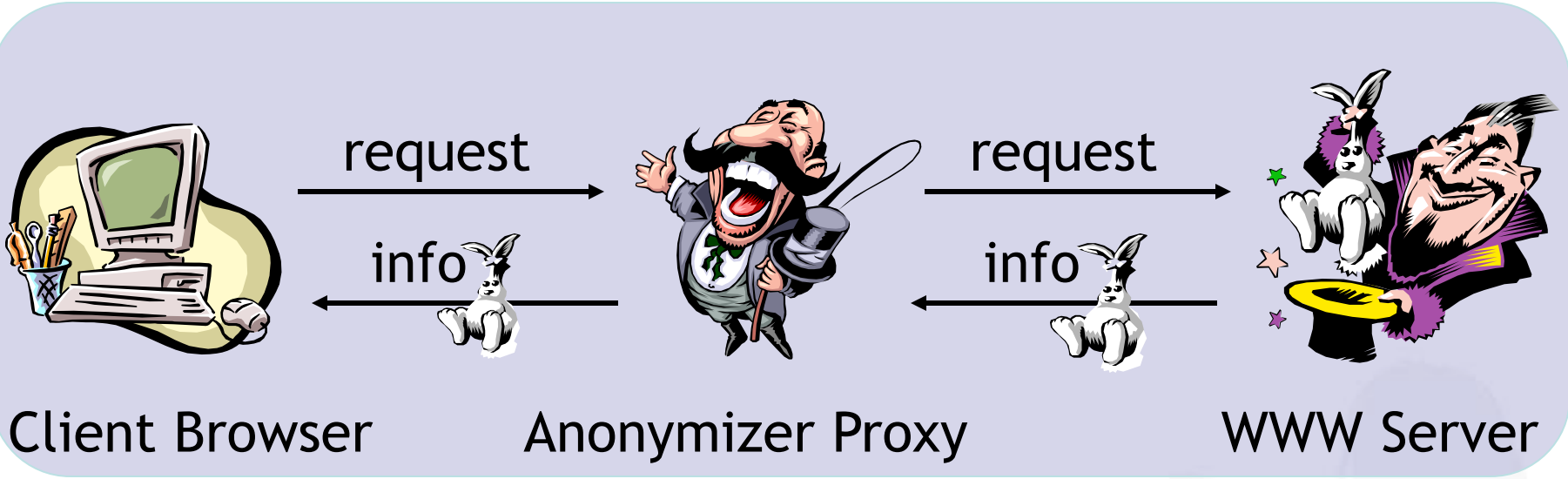
  www.w3.org/P3P

- Idemix

  www.zurich.ibm.com/security/idemix

- Online Tracking Protection
  - DoNotTrack
  - IE Tracking Protection

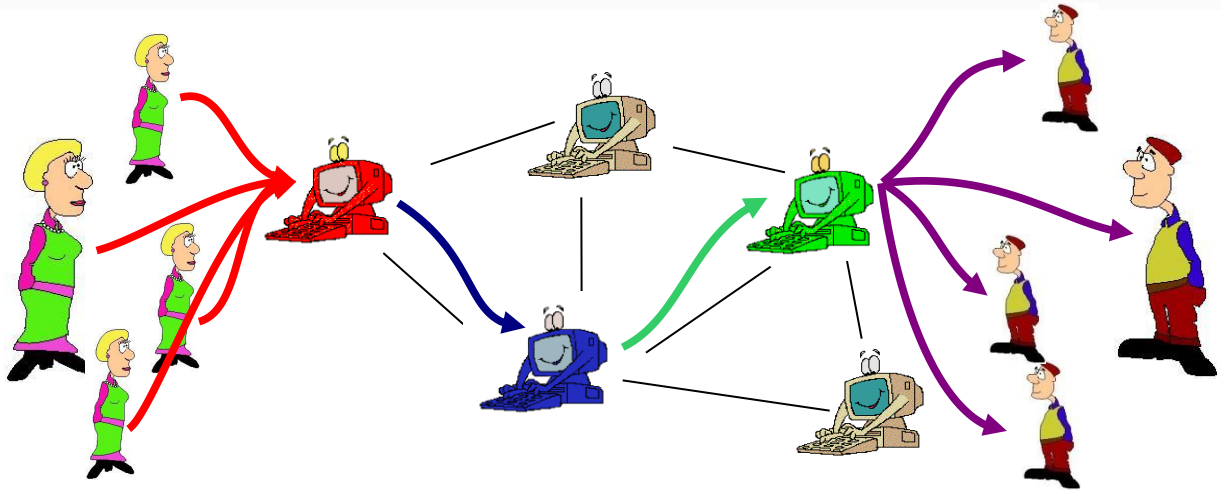
www.anonymizer.com



www.anonymizer.com

- ↑ Client (anonymity) is protected in an "**anonymity set**" of all possible proxy clients.
- ↓ Anonymizer learns about client's activities / interests.
- ↓ No protection against attackers with global view.

Mixes and Onion Routing



- Communication is anonymised by multiple mix servers, also called onion routers.
  - Both onion routing and JAP are based on the same Mix concept.

Java Anonymity Proxy (JAP)

- Users can choose between multiple mix-cascades.
- Number of active users is a heuristic for level of anonymity achieved.
- Current version does not achieve security against a global attacker but can protect against local attackers
  - your boss
  - your provider
  - operator of a mix

http://anon.inf.tu-dresden.de

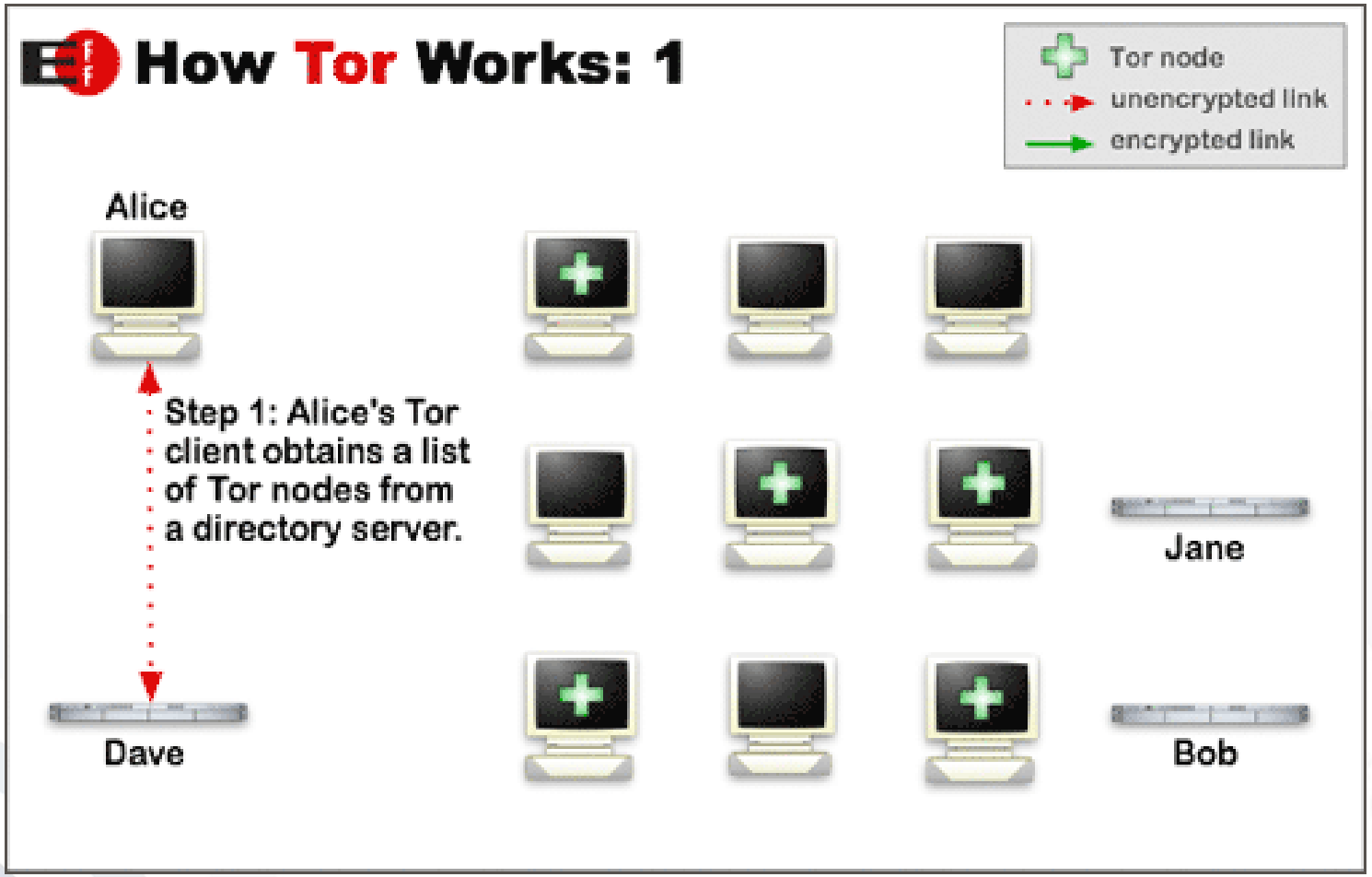
Tor Network

- Tor is a network of virtual tunnels that allows people and groups to improve their privacy and security on the Internet.

- Distributed anonymous network

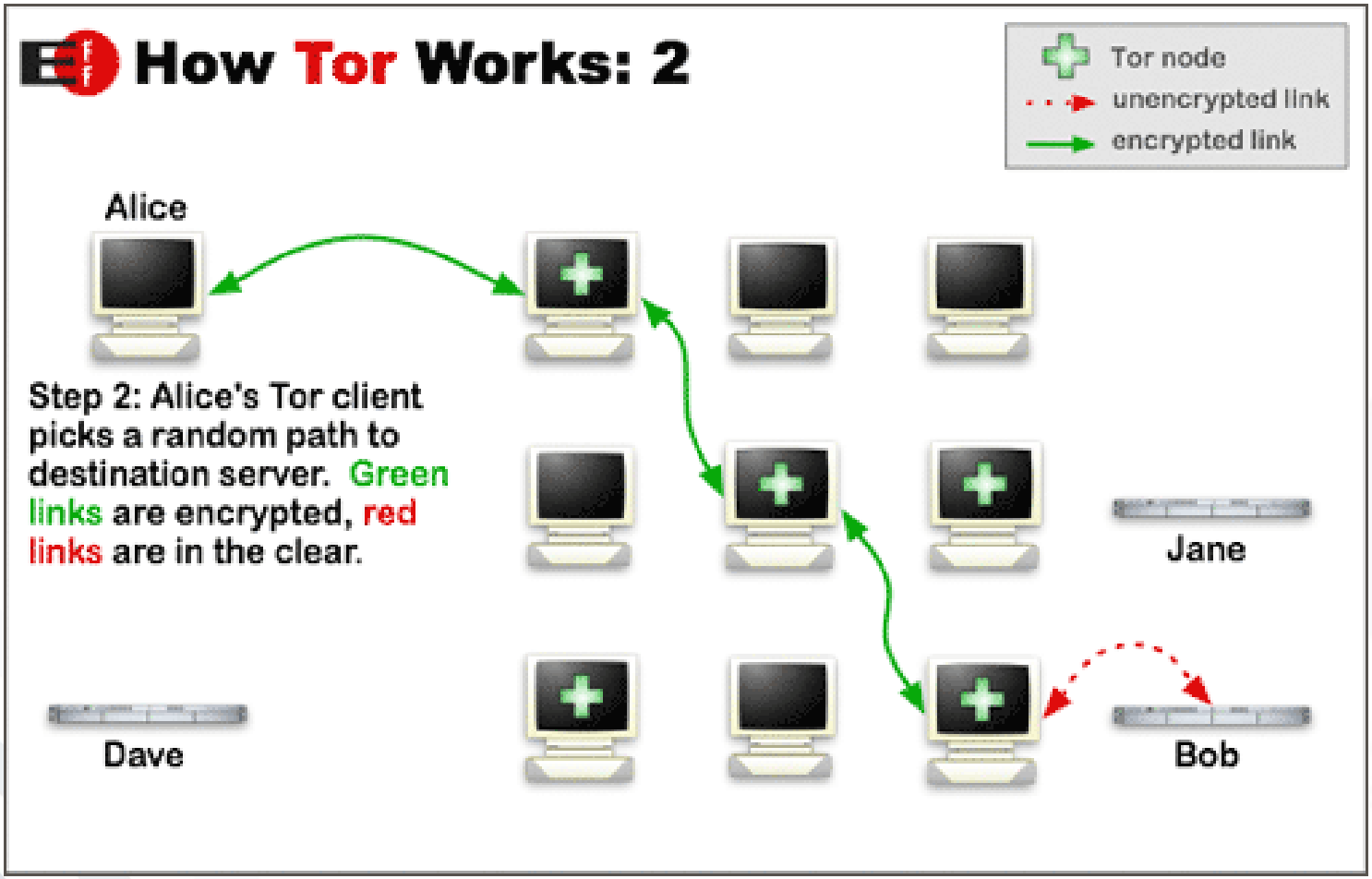
- Tor allows users to change circuits during sessions
  →Aims to minimize linkability of actions

How Tor works (1)



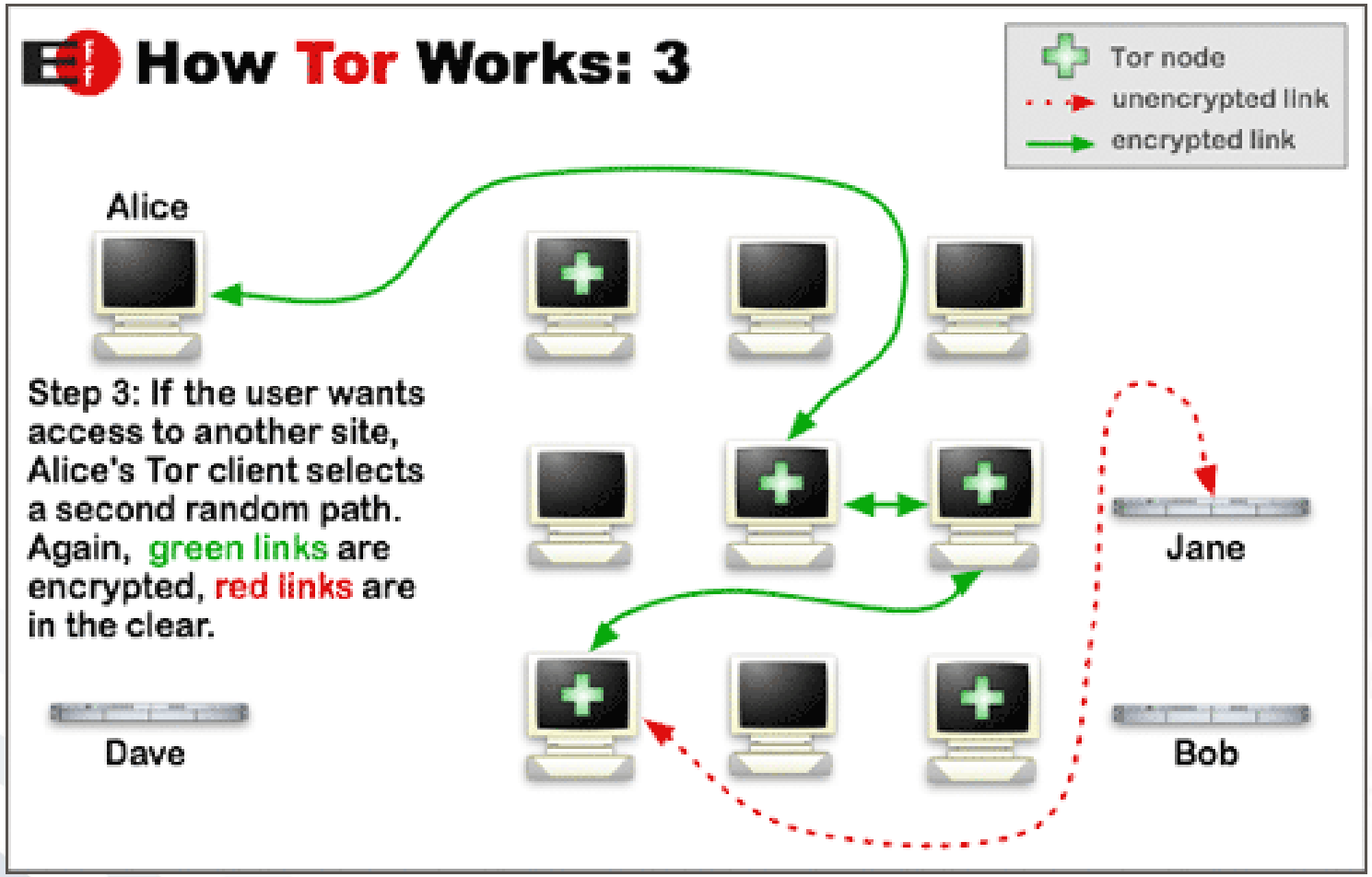
http://tor.eff.org

How Tor works (2)



http://tor.eff.org

How Tor works (3)



http://tor.eff.org

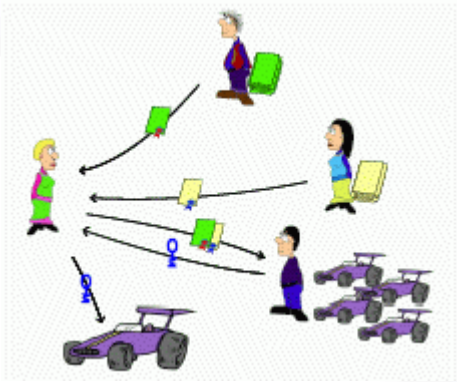
CookieCooker

- Confuses data collectors
  - Exchange of cookies between users
  - Exchange of identities
  - Use of „faked“ data
- User-defined identity management
  - Assistance for the registration
  - Application of „real“ and „faked“ data
- Spam protection through disposable email addresses
- Ad blocking
- Integrated with JAP Anonymizer

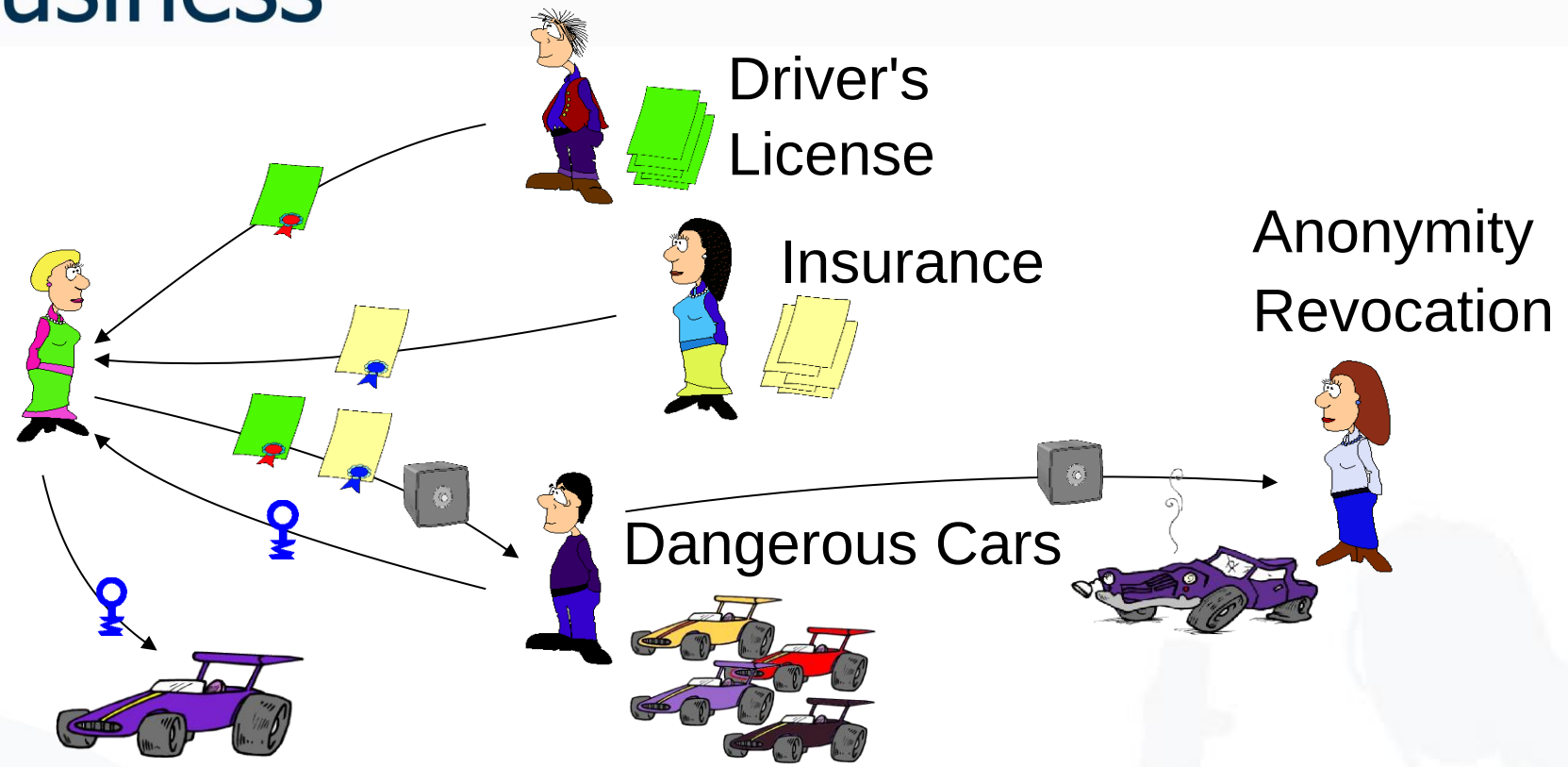
Idemix: Anonymous Credentials

- Anonymous Credentials are used to prove privileges or attributes of their owner without revealing its identity, e.g. to prove, that
  - a device contains an unrevoked Trusted Platform Module (TPM); this is also called Direct Anonymous Attestation
  - the owner possesses a subscription and is of the required age, e.g. for an identity management system supporting anonymous video download

- Such a system needs to have the following properties:
  - Unforgeability of credentials
  - Unlinkability of credentials
  - No credential sharing
  - Consistency of credentials



Idemix: Car Example



- What happens if a user exchanges information with multiple parties?
- The user has different pseudonyms with different parties.
- The user uses credentials to prove that he has a driver's license and an insurance.

Platform for Privacy Preferences (P3P)

- Standard of declaring privacy preferences in a standardized way
  - Snapshot of how a web site handles personal information about its users
  - P3P enabled browsers can "read" this snapshot and compare it to the consumer's set of privacy preferences.

- P3P enhances user control by
  - putting privacy policies where users can find them,
  - in a form users can understand, and
  - enables users to act on what they see.

Source: W3C P3P

- Unfortunately, this promise has not been fulfilled, yet.

Online Tracking Protection

DoNotTrack Flag

- Browsers are signaling advertising networks via the DoNotTrack Flag not to track the online behaviour of their users
- Problem: Advertising networks and other data collectors can either respect or ignore the DoNotTrack Flag
- Solution Approach: Privacy Badger – Enforcing respect for the DoNotTrack Flag (www.eff.org/privacybadger)

Online Tracking Protection

Tracking Protection Lists

- Tracking Protection is build into a web browser (e.g. Internet Explorer 9 or later).
- Based on black lists, browsers prevent tracking data being transferred to advertising networks or other data collectors while Online advertisements are still being displayed.

- Problems
  - Who maintains and updates the lists?
  - Do users understand the black list concept?
  - What if tracking protection is turned on by default?

Privacy-enhancing Technologies (PET) alone are not sufficient

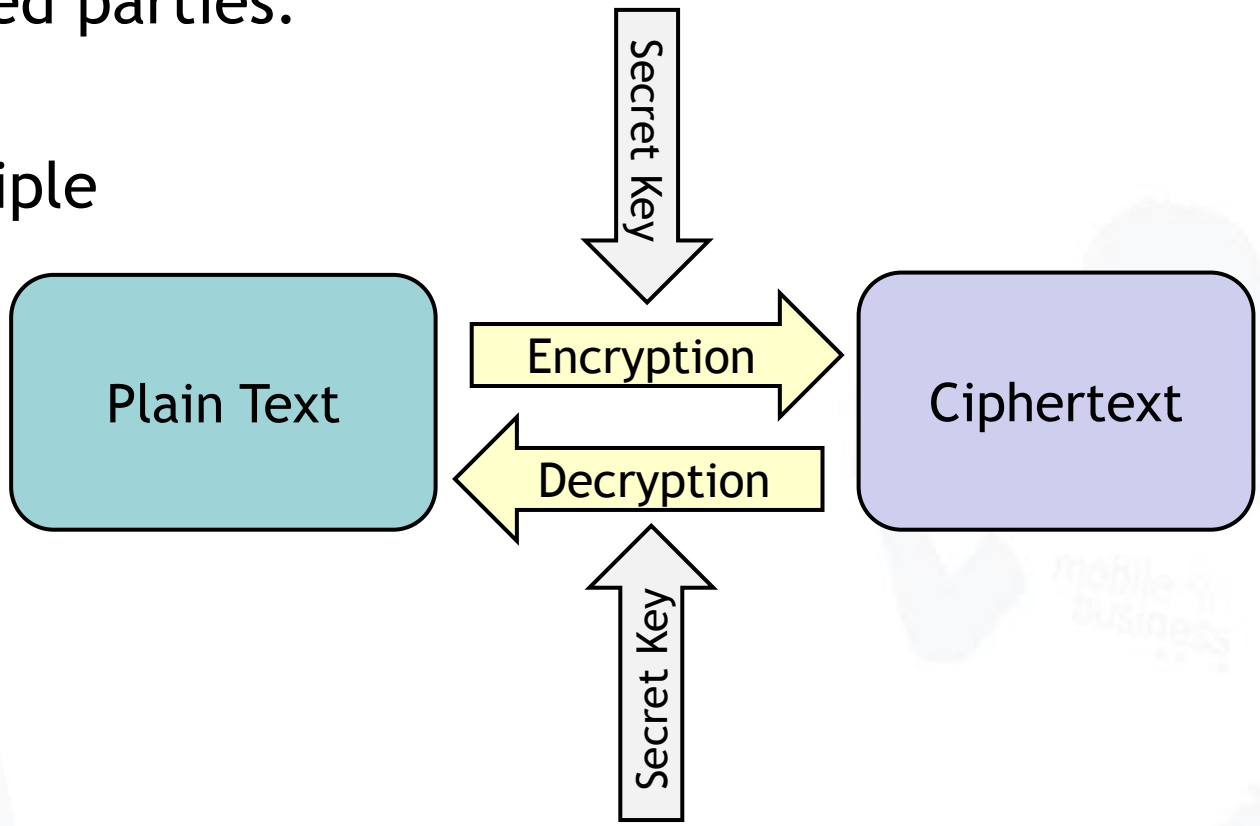
- Anonymisation and Pseudonymisation
  - Mix-Master, Onion Routing, Anonymous Payment, Anonymous Credentials
  - A myriad of techniques and algorithms

- Issues of PETs
  - Lack of integrated privacy protection for
    - business processes
    - for user interfaces
    - for applications in general
  - PETs affect service quality and performance
  - Playing Cat and Mouse with “Big Brother”
  - Privacy and service quality are always a trade-off. Which personal data to disclose and which to hide?

Data Security / Encryption

- Encryption is the transformation of data into ciphertext, which typically cannot be understood by unauthorized parties.

- Basic Principle

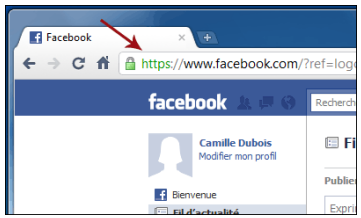


Data Security / Encryption (2)

- Symmetric Encryption
  - Sender and receiver share the same secret key used to encrypt/decrypt a message
  - Challenge: How to securely share the secret between sender and receiver?
- Asymmetric Encryption
  - Receiver has a private and public key. Senders encrypts message with public key while receiver of message decrypts message
  - Challenge: How to deal with the high computing/processing requirements?
- Hybrid Encryption
  - Mix of symmetric and asymmetric encryption, which combines the best of both worlds
  - Asymmetric encrypted message used to share the secret symmetric key between sender and receiver. Then switch to symmetric encryption for the actual messages

Data Security / Encryption (2)

- Encryption of data in different states
  - Encryption of stored data
    - BoxCryptor, Spideroak
    - Zero Knowledge Encryption
  - Encryption of data in transit
    - Transport Layer Security
  - Encryption of data being processed
    - Hardly possible, mainly still research

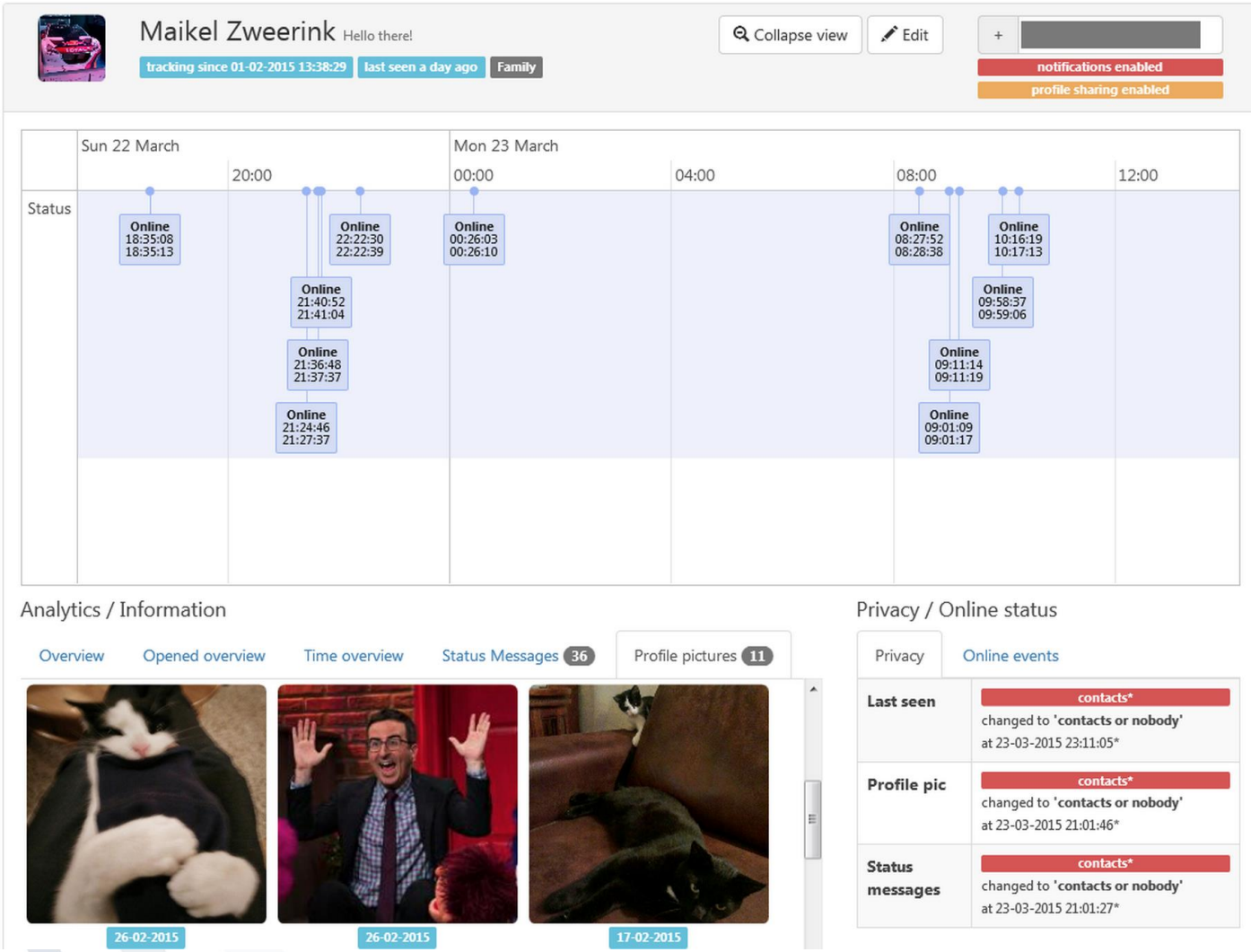


- Threads to encrypted data
  - Weak passwords
  - Carelessly revealed keys
  - Man-in-the-Middle Attacks (symmetric encryption)
  - Weak implementation of encryption algorithms
  - Weakly designed encryption algorithms
  - Implementations contain backdoors

Public WhatsApp Spy Example for Security Issues

- Software that allows to reveal for ANY WhatApp user (if mobile phone number is known) the following information:
  - Online/Offline status (even with privacy options set to „nobody“)
  - Profile pictures (if set to everyone; set by default)
  - Status messages (if set to everyone; set by default)
  - Privacy settings
- Collected meta data from users could be cross-referenced with other information about the user
- **More information (last access:16.4.15):** https://gitlab.maikel.pro/maikeldus/WhatsSpy-Public/wikis/home

Public WhatsApp Spy (2) Example for Security Issues



Agenda

- What is Privacy?
- Data Protection Directives and Laws
- Technical Data Protection
- General Requirements for Privacy
- Selected Privacy Concepts
- One more thing about Privacy

General Privacy Requirements

- **Anonymity**
  is the condition of not being identifiable within a set of subjects. The anonymity set for a given action is the set of all subjects who might have triggered the action.

- **Pseudonymity**
  is an identifier used in place of the “real” identities, e.g., name, unique id number, of a given user. Pseudonymous identifiers can be made conditional and accountable using cryptographic building blocks.

- **Unlinkability**
  is the condition in which a third party cannot determine whether two actions or two data items belong to a single user. Unlinkability is central to another privacy related concept called the separation of identities.

Source: GINI (2011)

General Privacy Requirements (2)

- **Separation of Identities**
  the condition of guaranteeing that separate partial identities of a given user are unlinkable.

- **Separation of Audiences**
  the condition in which a user can control the audience of the information s/he reveals. The flexibility of the access control models determine the type of separation of audiences that can be practiced by the user.

Source: GINI (2011)

Agenda

- What is Privacy?
- Data Protection Directives and Laws
- Technical Data Protection
- General Requirements for Privacy
- Selected Privacy Concepts
- One more thing about Privacy

Privacy by (Re)Design

- Privacy by Design advances the view that the future of privacy cannot be assured solely by compliance with regulatory frameworks; rather, privacy assurance must ideally become an organization's default mode of operation.

- Privacy by Design states that privacy and data protection are embedded throughout the entire life cycle of technologies, from the early design stage to their deployment, use and ultimate disposal.

- The objectives of Privacy by Design — ensuring privacy and gaining personal control over one's information and, for organizations, gaining a sustainable competitive advantage — may be accomplished by practicing the 7 Foundational Principles

Source: Ann Cavoukian, Privacy Data Commissioner of Canada

Privacy by Design Principles

1. Proactive not Reactive; Preventative not Remedial
2. Privacy as the Default Setting
3. Privacy Embedded into Design
4. Full Functionality — Positive-Sum, not Zero-Sum
5. End-to-End Security — Full Lifecycle Protection
6. Visibility and Transparency — Keep it Open
7. Respect for User Privacy — Keep it User-Centric

More Privacy Concepts

- Privacy by Default
  - The collection, display and sharing of personal data is not allowed without explicit consent of a user
- Transparency
  - Organisations need to make transparent to their users what personal data they collect and how they process and use it
- Right to be forgotten
  - Users have the right to have their personal data (stored by organisations) deleted at their command

User Awareness as Prerequisite for Privacy Protection in General

- Take This Lollipop, www.takethislollipop.com
  (Scary visualisation of a user's Facebook profile)
- Facebook Privacy App, http://app.europe-v-facebook.org
  (what Facebook shares about you)
- Please Rob Me, http://pleaserobme.com/
  (Determining if someone is not home, based on his check-ins)
- We know what you are doing, www.weknowwhatyouredoing.com
  (Analysis of public Facebook status updates and their meanings)
- Need a debit card, www.twitter.com/NeedADebitCard
  (Tweets about pictures of debit cards taken by their owners)
- Lamebook, www.lamebook.com
  (Funny & embarrassing (un)-intented public Facebook posts
- Ghostery, www.ghostery.com
  (Alerts about online trackers on Websites, Browser Plug-In)
- Panoticlick, panopticlick.eff.org/
  (Tests the uniqueness of your browser fingerprint)

Agenda

- What is Privacy?
- Data Protection Directives and Laws
- Technical Data Protection
- General Requirements for Privacy
- Selected Privacy Concepts
- One more thing about Privacy

NSA & Co.
Some press statements

Only technology is the limit …

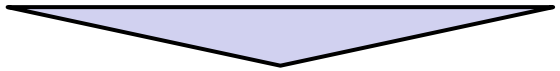
- NSA: Large scale attack on encryption technology
- NSA: Surveillance of German chancellor's mobile phone
- NSA: Access to all major smartphones (operating systems)
- NSA: XKeyScore as real-time analysis tool to handle massive amount of collected data
- NSA: Access to Telecommunications providers infrastructures
- NSA: PRISM program: Direct access to user data of Microsoft, Google, Yahoo, Facebook, Paltalk, Youtube, Skype, AOL und Apple
- GCHQ: Storage of all data transmitted through certain Internet backbone (sea) cables
- Direction Générale de la Sécurité Extérieure (DGSE): Storage of communication meta data and user activities on Google, Apple, etc.
- NSA: Surveillance of meta data for phone calls & scanning of letter post
- Latest on NSA: Possible access to some SIM-Card crypto keys and possible ability to compromise hard drive BOOT-ROMs with malware

Further information:

- http://www.heise.de/newsticker/meldung/Was-bisher-geschah-Der-NSA-Skandal-im-Jahr-1-nach-Snowden-2214943.html (1 years summary of NSA scandal)
- http://www.heise.de/extras/timeline (Interactive timeline of NSA events)

NSA & Co. and Privacy
Not all is lost …

- As it appears, the NSA is able to virtually access and also manipulate all communications running across the Internet or Telecommunication infrastructures
- Why should we then start/continue to care about our privacy?



- Privacy is always context dependent
  - You still want to protect your personal data against online enterprises, other individuals or other entities
- Aiming at End-2-End encryption for all sensitive (Internet) communications can become the first level of defence or at least a start
  - Reliable encryption mechanisms are hard to break
  - Exception: Via known exploits or backdoors
  - So how can we know?

NSA & Co. and Privacy
Not all is lost … (cont'd)

- Trusted, reliable communication infrastructures need to be established
  - How does one know that an infrastructure can be trusted?
    - Infrastructures could be audited by independent Third Parties, but why should the latter be trusted then?
  - Open Source software, which could be inspected by everyone (theoretically), should be fostered
    - Could such inspections work for enterprises/NGOs: Yes/Maybe. For individuals: Most likely no. So it's trust again, right?
  - Switching to more secure software alternatives
    - E.g. switching from WhatsApp to Threema
    - No 100% guarantee of privacy protection, but several indications for a more secure approach
    - Does not protect against backdoors on a phone
- As always there is no 100% security for protecting privacy, but with some effort (by individuals, industry and (EU-)regulators) it can be significantly increased.

Literature

- Bell, Tom W.: Internet Privacy and Self-Regulation: Lessons from the Porn Wars, Cato Institute Briefing Papers, No 65., 2001, www.cato.org/pubs/briefs/bp65.pdf
- Blarkom, G. W., Borking, John J., and Olk., J.G.: Handbook of Privacy and Privacy-Enhancing Technologies - PISA Privacy Incorporating Software Agent. The Hague, 2003.
- David Chaum: *Untraceable Electronic Mail, Return addresses, and Digital Pseudonyms*; Communications of the ACM February 1981 Volume 24 Number 2
- Hoofnagle, Chris Jay: Privacy Self Regulation: A Decade of Disappointment, 2005, www.epic.org/reports/decadedisappoint.html
- Reagle Jr, Joseph M., Boxed In: Why US Privacy Self Regulation Has Not Worked, Berkman Center for Internet & Society, Harvard Law School, 1998, http://cyber.law.harvard.edu/people/reagle/privacy-selfreg.html
- Self-Regulation: Regulatory Fad or Market Forces? Paper prepared for Cato Roundtable „Privacy vs, Innovation“ by Solveig Singleton, May 7, 1999, www.cato.org/pubs/wtpapers/990507report.html
- Varian Hal, Economic Aspects of Personal Privacy, Berkley University, 1996.
- Warren and Brandeis, The Right to Privacy”, Harvard Law Review., Vol. IV, December 15, 1890, No. 5