Data Security

Basic Overview and Exercises
Security Techniques

Department of Media Studies
University of Amsterdam

Digital Practices Worksheet

Version: November 2021

Department of Media Studies
University of Amsterdam
Turfdraagsterpad 9, 1012 XT Amsterdam, the Netherlands
https://www.uva.nl/en/disciplines/media-studies
http://www.mediastudies.nl/

Created by Lonneke van der Velden, building on basicprivacysecurity.org and instructions written by Emile den Tex (2018).

Contents

Table of Contents

Introduction

In this worksheet you learn:

- What data security is and why you may need it;
- To understand how security practices contribute to ethical research practices;
- Different types of security tools and how to approach choosing one that best fits your needs;
- How to use security tools to safely store and communicate data.

This worksheet introduces you to some basic principles of internet communication and presents a selection of security practices. You will learn how to translate that knowledge into academic practice, which increasingly takes place in online settings. The first unit of Digital Practices focused on ‘working with data’. In unit 2-4 you learned various data skills. This fifth unit is about ‘working with data *securely*’, which is important in the context of research ethics.

The material aspects of the internet and internet protocols are usually somewhat hidden when using the internet on an everyday basis. That means that potential security problems may remain invisible too.

This worksheet aims to familiarize yourself with these invisible infrastructures, it explains why security practices are important when doing research (and when ‘working with data’), and it provides you with practical exercises that teach you secure data practices.

Before presenting the security tools, this worksheet runs you through basic information about internet communication (section 1) in order to understand why there can exist security problems from a technological standpoint.

Next it walks you through different pieces of software (section 2 and 3). There are many different security techniques which depend on one’s context and communicative demands. This worksheet presents only a small selection of tools that are used to store and communicate data. For us, this will mostly concern research data. Security tools are continuously updated. After the ending of this unit, students are advised to update the software and keep themselves informed about the state of these technologies in order to have them work as desired.

After the tools, you are introduced to issues of data ethics and data ethics requirements at the UvA (section 4), which relate closely to how we collect, process and store research data.

Section 5 presents the assignment for unit 5, which consists of a component on data ethics and two practical components.

This worksheet is based on previous worksheets in the course of Digital Objects and Research Practices and a guide to online privacy (basicprivacysecurity.org).

1. Internet Security

1.1 Trusting the Internet

When verbally passing a message you usually need to assess your contact persons to know if your message arrives in the way it was intended. Similarly, you have to know your technology a little to know if you can trust it. Technologies can leak or distort your message just as humans can. Technologies are invested in types of trust relations: some devices are safer than others, some can be modified, and others are better avoided. This worksheet tries to address these different layers by giving hands-on explanations on how to make your digital communication and data more secure and by providing you with a basic understanding of the concepts of digital communication and data security.

In the eighties when the Internet was in its infancy, people did not foresee the internet as a way to organise financing, health, commerce, and private communication. Despite the development of the internet being funded through the American ministry of Defense, its main usage came from university students and professors in an atmosphere of implicit trust. This means that security was not the first thing in mind when the basic uses and functions of the Internet were first developed. Hence, no security was built in.

Nowadays the Internet is everywhere both in public and in private life. It has become a vital means for professional and personal - often confidential - communication. This has required security enhancements to be added to the various communication methods used on the internet after it became widely used. A lot of these enhancements are not implemented by default or require additional configuration.

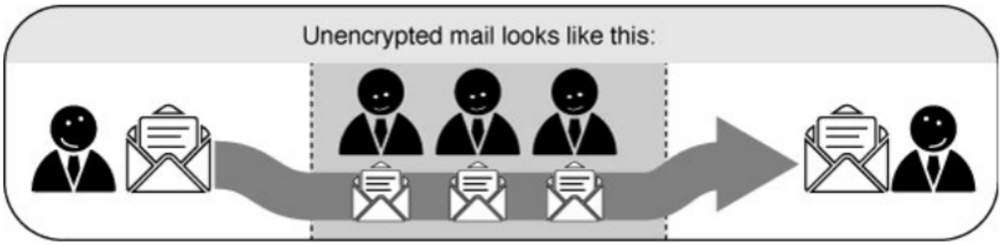
1.2 What is security?

Absolute security does not exist, and security is always context dependent: who or what do you need to secure your data for? People may have different reasons for taking security measures. People might prefer to secure their data out of privacy reasons (not wanting to share personal information to third parties), platform criticisms (not wanting to feed big data platforms with information), professional reasons (for instance, a journalist wanting to protect his or her source), fear (having to operate under surveillance risks or censorship), or out of obligation (having to meet privacy rules, for instance The General Data Protection Regulation) when applying for research funding.

Security tools secure only certain aspects of your communication. Therefore using one tool is not a guarantee to full security (or anonymity). The success of the tool also depends on whether it is used properly. Security is therefore about informing yourself and assessing the possible risks you, and others you communicate with, are facing. Make sure you reserve some time to choose the right tools, install everything properly, and test if it works. Compare it with driving a car: it takes a little bit of practice, and some judgement on others' behaviour, but as soon you are in control it can safely get you where you want.

1.3 Understanding basic Internet security

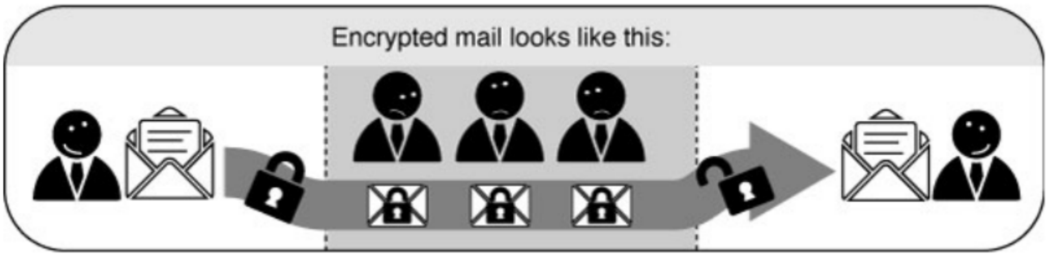
To understand basic internet security we should have a basic understanding of how the Internet is organised and which path our information travels. To have a notion of how the Internet works you can compare it with the normal world wide mail network. If you want to communicate with a friend you can send her a letter and post it to the nearest mailbox; it then travels through an extensive network to (hopefully) reach the person the information is intended for. Internet is just like that, *however, the message is sent in an open envelope and every postman on the way can read the message, alter its content and/or the destination without you knowing.*



To counter this, people have long used secret languages to communicate safely. In this worksheet we will explain two methods of encryption. The first method explains an end-to-end encryption: encrypting the whole way from sender to receiver. The second method partly encrypts the route.

1.4 End-to-end encryption

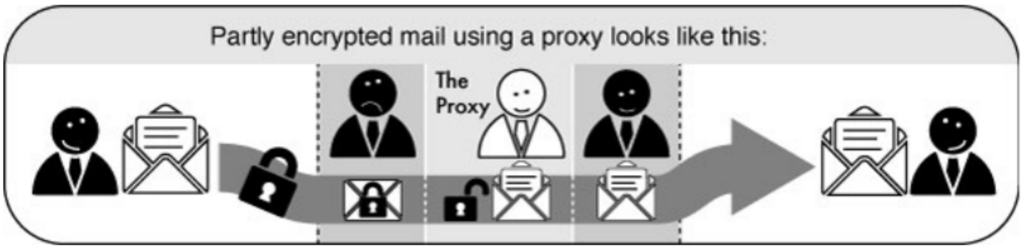
If you encrypt your message and only the recipient can read it, it will be meaningless to all the postmen in between, and if they alter it you will notice it directly. In order to make such an encryption work, you still have to be sure to trust the recipient and be sure that you are really exchanging information with her and not with someone pretending to be her. This method is called end-to-end encryption and is the safest way of communication. You also have to be sure that no one is watching over your shoulder while you write your message. Some of the end-to-end encryption methods that we cover in this worksheet are HTTPS)(for browsing in our example) and PGP for e-mailing.



Unfortunately for end-to-end encryption to work, both you and your friend (colleague, source, co-worker) need to have the tools to use it and have to agree on the secret language used. On the internet this means the website you are visiting or the people you are e-mailing. This is not always the case, still, we can considerably increase our online safety by encrypting a part of the route.

1.5 Partly encrypted mail through a proxy

To get back to the mail analogy: you might be on a field trip in a repressive country and want to send a message to your friend at home. You don't trust the post offices and the postmen in this country. So before you left, you asked your local post office to act as an intermediary (the proxy) and agreed to use a secret language. Now you can just write a message to your friend in the secret language of your post office. You will send this to your post office and they will take care of the delivery of the message to your friend. In this scenario you have to trust your local post office, all the postmen after that, and of course your friend.



1.6 Visiting websites is communicating

Because in this example an analogy was drawn with mail messages, you probably thought of ‘e-mails’ when reading this. While this is true, the example also counts for other internet communications. Until recently, visiting a website or using an app was just like sending the message to your friend "please mail me your copy of the book 1984", after which she sent it to you. Let's follow the example of visiting a website from your home computer:

1. You type in http://mediastudies.nl/.

2. The request goes through a series of routers, each one forwarding a copy of the request to a router closer to the destination, until it reaches a router that finds the specific computer needed.

3. This computer sends information back to you, allowing your browser to display the page. The message that is transmitted from the website to you travels through other devices (computers or routers). The amount of devices your message comes in contact with along its way is often between 5 and 30.

By default, information travels on the internet in an insecure way. This means that your message can be eavesdropped or tampered with on every device. If you are connecting wirelessly, people can also just "tune in" to the information that is being sent through the air. Have a look at what these artists managed to do in an internet cafe a couple of years ago:

https://criticalengineering.org/projects/men-in-grey/videos/men_in_grey-480p.mp4

Luckily, a large part of the internet traffic is currently encrypted via SSL/TLS. (Represented with the little lock in the URL bar of your browser.) However, there are still some websites (and also web shops!) that have not encrypted their website.

In sum, to keep information from being compromised, you have to be careful to make sure of the following: Can you trust the entry point (your internet connection) to the internet? If this is an insecure wireless connection anyone can eavesdrop on it, if it is a physical (cable connection) it can be eavesdropped by the operator. Can you trust the exit point (the site you will be visiting) of your information? Are you really communicating to the right destination? Or did your request end up on a server trying to appear like the server you were looking for, but really isn't.

If you want to read more about the internet, and the complex system of protocols that make the internet behave as it does, you can find more information under the suggested readings.

Suggested readings

- Burrington, Ingrid. “Networks of New York”. http://lifewinning.com/projects/networks-of-new-york/
- Global Commission on Internet Governance. 2016. *One Internet*. Centre for International Governance Innovation and Chatham House. Chapters: "What Do We Mean When We Say "The Internet”, and "A Fine Balance: Promoting a Safe, Open and Secure Internet (including 'Internet Governance: A Complex and Distributed Landscape')
- D.R.E.A.M. Peering into Internet Infrastructure with Ingrid Burrington. https://www.youtube.com/watch?v=E5f7Jikg7ZU
- Galloway, Alexander R. 2004. *Protocol: How Control Exists after Decentralization*. MIT Press. Chapter 1: Physical Media
- Parks, Lisa.2009. Around the Antenna Tree: The Politics of Infrastructural Visibility. *Flow Journal.*

http://www.flowjournal.org/2009/03/around-the-antenna-tree-the-politics-of-infrastructural-visibilitylisa-parks-uc-santa-barbara/
- Mattern, Shannon. 2013. “Infrastructural Tourism.” *Places Journal*, July. https://doi.org/10.22269/130701.
- Sandvig, Christian. 2013. “The internet as infrastructure.” *The Oxford Handbook of Internet Studies* https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199589074.001.0001/oxfordhb-9780199589074-e-5

2. Secure your computer

There are steps that everyone with a computer should take to keep it secure. This may involve protecting information about your research subjects, your digital methods data collection, your colleagues or your credit card number, but some of the tools you need are the same. Your computer holds valuable information and this needs to be protected. Beware of programs or people that promise perfect security: online safety is a combination of good software and human behavior. Knowing what should be kept offline, who to trust, and other security questions cannot be answered by technology alone. Look for programs that list risks on their Web sites or have been peer reviewed.

2.1 Keep your OS updated

Keep your operating system up-to-date: the developers of operating systems provide updates that you should install from time to time. These may be automatic or you may have to request them by entering a command or adjusting your system settings. Some of these updates make your computer more efficient and easier to use, and others fix security holes. Attackers learn about these security holes rapidly, sometimes even before they're fixed, so fixing them promptly is crucial. Luckily most operating systems do a quite good job in keeping the system updated and safe, if at least you allow them to do so. Installing new updates on a new computer is very important. A new computer you buy in the shop, can be there for some months already. This means the computer is often behind with the security updates. So when buying a new computer, please take some time to update your Operating System.

2.2 Passwords

Every computer needs an account to login. This account is needed to access your data and use the functions of your computer. Please be sure to set up a password for every account. Use good passwords, also for clients and online services.

Many people forget their passwords or use the same password for several online services. This is problematic, because if a platform gets hacked, the hackers could try out this password for all the other platforms for which you used the same username. This has happened in the past for example with LinkedIn. LinkedIn got hacked and many account details ended up online.

There exist databases that document whether accounts have been hacked, for instance https://haveibeenpwned.com/.

One solution for using passwords is to use a dedicated application to manage most of your passwords. A password manager is a program that can generate strong passwords and remembers them. You only need to remember one master password. Operating systems often have in-built password managers. An open alternative is KeePassXC.

Exercise:

Check whether your account details have been leaked in the past via https://haveibeenpwned.com/

Exercise:

Check whether your operating system has an in-built password manager and start using it, or install the alternative KeePassXC.

3. Applications for Internet security

3.1 Browsers

A browser is a piece of software that helps you access the world wide web and which renders it visible in a graphical way. Web browsing is one of the key activities we engage in while using the internet. Our browsing histories, the things we search for, the sites we visit and the things we post might be of interest to others, either for commercial or political reasons.

The first thing to consider is which web browser to use. Windows comes pre-installed with Internet Explorer while Apple computers come shipped with Safari. Freely available and more privacy-friendly browsers are Firefox and Brave. Brave's has developed a - from their perspective - privacy friendly business model, through so called ‘Basic Attention Tokens’ (https://basicattentiontoken.org/)

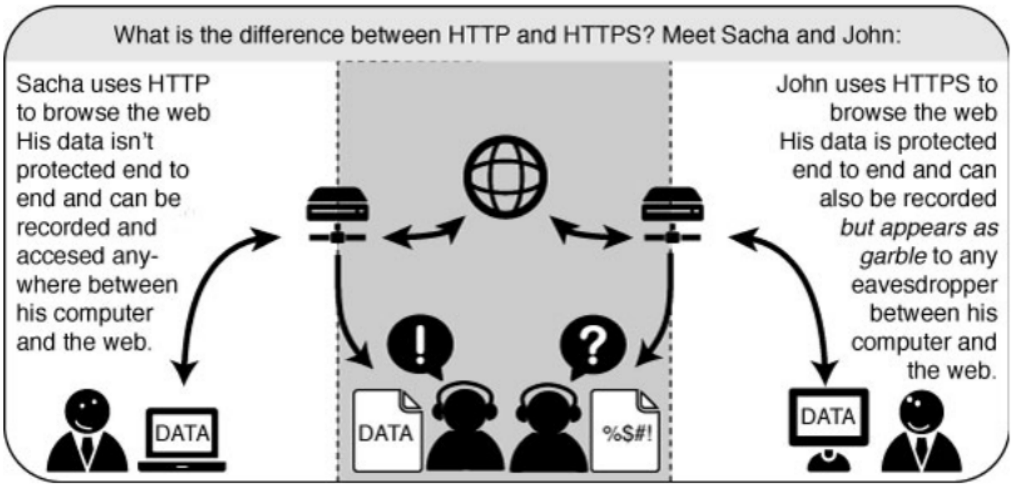
In the worksheet we will also recommend a couple of extensions. Not all extensions work in each browser. If you are used to using only one browser, take the effort to install an alternative one.

Exercise:

Install Firefox and Brave. Compare their performance to the other browser(s) you know.

3.2 HTTP and HTTPS-Everywhere

The Hypertext Transfer Protocol (HTTP) is the networking protocol used by browsers that allows communication between you and a site you are visiting. Because communication is transmitted in plain text it is unsafe, especially when using wireless networks. It is like transmitting a message with personal information on a postcard. Data, such as usernames and passwords, sent to and received by Web sites, are easy to read by third parties. To solve this problem the Hypertext Transfer Protocol Secure (HTTPS) was invented to provide encrypted communication and secure identification of a network web server. Most major Websites, including Google, Wikipedia, and popular social networking platforms such as Facebook and Twitter can also be reached via a secure connection.


What is the difference between HTTP and HTTPS? Meet Sacha and John:
Sacha uses HTTP to browse the web His data isn't protected end to end and can be recorded and accesed any-where between his computer and the web.
John uses HTTPS to browse the web His data is protected end to end and can also be recorded *but appears as garble* to any eavesdropper between his computer and the web.
DATA
DATA
%$#!
DATA

Exercise

Install the plugin “HTTPS-everywhere” in your browser.

Explain what it does. Try to use the tool with the different settings. Visit a range of websites to see if there is any difference. For a (recent) list of unencrypted websites, see: https://whynohttps.com/

If you own your own website, you can check here whether it is conforms current internet standards: https://internet.nl/

3.3 Online tracking and tracker blockers

Nowadays, the majority of web sites are populated with trackers. Trackers are pieces of code that collect information about browsing behaviour. They are installed on a website and allow third party connections. This means that your computer connects not only to the server on which this website is hosted, but also to the servers of those other parties/companies. Those companies usually collect several types of data (such as: search terms, device type, user data) in order to build profiles or sell the data.

There exist many tracker blockers that prevent these connections from taking place. You may have already installed a couple of these extensions or add-ons, such as Lightbeam or Ghostery.

Exercise:

Watch this video: https://youtu.be/2aesvIjQyAo
Start Firefox. Install Lightbeam: https://addons.mozilla.org/nl/firefox/addon/lightbeam-3-0/
Use it over the course of a day. What have you learned about your online profile?

Note:
If you want to do more research into online profiling, know that the Digital Methods Initiative has developed a piece of software called Brightbeam. Brightbeam is a Firefox extension based on Lightbeam. Brightbeam looks even further behind the tracker names and offers you insight into the companies that these trackers belong to. Hence, you can see which companies collect the data that is being produced by your browsing behaviour.

Suggested readings:

- Gerlitz, Carolin, and Anne Helmond. 2013. “The like Economy: Social Buttons and the Data-Intensive Web.” New Media & Society 15 (8): 1348–65. https://doi.org/10.1177/1461444812472322.
- Lupton, Deborah. Feeling your data: Touch and making sense of personal digital data. https://journals.sagepub.com/doi/abs/10.1177/1461444817717515
- Van der Velden, Lonneke. 2014. “The Third Party Diary: Tracking the Trackers on Dutch Governmental Websites,” no. #Traces. https://necsus-ejms.org/third-party-diary-tracking-trackers-dutch-governmental-websites-2/

3.4 VPN

A VPN (Virtual Private Network) encrypts and tunnels all Internet traffic between yourself and another computer (VPN server). This computer might belong to a commercial VPN service, your organization, or a trusted contact.

Because VPN services tunnel all Internet traffic, they can be used for e-mail, instant messaging, Voice over IP (VoIP) and any other Internet service in addition to Web browsing, making everything that travels through the tunnel unreadable to anyone along the way. This makes your connection more secure by default.

If the tunnel starts at your laptop in a country that is known for surveillance and censorship (eg. China) and ends at your VPN-provider in a country that has more internet freedom (eg. Iceland), this can be an effective method of circumvention, since all the hops in China will only see encrypted data and have no way of knowing what data is passing through the tunnel. It has the additional effect of making all your different kinds of traffic look similar to an eavesdropper. It is important to note that the data is only encrypted until the end of the tunnel, and then the data travels unencrypted to its final destination.

For a more complete explanation of the workings of VPN, and how to use it when you communicate with another person, please read Basic Privacy Security (https://basicinternetsecurity.org/book/basic-internet-security.pdf), ch. 34.

Another reason for using a VPN is more pragmatic, for instance: you want to make use of academic articles that are behind a paywall. If the University of Amsterdam grants you access to these articles, you can use the UvA VPN. The VPN will tunnel you through the UvA network (and hence it gives you an IP-address that belongs to this network) and you will be able to read the material.

Exercise:

Install the UvA-VPN - if you haven't done so yet - and use it whenever you are in need of academic literature behind a paywall.

3.5 Tor Browser

Tor is a system intended to enable online anonymity, composed of client software and a network of servers which can hide information about users' locations and other factors which might identify them. Imagine a message being wrapped in several layers of protection: every server needs to take off one layer, thereby immediately deleting the sender information of the previous server. Because the analogy with 'wrapped layers' this process is also called 'onion routing'.

Use of this system makes it more difficult to trace internet traffic to the user, including visits to Websites, online posts, instant messages, and other communication forms. It is intended to protect users' personal freedom, privacy, and ability to conduct confidential business, by keeping their internet activities from being monitored. TOR is also notorious as it is being used for criminal activities as well. See also the suggested readings about this tension.

The software is open-source and the network is free of charge to use.

Like all current low latency anonymity networks, Tor cannot and does not attempt to protect against monitoring of traffic at the boundaries of the Tor network, i.e., the traffic entering and exiting the network. While Tor does provide protection against traffic analysis, it cannot prevent traffic confirmation (also called end-to-end correlation). Caution: As Tor does not, and by design cannot, encrypt the traffic between an exit node and the target server, any exit node is in a position to capture any traffic passing through it which does not use end-to-end encryption such as TLS. (To go back to our analogy earlier in the worksheet: If your postman is corrupt he might still open the envelope and read the content). While this may or may not inherently violate the anonymity of the source, if users mistake Tor's anonymity for end-to-end encryption they may be subject to additional risk of data interception by third parties. So: the location of the user remains hidden; however, in some cases content is vulnerable for analysis through which also information about the user may be gained.

Therefore a combination of Tor with end-to-end encryption technologies is recommended in high risk cases.

You can self-install Tor. For instructions we recommend to read the manual at https://www.torproject.org/ and download it from there.

You can also choose for the easy option use the in-built TOR option in Brave. Brave has the possibility to open a window that will channel your browsing activities via TOR.

For instructions, see: https://brave.com/new-onion-service/

Exercise:

Read Hoffman (2007, under the suggested readings) and try to visit the Facebook Onion Service. Note that the article is from 2007 and that Facebook by now has changed the link: facebook.com/onion-service . See also “Facebook over Tor”.

Suggested readings:

- “Facebook over Tor” https://www.facebook.com/facebookcorewwwi/posts/3727741430665883
- Hoffman, Chris. How to Access .onion Sites (Also Known as Tor Hidden Services), 12 July 2007. https://www.howtogeek.com/272049/how-to-access-.onion-sites-also-known-as-tor-hidden-services/
- Galloway, Alexander. 2014. “Protocol Futures,” in: *Protocol: How Decentralization Exists after Control*, Cambridge, MA: MIT Press.
- Gehl, Robert.2016. "Power/Freedom on the dark web: A digital ethnography of the Dark Web Social Network." New Media & Society 18 (7), 1219-1235.
- Rap News, "Big Brother is WWWatching You - feat. George Orwell”, RAP NEWS 15. Season 1, https://thejuicemedia.com/season-1/

3.6 Privacy-aware and privacy monitoring Apps

Many apps leak user data. Apps often ask for permissions to your sd-card, device number, connection data, and sometimes contacts. As we cannot cover everything in our worksheet, we limit ourselves only to the following suggestions:

Signal is a free and open source alternative for What's App: https://www.signal.org/

Signal offers an (academically reviewed) protocol for end-to-end encryption. For each communicative session, it generates a pair of keys that are deleted after the session. So, different from public key encryption that is being used in PgP, you don't have a single personal or public key to store, to distribute, or to lose.

We will install Signal during our seminar. An interview with the founder of this app is under the suggested readings.

Secuso, a university based app development initiative.
https://secuso.aifb.kit.edu/english/Net_Monitor.php

Privacy friendly apps can be found at:

The Guardian Project, a human rights focused app development initiative:
https://guardianproject.info/nl/

Exercise:

Install Signal and set up a (small) group chat.

Suggested readings:

- Wiener, Anna. "Taking Back Our Privacy." Interview with Moxie Marlinspike. *The New Yorker*, 19 October 2020. https://www.newyorker.com/magazine/2020/10/26/taking-back-our-privacy.
- ShareLab. Invisible Infrastructures : Mobile permissions. March 2, 2015. https://labs.rs/en/invisible-infrastructures-mobile-permissions/

3.7 Secure video conferencing

Since the COVID-19 pandemic, a lot of meetings have continued to take place in online form. At our university, the default teaching form is now online. One of the tools that gained enormous popularity is Zoom, and it has become the standard tool for lectures and seminars.

Yet, the use of Zoom has been criticised for several reasons. It has been caught up in controversies around the question of whether its communications were encrypted, and in what sense: whether this was end-to-end, whether Zoom was making incorrect promises about its encryption, and whether it concerned only chats or also calls. Other criticisms highlight the point that a tool that was originally meant for companies was transferred to educational settings. This is a discussion about whether the goals of Zoom, and some of its features and its business model, are fitting to the goals of (public) education. You can read more about these controversies and criticisms in the suggested readings.

In this worksheet we invite you to experiment with alternative tools for videoconferencing. If you would like to do a video call one-on-one, Signal provides you the most straightforward option providing end-to-end encryption. For small groups, we will work with Jitsi.

Jitsi is an open source video conferencing tool. The standard implementation is offered at https://meet.jit.si/ but there are several local implementations as well. See for instance, de Waag, an Amsterdam based institute of researchers and designers working on the intersection of technology of society. De Waag offers Jitsi as part of what they call the 'public stack', a collection of open, fair and safe alternatives.

After using Jitsi, we will compare it to Zoom.

(In class) Exercise:

You will be organized into groups. Assign one person who will host the meeting and invite the others. Set up a group-chat in signal so you can easily share the Jitsi and Zoom links.

Go to https://meet.jit.si/ and read the steps to take. Use Chrome browser preferably. Don't forget to allow audio and video permissions. You may want to use the mobile Jitsi-app if you encounter problems.

Start your meeting and test the various features. Such as: screen sharing, background blurring, hand raising, recording, camera, sound, chat. Make notes of your experiences.

Do the same exercise with Zoom. https://zoom.us/

What is needed to start and end-to-end encrypted conversation in Zoom, as compared to Jitsi? In order to answer this question, you need to read the suggested sources.

Suggested material:

- End-to-end encryption in Jitsi: https://jitsi.org/e2ee-in-jitsi/

- End-to-end encryption in Zoom: https://support.zoom.us/hc/en-us/articles/360048660871-End-to-end-E2EE-encryption-for-meetings#h_01ENGDFYWDXHSE16E9BDZ46HKQ
- Gürses, Seda. Rectangles-R-Us - What Happened When the University Went Online? - Part 3 - YouTube." 8 May 2020. https://www.youtube.com/watch?v=fLuCDj3EOKM.
- Gürses has also given a follow up to her talk: Gürses, Seda. 2020. "Digital Public Infrastructures." September 23. https://globaldigitalcultures.org/2020/10/12/full-video-digital-public-infrastructures-online-seminar/.
- Brodkin, Jon. 2020. "Zoom Lied to Users about End-to-End Encryption for Years, FTC Says." Ars Technica. November 9, 2020. https://arstechnica.com/tech-policy/2020/11/zoom-lied-to-users-about-end-to-end-encryption-for-years-ftc-says/.
- Wiener, Anna. "Taking Back Our Privacy." Interview with Moxie Marlinspike. *The New Yorker*, 19 October 2020. https://www.newyorker.com/magazine/2020/10/26/taking-back-our-privacy.

3.8 E-mail security

3.8.1 E-mail

E-mail is one of the oldest forms of communication on the Internet. We often use it to communicate personal and sometimes confidential information. It is important to know that e-mail in its default configuration is not secure.

Most people do not realize how trivial it is for any person on the Internet to forge an e-mail by simply changing the identity profile of their own e-mail program. This makes it possible for anyone to send you an e-mail from some known e-mail address, pretending to be someone else. This can be compared with normal mail; you can write anything on the envelope as the return address, and it will still get delivered to the recipient (given that the destination address is correct). We will describe a method for signing e-mail messages, which prevents the possibility of forgery.

An e-mail message travels across many Internet servers before it reaches its final recipient. Every one of these servers can look into the content of messages, including subject, text and attachments. Even if these servers are run by trusted infrastructure providers, they may have been compromised by hackers or by a rogue employee, or a government agency may seize equipment and retrieve your personal communication.

Then, there is also a range of e-mail hoaxes, phishing e-mails, and e-mails that contain malware. So read your mail critically and don't click links or open attachments that you don't trust.

3.8.2. E-mail client

An e-mail client is a program that can import and organise your e-mails on your computer. For instance: Outlook (on Windows) and Mail (for Macs). A free and open source e-mail client is Thunderbird. Many people also manage their e-mails is via the browser.

Exercise (only for those who prefer a new e-mail client):

Install Thunderbird. Visit your e-mail provider for selecting the right import settings. Take care not to import all your e-mails at once. In the settings menu you can limit the amount of e-mails (for example max. 30 days).

3.8.3 PgP (Suggested)

Note: We will not do PgP in class, but you are free to do this at home by following the referenced tool guide. If you have set it up and want to test it, contact your instructor for a key exchange and write an encrypted e-mail.

Introduction:
This section will introduce you to some basic concepts behind mail encryption. It is important to read to get some feeling of how mail encryption actually works and what its caveats and limitations are. PGP (Pretty Good Privacy) is a well known protocol for e-mail encryption. This protocol allows us to digitally sign and encrypt mail messages. It works on an end-to-end basis: messages will be encrypted on your own computer and will only be decrypted by the recipient of the message. There is no possibility for a 'man- in-the-middle' to decipher the contents of your encrypted message. This excludes the subject lines and the 'from' and 'to' addresses, which unfortunately are not encrypted in this protocol by default, although some e-mail clients now do so.

We only introduce the basic concepts, and won't practice this in the seminar. Due to online teaching, we cannot provide you the support needed. (Instead, we practically engage with secure video conferencing). If you want to try this at home, please consult the tool guide of the Electronic Frontier Foundation, as this one is being kept up to date.

The encryption method in PgP is asymmetric public key encryption. We recommend you to first watch this metaphorical explanation for a basic understanding of the form of encryption behind PgP: ‘How asymmetric (public key) encryption works’

https://www.youtube.com/watch?v=E5FEqGYLL0o

Explanation:

Using a Key-Pair to encrypt your mail
A crucial concept in mail encryption is the usage of so-called key-pairs. A key-pair is just two separate files sitting on your hard disk or USB stick. They can be generated via software that you

integrate with your e-mail client or webmail application. Whenever you want to encrypt mails for a certain mail-account, you will need to have these files available to yourself in some form. A key-pair consists of the two different keys: a public key and a secret key.

The public key: you can give this key to other people, so they can send you encrypted mails. This file does not have to be kept secret.

The secret key: this basically is your secret file to decrypt emails people send to you. It should never be given to someone else.

Sending encrypted mails to other people: You need their Public Key

I have five colleagues at work and I want to send encrypted mails to them. I need to have public keys for each of their addresses. They can send me these keys using ordinary mail, or they can give them to me in person, or put them on a USB stick, or they can have their keys on a website. It doesn't matter, as long as I can trust those keys really belong to the person I want to correspond with. My software puts the keys on my ` keyring', so my mail application knows how to send them encrypted mails.

In the video listed above the metaphor of a ‘padlock’ is used to explain the process. To encrypt messages, you lock your message with your friend’s padlock (the ‘public key’), of which your friend has distributed several copies. Your friend is the only one that can open the padlock with the corresponding private key.

Receiving encrypted mails from other people: They need your Public Key

For my five (or thirty) colleagues to be able to send me encrypted mails, the process goes the other way around. I need to distribute my public key to each of them.

So, if we go back to the padlock, if your friend wants to encrypt a message to you, your friend encrypts the message with your public padlock (public key) and you can open it with your private key.

Conclusion: Encryption requires Public Key distribution

All the people in a network of friends or colleagues wanting to send each other encrypted emails, need to distribute their public keys to each other, while keeping their secret keys a closely guarded secret.

Suggested readings:

- Zimmerman, Phil. 1999. “Encryption, Privacy, and Crypto-Anarchism” and “How PGP Works/Why Do You Need PGP?” In: Peter Ludlow (Ed.) *High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace*, 173-184. (Available on Canvas).

Exercise:

Try to answer:
1a. What is a public key server?
1b. How could a public key server be used to conduct a social network analysis?

4. Data ethics

When doing research we encounter several ethical questions. Think of: What is the impact of research? Are the ways we use data legitimate? If we collect data about people, do they need to be informed? Does the data collection concern ‘personal data’? Can the results of certain studies do harm after publication?

In this unit we focus ourselves primarily on ethical concerns that may arise when data is not kept securely. In the lecture and in the required literature you can find examples of why research data should be treated with care. In addition, this worksheet has suggested several tools that help you to work with data more securely.

Research that takes place at the University of Amsterdam needs to comply with certain standards regarding data security. Especially when it concerns the processing of personal data. This includes digital research. If a researcher at the UvA wants to do a study that requires the processing of personal data, the Ethics Committee has to be informed and will only agree if the researcher outlines the necessary steps that will guarantee an ethical research process. Think of, for instance, guaranteeing the anonymity of data subjects and the secure storage of data. If one intends to do interviews, one has to document that the respondents gave their informed consent.

If a student does a research project in the context of the BA thesis that should pass the ethics committee, the supervisor will do this application. In this application, the supervisor will be asked whether the researcher - in this case: the student - will be working with human participants and whether the researcher will collect personal data. 'Personal data' includes all information on the basis of which someone can be identified or which can be directly or indirectly traced back to a natural person. This information includes a name, identification number, telephone number, assessments and research data, but also a combination of data which can jointly result in an image so unique that it can only relate to one person.

Therefore, the student needs to always inform the supervisor on time about the plans for data collection and data processing.

Exercise:

Read the suggested material listed below. Discuss with your fellow students if you ever encountered ethical concerns or ethical debates during your studies. These can be concerns relating directly to your own research projects or debates that relate to research at the university more generally.

Suggested material:

- The web page of the UvA ethics committee: https://aihr.uva.nl/about-aihr/ethics-committee/ethics-committee.html

- The document that explains the GDPR instructions: https://aihr.uva.nl/binaries/content/assets/subsites/amsterdam-institute-for-humanities-research/ethics/ethics--gdpr-instruction-240621.pdf

5. Practicing security

At home and in the seminars:

1. Read this worksheet thoroughly and spend some time exploring the linked resources and practicing with the exercises;

2. In the first seminar we will very briefly discuss the worksheet, and there is room for questions about the lecture. Then you can start working on the worksheet with the help of the teacher and the teaching assistants.

3. In the second seminar, we set up secure video conferencing and discuss data ethics in research practice. There is time to get feedback on what you have prepared for your assignment.

Exercise (complete/incomplete)

Answer the following **three questions** (all together 1000 words +/-10%). Reference in your assignment at least 5 literature sources out of which at least 3 sources from the required reading list (see next page) of this unit. Other sources can be taken from the suggested sources mentioned in this worksheet. Don't forget to reread the appropriate sections in this worksheet.

Hand in your assignment via Canvas before Monday 6 December 17:00.

1. Data ethics

Read the web page of the UvA ethics committee:
https://aihr.uva.nl/about-aihr/ethics-committee/ethics-committee.html

Focus in particular on the document that explains the GDPR instructions:

https://aihr.uva.nl/binaries/content/assets/subsites/amsterdam-institute-for-humanities-research/ethics/ethics--gdpr-instruction-240621.pdf

Answer the following questions:

a. What is the Ethics Committee of the Faculty of Humanities? Why do you think such a committee is needed? (max 100 words)

b. If you do a research project for your BA thesis, your instructor might need to do an application with the Ethics Committee. Explain why. (Max 30 words)

c. What is the GDPR? (Max 50)

d. Give an example of a possible *offline research project* that would require the processing of personal data? What measures would you take? For example: How would you organize data collection, storage and transmission? (max 100 words)

e. Give an example of a possible *digital research project* that would require the processing of personal data. What measures would you take? For example: How would you organize data collection, storage and transmission? (max 100 words)

2. Secure video conferencing

Set up a test video call with a fellow student. First try Jitsi and then Zoom. Make use of the suggested readings to better understand all the features.

a. Did things work well? Which things did not? Do you know why?

b. Try to enforce *end-to-end* encryption in Jitsi. Explain your steps and whether you succeeded. If not, what could be the reason for it?

c. Try to enforce *end-to-end* encryption in Zoom. Explain your steps and whether you succeeded. If not, what could be the reason for it?

3 Anonymity

Go to facebook.com/onion-service and try to get to the login page. You don't need to login with your credentials, only open the login page.

a. How did you manage to visit the login page?

b. Give two reasons why people would like to use a version of Facebook that is accessible via Tor?

6. Readings

6.1 Required readings and sources

- Gürses, Seda: Rectangles-R-Us - What Happened When the University Went Online? - Part 3 - YouTube." 8 May, 2020. https://www.youtube.com/watch?v=fLuCDj3EOKM.
- Hoffman, Chris. How to Access .onion Sites (Also Known as Tor Hidden Services), 12 July 2007. https://www.howtogeek.com/272049/how-to-access-.onion-sites-also-known-as-tor-hidde

n-services/

- Tanczer, L. “Introduction” (p. 3-7) and Kazansky, Becky and Stefania Milan “Infrastructure and Protocols for Privacy-Aware Research (p. 26-36). In: Tanczer, L., R. J. Deibert, D. Bigo, M. I. Franklin, L. Melgaço, D. Lyon, B. Kazansky, and S. Milan. 2019. “Online Surveillance, Censorship, and Encryption in Academia.” International Studies Perspectives. https://doi.org/10.1093/isp/ekz016.
- UvA Ethics Committee: https://aihr.uva.nl/about-aihr/ethics-committee/ethics-committee.html
- Wiener, Anna. “Taking Back Our Privacy.” Interview with Moxie Marlinspike. The New Yorker, 19 October 2020. https://www.newyorker.com/magazine/2020/10/26/taking-back-our-privacy

6.2 Suggested readings

Suggested readings are listed per section in the worksheet.