

Module 4: Life on the net

The fourth module focuses on learning the characteristics of online human - human communication and information exchange through social networks, highlighting the roles AI plays in them. This module aims to equip educators and learners with the necessary tools to work towards digital wellbeing and healthy digital consumption. Concepts such as privacy, anonymity and confidentiality will be explored. Moreover, activities on this module will offer a broad understanding of the term 'fake news', training the skills to identify sources of knowledge and to exercise critical thinking. The module will include notions for respectful communication (netiquette), exploring the concept of *dangerous* and *hate speech online*. In class, the trained teachers will be able to run activities such as analysing news from social networks, identifying reliable sources in fact-checking, debunking viral misinformation and analyse extracts from real-life chats taken from instant messaging platforms to understand language and communication patterns and countering hate. The key element of the learning process is gaining acquaintance with the notion of "Digital Citizenship"; informed digital citizens will make informed choices once conducting their lives on the net. Students will be trained to engage, critically and competently in the digital world, while exploring how the use of digital and AI technologies can impact not only everyday life but also participation and engagement at the societal level.

Topic 4.1 - Knowledge on the web: disinformation, fake news, fact checking and critical thinking

This topic focuses on the role of AI technologies in the creation and dissemination of knowledge and information on the web. It addresses issues like fake news, fact checking and critical thinking. Its aim is to foster students' ability to identify misinformation mechanisms and to respond to those adequately, i.e. according to the ethical principles explored in Module 3.

The exposure to and increasing volume of information in the age of the Internet has two major interrelated risks. Firstly, it runs the risk of overwhelming people with information overload. And secondly, this overwhelming volume of information can be used intentionally to spread disinformation or manipulate facts to gain political and economic power (infodemic). Given these premises, there is the need to better equip young citizens with resilience to "information disorder" (Wardle and Derakhshan 2017, McDougall 2019) by investing in critical media literacy to detect fake news through fact checking and de-bunking processes, to evaluate evidence critically, to prevent "filter bubbles" and "echo chambers" and create a level playing field for an

inclusive online environment. This module will also explore how AI tools can contribute both to spreading misinformation and to tackling it.

Useful **resources**

Fake news, information disorder and fact checking:

SOMA, Social Observatory for Disinformation and Social Media Analysis.

<https://www.disinfobservatory.org/>.

Including tool several resources and tools, e.g. to examine fake news:

<https://www.disinfobservatory.org/investigations/>

Wardle, C., & Derakhshan, H. (2017) "Information disorder: Toward an interdisciplinary framework for research and policy making." Council of Europe.

<http://tverezo.info/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-desinformation-A4-BAT.pdf>

Foundation European Institute Outsourcing - FEIO, Combating Fake News: Handbook for Youth Workers (2019).

https://www.salto-youth.net/downloads/toolbox_tool_download-file-2402/Handbook%20-%20How%20to%20Fight%20Fake%20News.pdf

Council of Europe. Resources on Dealing with propaganda, misinformation and fake news

<https://www.coe.int/en/web/campaign-free-to-speak-safe-to-learn/resources-on-dealing-with-propaganda-misinformation-and-fake-news> E.g. Council of Europe, [Digital Citizenship Education](#) (2017)

The Conversation. How to spot fake news: an Expert Guide for Young People

<https://theconversation.com/how-to-spot-fake-news-an-experts-guide-for-young-people-88887>

Pew Research. The Future of Free Speech, Trolls, Anonymity and Fake News Online

<https://www.pewresearch.org/internet/2017/03/29/the-future-of-free-speech-trolls-anonymity-and-fake-news-online/>

WHO "Let's flatten the infodemic curve"

<https://www.who.int/news-room/spotlight/let-s-flatten-the-infodemic-curve>

Filter bubbles and echo-chambers:

Fondation Descartes, "Filter Bubbles and Echo Chambers"

<https://www.fondationdescartes.org/en/2020/07/filter-bubbles-and-echo-chambers/>

DiFranzo, Dominic and Gloria-Garcia, Kristine (2017) "Filter Bubbles and Fake News" Association for Computing Machinery Magazine for Students 23 (3): 32-35. <https://doi.org/10.1145/3055153>

Critical thinking and media literacy

Mcdougall, J. (2019) "Media literacy versus fake news: critical thinking, resilience and civic engagement." Media Studies, 10(19), 29-45.

<https://hrcak.srce.hr/ojs/index.php/medijske-studije/article/view/8786>

Columbia University - Center for Professional Education of Teachers, "Information Overload: Combating misinformation with critical thinking."

<https://cpet.tc.columbia.edu/news-press/information-overload-combating-misinformation-with-critical-thinking>

Garcia, Laura and Tommy Shane (2021) "A guide to prebunking: a promising way to inoculate against misinformation." First Draft.

<https://firstdraftnews.org/articles/a-guide-to-prebunking-a-promising-way-to-inoculate-against-misinformation/>

GOViral! Game to help protect against Covid-19 misinformation

<https://www.goviralgame.com/en/intro>

Role of AI

Horizon, the EU Research and Innovation Magazine, "Can AI help end fake news?"

<https://ec.europa.eu/research-and-innovation/en/horizon-magazine/can-artificial-intelligence-help-end-fake-news>

Politico, "AI Decoded: How Cambridge Analytica used AI."

<https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-how-cambridge-analytica-used-ai-no-google-didnt-call-for-a-ban-on-face-recognition-restricting-ai-exports/>

Supasorn Suwajanakorn, Steven M. Seitz, Ira Kemelmacher-Shlizerman SIGGRAPH 2017, "Teaser -- Synthesizing Obama: Learning Lip Sync from Audio"

https://www.youtube.com/watch?v=MVB6_o4cMI

See also: <http://grail.cs.washington.edu/projects/AudioToObama/>

Udupa, Sahana, Elonnai Hickok, Antonis Maronikolakis, Hinrich Schuetze, Laura Csuka, Axel Wisioerek, Leah Nann (2021) "AI, Extreme Speech and the Challenges of Online Content Moderation". AI4Dignity Project. <https://doi.org/10.5282/ubm/epub.76087>

Topic 4.2 - **Identifying and countering hate speech online**

Online hate speech is a growing phenomenon. Every day, the media, especially the Internet, convey messages that encourage intolerance towards individuals or social groups on the basis of their social and cultural characteristics and/or their personal and political opinions. Although there is not an international legal definition of hate speech, in broad terms we can define this concept by recalling "any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor" (see the *UN Strategy and Plan of Action on Hate Speech*).

Increasingly, hate speech is targeting religious communities and/or their beliefs. As is the case with ethnic and racial discrimination, which are often associated with character or even biological aspects of individuals, hate messages tend to identify members of a particular religious community with certain negative characteristics (e.g., anti-Semitism has generally stigmatized Jews as greedy and dishonest, or even cruel). However, hate speech can also target beliefs and precepts or even values associated with them. In the case of anti-Muslim discourse, to exemplify, the hostility is based on the representation of Islam as a religion incompatible with Western values and that provides violence as a tool to advance the faith.

In addition, there is an increase of those forms of hostility targeting ethnicity, migratory status, disability, gender identity and sexual orientation (misogyny and homo-trans-phobia).

It is less and less possible to deal separately with online and offline hate phenomena, because communication technologies and social media are part of our daily life and, more generally, of our identity. In this sense, if some dynamics peculiar to online communicative interaction have been identified (such as the escalation of negative emotions and the polarization of arguments), the orientations expressed on the web must be addressed in their permeable relationship with the lifestyles and face-to-face interactions experienced in everyday life.

Indeed, there is evidence about the possibility of a link between online hate speech and hate crimes or acts of violence in "real life". Scholars are studying the possible responsibility of AI in this. The algorithms that mediate users' experiences online are designed to maximize their engagement, which often inadvertently promote extreme content. As an example, YouTube's

autoplay function, in which the player, at the end of one video, tees up a related one, can be especially pernicious.

Young people are overexposed to hate speech, both because of their now stable online presence, and because they are not yet in possession of all the tools necessary for a critical analysis of these communicative phenomena. Young people tend to show little awareness about the possible effects of hate speech for the victims, as well as of the fact that the internet is a public space and it is not possible to guarantee true anonymity to their posts and interventions in it.

Digital education is therefore key in order to reduce the exposure of young people to hate speech. Contrary to the myth of the "digital natives", in fact, young people do not automatically and critically master new technologies just because they use a smartphone. On the contrary, as a paradox, children who make exclusive use of this device acquire fewer technical and creative skills than those who also use other means of navigation. Furthermore, as the Council of Europe has repeatedly stated, we can remark the importance of intercultural education and education about religions as a means of combating prejudice and intolerance and promoting mutual understanding. In fact, it is necessary to transmit to new generations knowledge, interpretative tools and social competences for the understanding of a plural social scenario, in which cultural and religious diversities play a crucial role both in the genesis of conflict and in overcoming it.

With respect to the need for countering online hate speech, a balance between safeguarding individuals' freedom of expression, at the basis of our democracies, and restricting the exercise of this freedom in order to preserve the dignity of the other must be found.

In response to this challenge, it is widely acknowledged to be effective to use counter-narratives. Counter-narratives (or alternative narratives) combat hate speech by discrediting and deconstructing the narratives on which they are based. They also propose (alternative) narratives based on human rights and democratic values, such as openness, respect for difference, freedom, and equality. Possible strategies to adopt in producing counter-narratives are: working to modify terms or 'de-emphasize' them in their effects; humanizing victims; combating prejudice through verifiable information and data.

In recent years, at the international level several attempts to improve the AI-based tools potential in combating hate speech have been put in place. Most hate speech monitoring technologies involve a mix of manual and automated processes. Social media platforms rely on a combination of artificial intelligence, user reporting, and staff known as content moderators to enforce their rules regarding appropriate content.

In fact, one of the main difficulties in the automatic detection processes of hate speech is that in order to identify and understand hate messages it is necessary to recognize elements of context in their production, circulation and reception. Let's think about the role of the political context and current events reported by the media in amplifying or reducing the generation of these

messages; the role of humor as a possible code of the message; the different possible understandings of the content by different targets, etc. Also to be considered are the additional difficulties of controlling cyberspace without altering the advantages of the internet (ease and speed communication).

Useful resources

Digital wellbeing

- <https://www.benesseredigitale.eu/pubblicazioni/>
- <https://www.mulino.it/isbn/9788815280992>
- Gui M, Büchi M. From Use to Overuse: Digital Inequality in the Age of Communication Abundance. *Social Science Computer Review*. 2021;39(1):3-19. doi:[10.1177/0894439319851163](https://doi.org/10.1177/0894439319851163)

Impact of online communication

- The impact of AI on communication networks and services, International Telecommunication Union, https://www.itu.int/dms_pub/itu-s/opb/journal/S-JOURNAL-ICTF.VOL1-2018-1-PDF-E.pdf
- The Future of Human Communication: How Artificial Intelligence Will Transform the Way We Communicate, <https://www.quantified.ai/blog/artificial-intelligence-in-communication>

Impact of social media on religious practices and values among young adults:

- Moberg, M., & Sjö, S. (Eds.). (2020). Digital Media, Young Adults, and Religion: An International Perspective (1st ed.). Routledge. <https://doi.org/10.4324/9781351010610>
- Two social lives: How differences between online and offline interaction influence social outcomes, https://escholarship.org/content/qt94n9w8b9/qt94n9w8b9_noSplash_293949a5e051fff8e1fdcc9ffc168c4.pdf?t=qdtezb
- Communicating across Cultures in Cyberspace, A bibliographical review of intercultural communication online, <https://core.ac.uk/download/pdf/12174297.pdf> \\
- Some data and findings on Online Harassment <https://www.pewresearch.org/internet/2014/10/22/online-harassment/>
- Council of Europe, “Model of Governance of Online Hate speech, On the emergence of collaborative governance and the challenges of giving redress to targets of online hate speech within a human rights framework in Europe”, <https://rm.coe.int/models-of-governance-of-online-hate-speech/16809e671d>

MEMES as a “case study” of on-line hateful communication among youngsters

- Facebook data set and policy on scrutinizing memes post in FB groups against hate speech with the help of AI and algorithms, <https://ai.facebook.com/blog/hateful-memes-challenge-and-data-set/>

- On the difficulties of AI to fully recognize hateful memes
<https://venturebeat.com/2020/12/01/ai-still-struggles-to-recognize-hateful-memes-but-its-slowly-im-proving/>
- “The hateful memes challenge: Detecting Hate speech in multimodal memes”,
<https://proceedings.neurips.cc/paper/2020/file/1b84c4cee2b8b3d823b30e2d604b1878-Paper.pdf>
- “Detecting hateful memes using AI”,
<https://www.eng.ed.ac.uk/about/news/20201201/detecting-hateful-memes-using-ai>
- <https://www.drivendata.org/competitions/64/hateful-memes/>
- And Policy Brief on AI and extreme speech:
<https://www.disinfoobservatory.org/artificial-intelligence-extreme-speech-and-the-challenges-of-online-content-moderation/>

Topic 4.3 - **Privacy, anonymity and confidentiality online**

This topic focuses on the concepts of on-line anonymity, privacy and confidentiality as they are experienced by students. The topic outlines the difficulties of internet anonymity and shows to what extent actions of self-disclosure, or lack thereof, can affect people’s privacy online.

The growing use of social media, commercial platforms and AI powered on-line platforms is fostering, as a side effect, technical advancement in data retrieval. Indeed, everyday there is a new opportunity to make use of people’s personal information for both ethical and unethical reasons.

Everyday, millions of people leave their information footprint while using internet services. This **footprint** consists of all of the information that people post, the **hidden data attached** to those posts by the services people use, the record of people’s online activities, and also the **inferences** that can be drawn from putting that collective information together.

It can be surprising to realize how much hidden information about your activities is stored and transferred in modern systems. Indeed, to begin with, every Internet connection automatically transmits the sender’s IP address along with whatever information the user sends. In other words, every time we visit a website, involuntarily we are leaving a **trace**. Websites and apps, but also digital devices (such as TVs, fitness trackers and tablets) store this and other data about users’ behavior. In addition, many apps and devices **(such as TVs, tablets, and mobile phones) exchange information, often without the user knowing**. Mobile apps often communicate the device’s **location**, as well as configuration information hidden from the user.

Even when we are not using the Internet directly, your everyday activities generate data. Finally, data collected about someone from many sources, including information posted intentionally, (and the **metadata** attached to these contents), can be combined to infer even more information about that person. For example, a **tweet saying we are on vacation** has the potential domino effect to reveal personal information concerning our very private lives (where our house is located, where we work or go to school, etc.) and becoming a good target for a robbery.

Useful resources

Commonsense education provides an interesting portfolio for digital citizenship education:

<https://www.commonsense.org/education/digital-citizenship/curriculum?topic=privacy--security>

<https://ethicsunwrapped.utexas.edu/video/causing-harm>

Teaching privacy is a useful resource to teach both children how to keep their privacy while surfing the internet. The website presents a set of incredible tools for teachers, informal educators and parents as well.

<https://teachingprivacy.org/>

The Need for Respectful Technologies: Going Beyond Privacy

<https://dighum.ec.tuwien.ac.at/perspectives-on-digital-humanism/the-need-for-respectful-technologies-going-beyond-privacy/>

The Real Cost of Surveillance Capitalism as emerging form of economy: Digital Humanism in the US and Europe

<https://dighum.ec.tuwien.ac.at/perspectives-on-digital-humanism/the-real-cost-of-surveillance-capitalism-digital-humanism-in-the-us-and-europe/>

Topic 4.4 - AI-related opportunities for participation and engagement (digital citizenship)

The aim of this Topic is to train students ability to **engage, critically** and competently in the digital world. Students will learn how the use of digital and AI technologies can impact not only everyday life but also participation and engagement at the societal level. The core of this module is centered around the notion of “Digital Citizenship” and how **good online practices** form a critically and competently engaged “Digital Citizen”.

The concept of Digital Citizenship has now entered the agenda of state authorities and students curricula as much as digital tools have entered the lives of people. The development of digital technologies and their use on a daily basis has become a necessity, as a consequence there is the need to prepare students in a way to have them **communicate and collaborate safely and responsibly in online environments**. The rules for **correct and responsible technology** usage have brought the concept of Digital Citizenship to the fore, and its importance is gradually increasing and will become an indispensable part for the world of tomorrow.

The definition of Digital Citizenship adopted by the Council of Europe reads as follows “the competent and positive engagement with digital technologies (creating, working, sharing, socializing, investigating, playing, communicating and learning); participating actively and responsibly (values, skills, attitudes, knowledge) in communities (local, national, global) at all levels (political, economic, social, cultural and intercultural); being involved in a double process of lifelong learning (in formal, informal and non-formal settings) and continuously defending human dignity.” In this sense, there is a strong emphasis on digital citizenship education noting that the competences which citizens need to acquire if they are to participate effectively in a culture of democracy are not acquired automatically but instead need to be learned and practised.

Useful resources

Digital Citizenship

Mossberger, Karen; Tolbert, Caroline J.; McNeal, Ramona S. (October 2007). Digital Citizenship, The Internet, Society & Participation. MIT Press.

Ohler, Jason B. (2010-08-31). Digital Community, Digital Citizen. SAGE Publications.

Öztürk, Gülcan. "Digital citizenship and its teaching: A literature review." Journal of Educational Technology and Online Learning 4, no. 1 (2021): 31-45.

<https://files.eric.ed.gov/fulltext/EJ1286737.pdf>

Martin, Florence, Tuba Gezer, and Chuang Wang. "Educators' perceptions of student digital citizenship practices." Computers in the Schools 36, no. 4 (2019): 238-254.

Choi, Moonsun. "A concept analysis of digital citizenship for democratic citizenship education in the internet age." Theory & research in social education 44, no. 4 (2016): 565-607.

<http://global-awareness.org/resources/misc/Choi2016ConceptAnalysisDigitalCitizenship.pdf>

Council of Europe

<https://www.coe.int/en/web/digital-citizenship-education/digital-citizenship-and-digital-citizenship-education>

Digital Natives, Citizens of a changing world,

<https://net-ref.com/white-paper-fostering-digital-citizenship-in-the-classroom/>

<https://www.oecd.org/education/ceri/21st-Century-Children-as-Digital-Citizens.pdf>

Digital footprint and Inequality

Micheli, Marina (June 2018). "Digital Footprints: An Emerging Dimension of Digital Inequality". Journal of Information Communication and Ethics in Society: 7.

Social Media rights

<https://safelab.medium.com/know-your-social-media-rights-7e87a6c45540>

E-Democracy and Society

Promise and Problems of E-Democracy: Challenges of Online Citizen Engagement (PDF). Paris: OECD. 2003. p. 162. ISBN 9264019480.

Taken-for-granted technologies of digital citizens

QR codes for payments

<https://www.docomodigital.com/resource/blog/covid-19-pushes-qr-code-payment-adoption/>

QR codes and Covid-19

<https://www.kaspersky.com/blog/secure-futures-magazine/qr-codes-business-future/39236/>