MediaFutures

LUISS Data Lab

LUISS Master in Giornalismo

Introduction to AI&Data Journalism
Automated Journalism, Fake News Detection & Social Data Intelligence

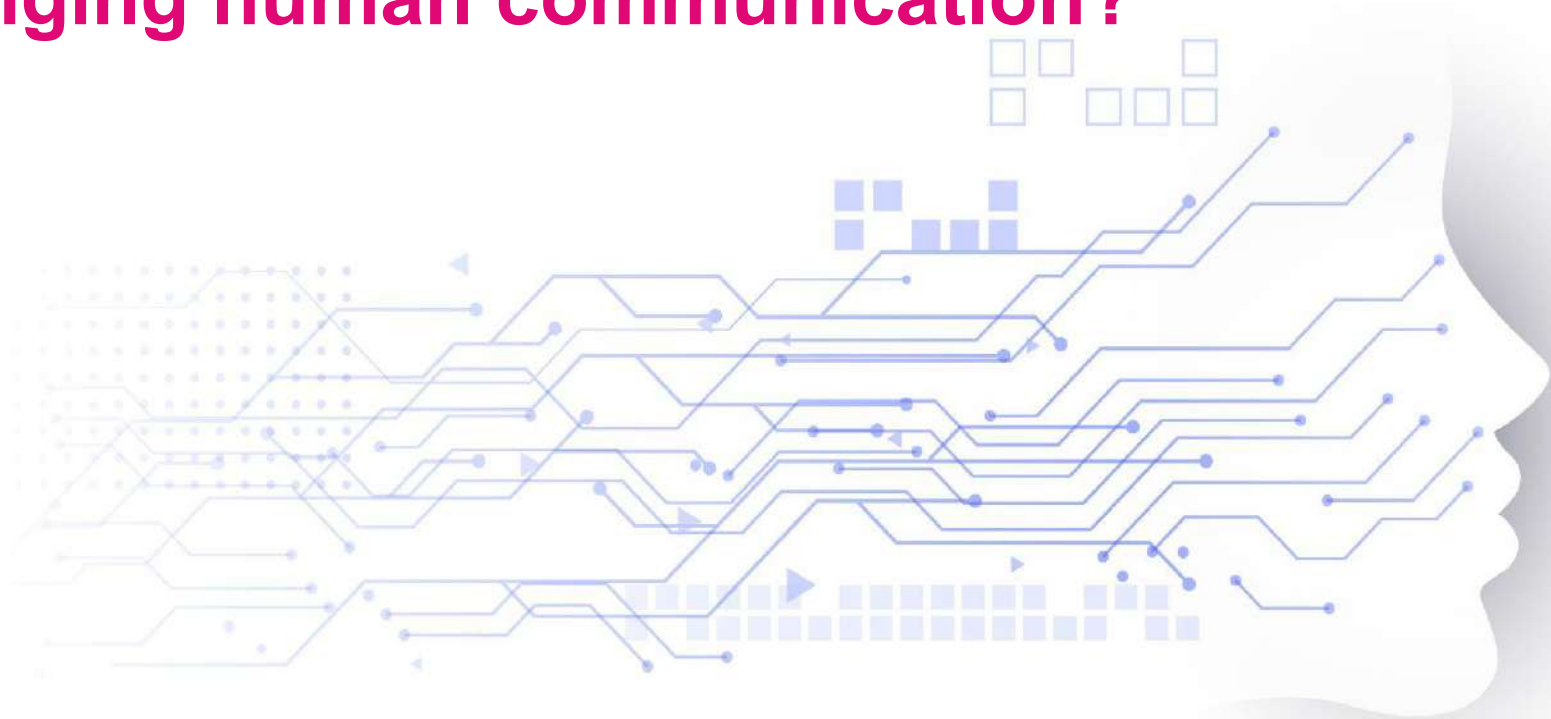
Italian Digital Media Observatory

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951962

Outline

1. **The Scenario**
    a. Why does technology, automation and artificial intelligence affect us?
2. **The Domain**
    a. First step: understand the domain
    b. Second step: actively use your knowledge
        i. Zeta Luiss
        ii. Catchy Summa
        iii. Social Data Intelligence
        iv. Compass & Bot Detection

How automation, AI and BIG Data are changing human communication?

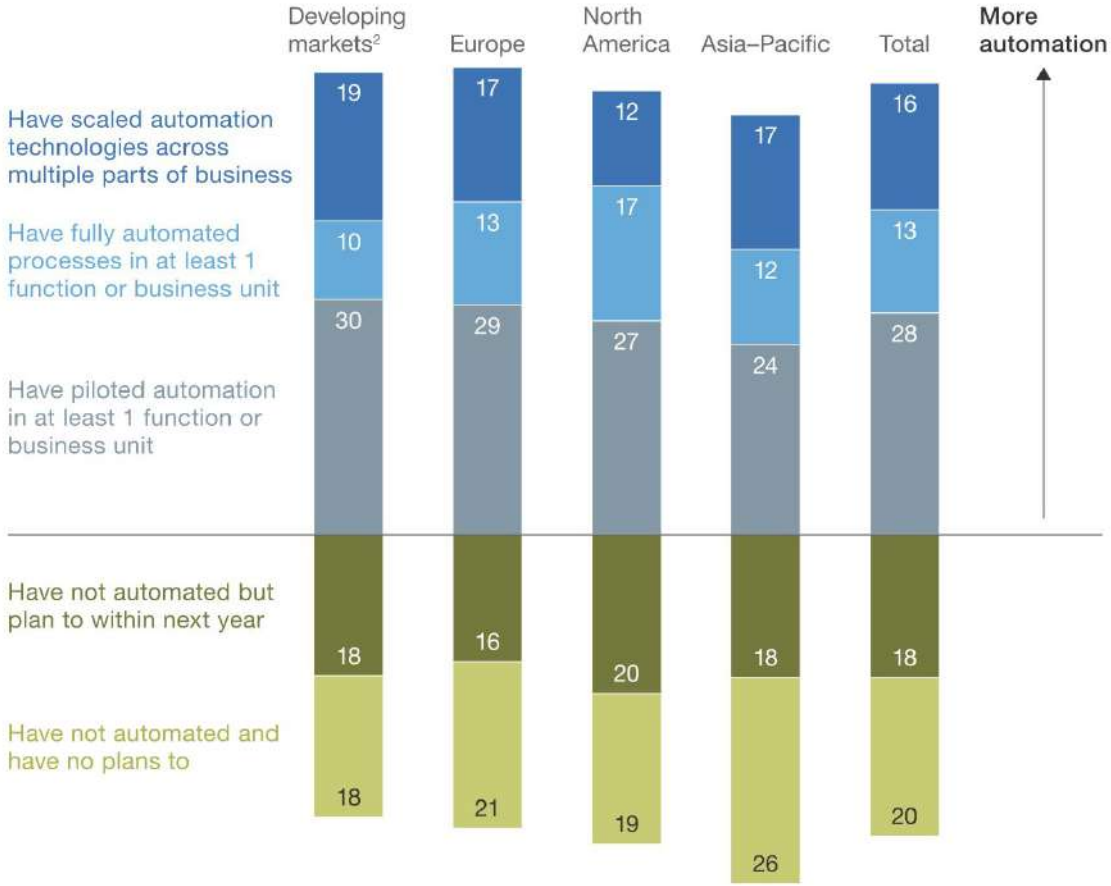


1. The current scenario in which we operate

The automation imperative

In September 2018, **Mckinsey & Company** titled one of its executives reports *The Automation Imperative* recognizing, precisely in the automation of processes and services, a phenomenon of global and profound transformative scope.

- **57% of enterprises have initiated automation processes** (16% at scale across different parts of their business, 13% partially, 28% for pilot projects)



[1]Respondents who answered "don't know" are not shown. Total n = 1,303; in developing markets, n = 373; in Europe, n = 479; in North America, n = 281; and in Asia–Pacific, n = 170.
[2]Includes respondents in China, India, Latin America, Middle East, and North Africa.

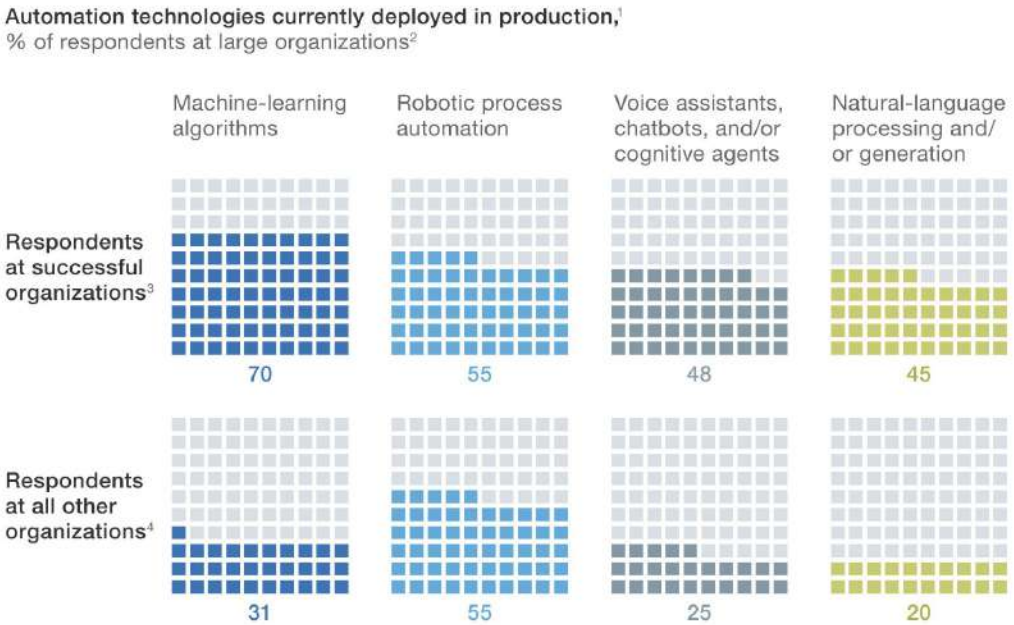
The automation imperative

70% of enterprises choose to use automation technology during the production of **machine-learning algorithms developments**

55% for robotic process automation (robotic automation of business processes, both mechanical-routine and cognitive activities)

48% for voice assistants, chatbots, and/or cognitive agents

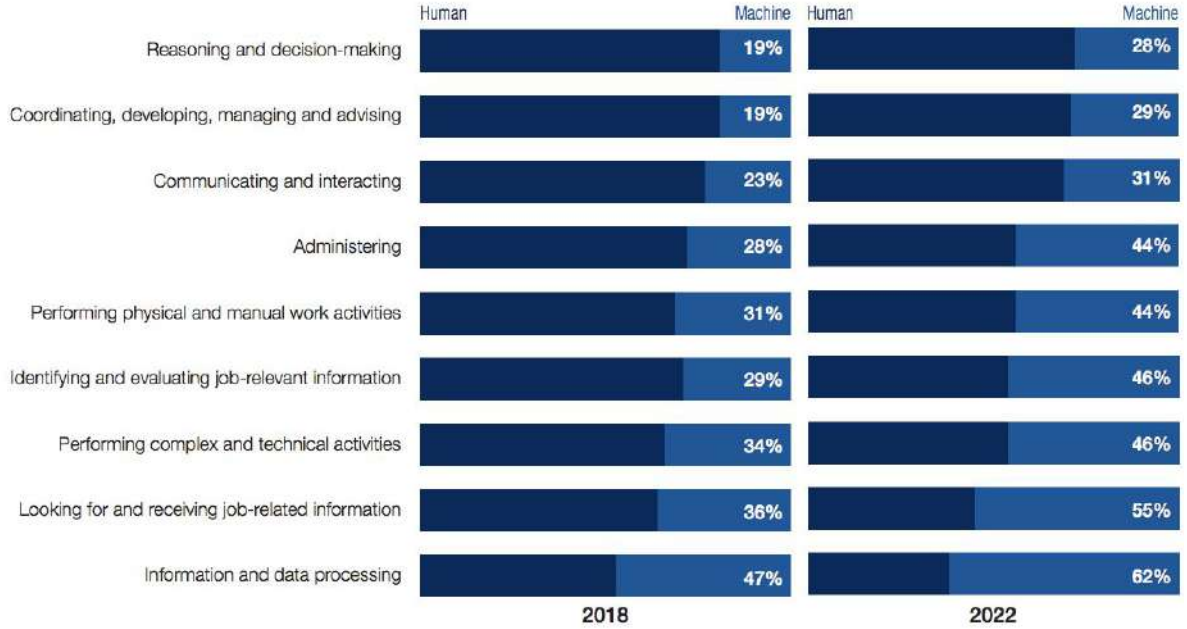
45% for natural-language processing and/or generation algorithms.



Jobs' automation

In 2025, for the first time in human history, **the share of the time performed by machines will exceed those performed by humans (48%) -** according to the World Economic Forum.

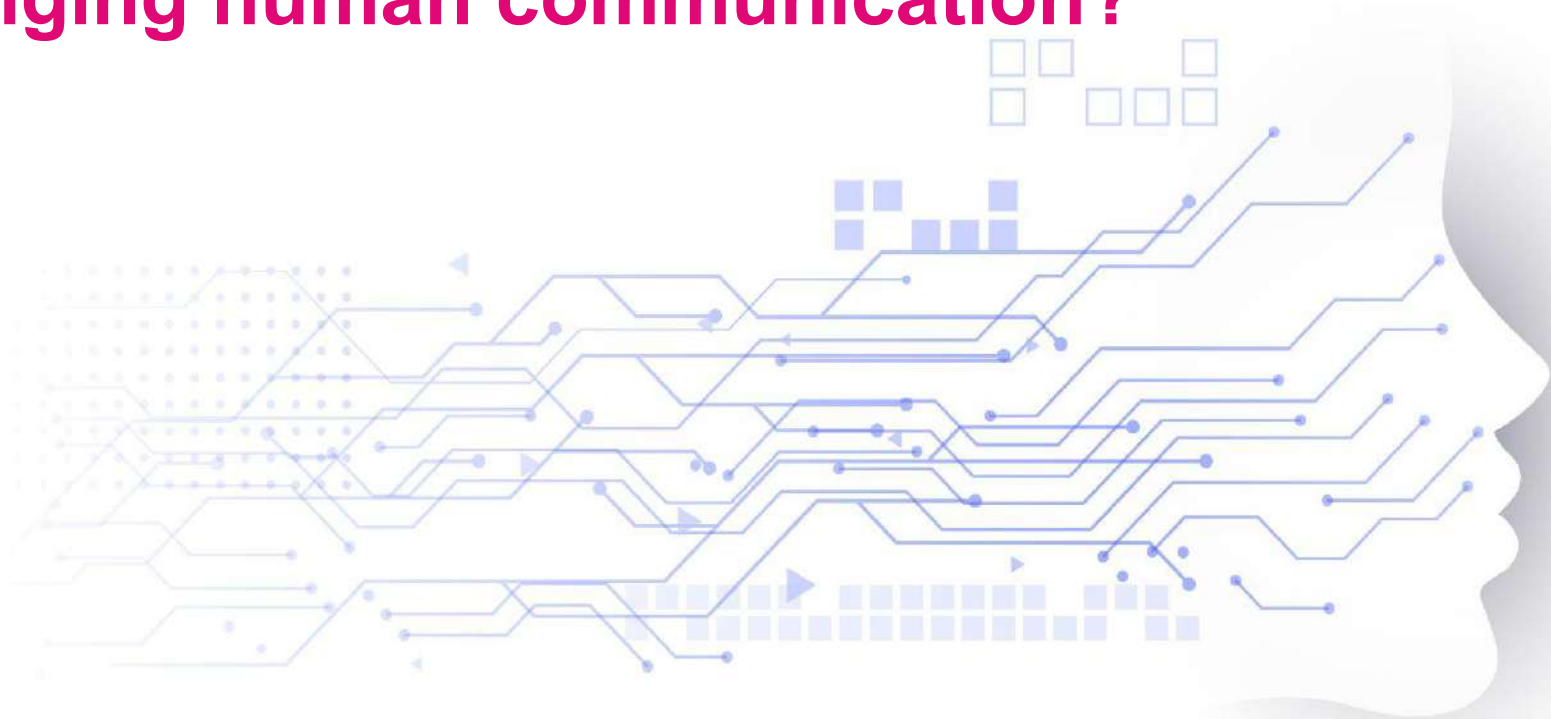
Figure 5: Ratio of human-machine working hours, 2018 vs. 2022 (projected)



Source: Future of Jobs Survey 2018, World Economic Forum.

https://reports.weforum.org/future-of-jobs-2018/?doing_wp_cron=1653749407.1691761016845703125000

How automation, AI and BIG Data are changing human communication?



2. Human, too human

Human, too human

Artificial intelligence and automation process deeply affect our lives, work,and political and institutional decisions. **We cannot leave the understanding, explaining and writing of algorithms only to technical experts.**

“Technically speaking, algorithms are a series of instructions telling a computer, precisely and unequivocally, what to do”. Cosimo Accoto, The Given World.

Ethical Dilemmas, Human:

- Who writes the algorithm?
- Is its point of view biased?
- How do we teach a self-driving car how to behave in the case of an accident?
- How do we distinguish correct from incorrect machine behavior?

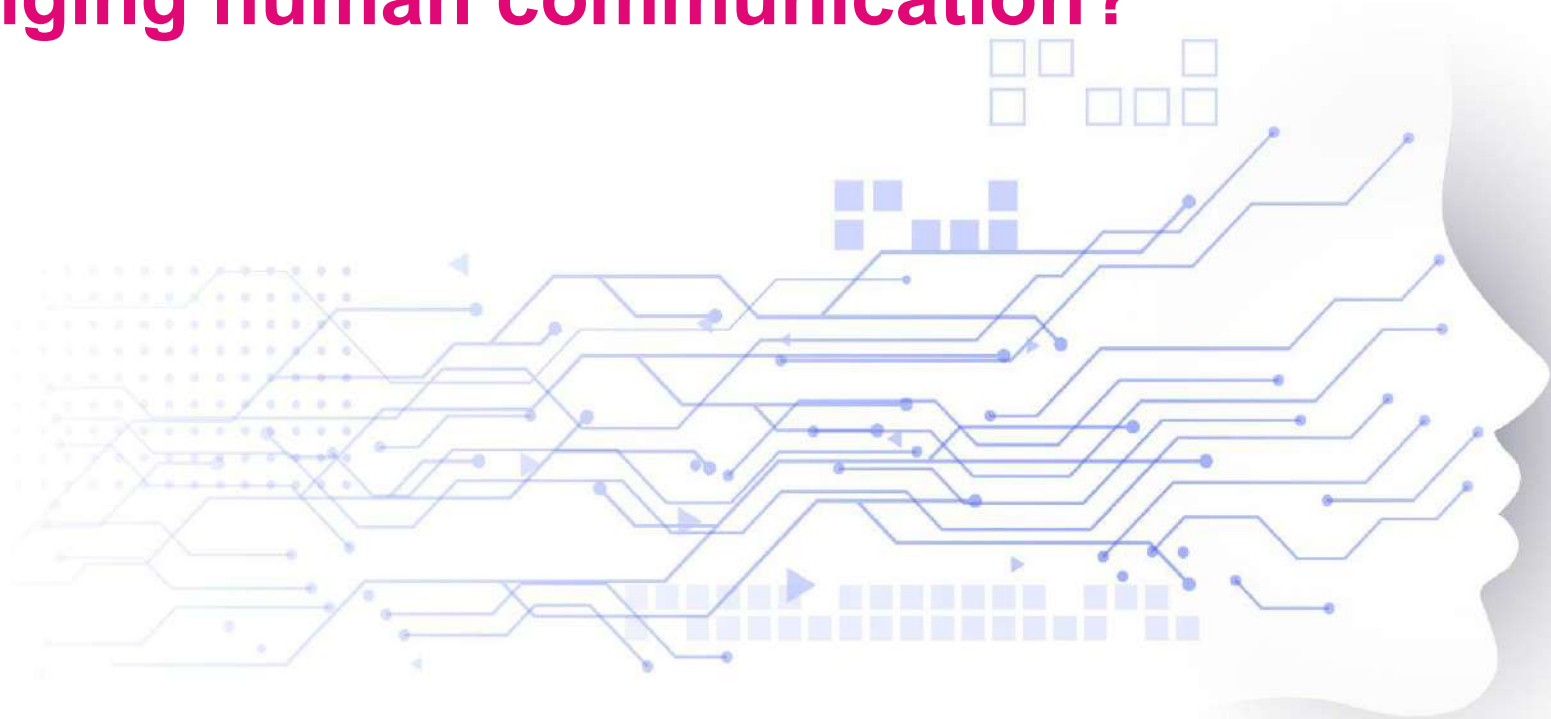
Examples of AI Biases

A cognitive bias is a systematic pattern of deviation from the norm. The most famous example of cognitive bias turned into an algorithmic bias is offered by the system developed by Google for **HR recruiting**.

The system consisted in screening CVs through text analysis, a technique of machine learning that identify the most valuable insights for a particular role.

The result showed on several occasions a preference for male candidates over female candidates, generating a «**gender bias**»

How automation, AI and BIG Data are changing human communication?

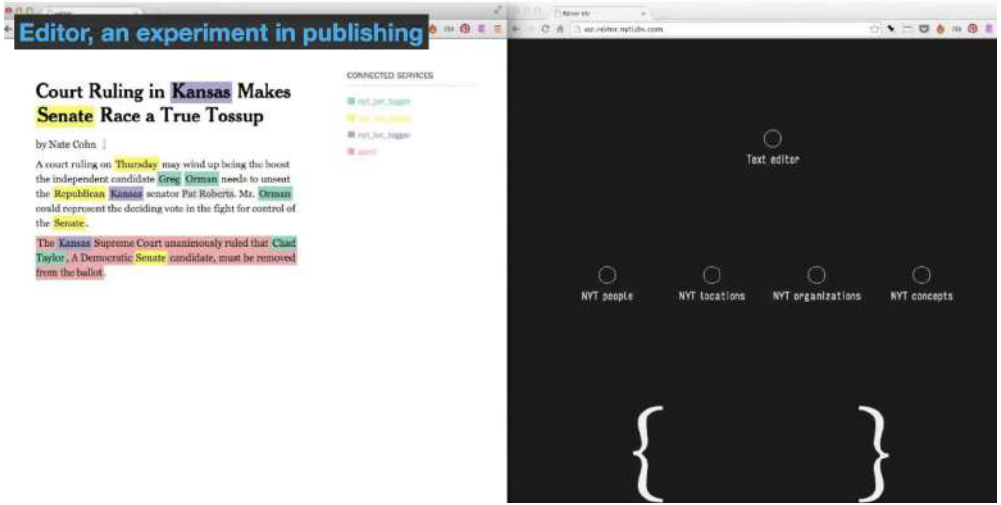


3. Automation is an opportunity

NYT Editor

In 2015, the New York Times implemented its experimental AI project known as **Editor**. The goal of the project was to simplify the journalistic process. When writing an article, a journalist can use **tags to highlight a phrase, a title, or the principal points of the text.**

Over time, the computer learns to recognize these **semantic tags** and earning the most **salient part of an article**. By searching data in real-time and extracting information according to the required categories, the editor will have more accessible information, simplifying the research process and providing quick and accurate fact-checking



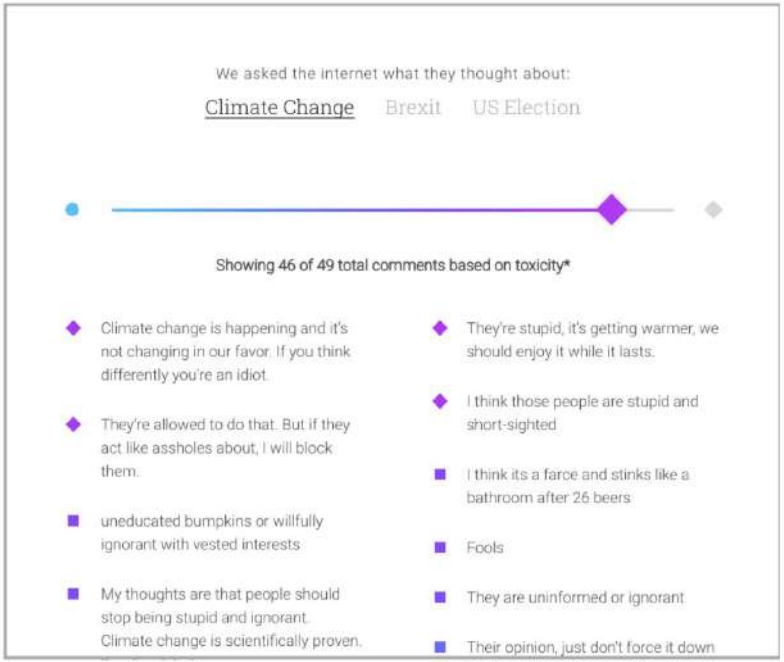
NYT Comments Organizer

Known as a friendly and often thought-provoking forum,**The Time's comments section** is managed by a team of **14 moderators** who manually review more than **11.000 comments each day**.

The **Perspective API tool** developed by Jigsaw (part of Google Alphabet) **organizes reader comments interactively.**

Perspective is a free API that uses machine learning to identify "toxic" comments, making it easier to host better conversations online.

This is a great way for users to read and interact with comments that interest them while avoiding the more aggressive ones.



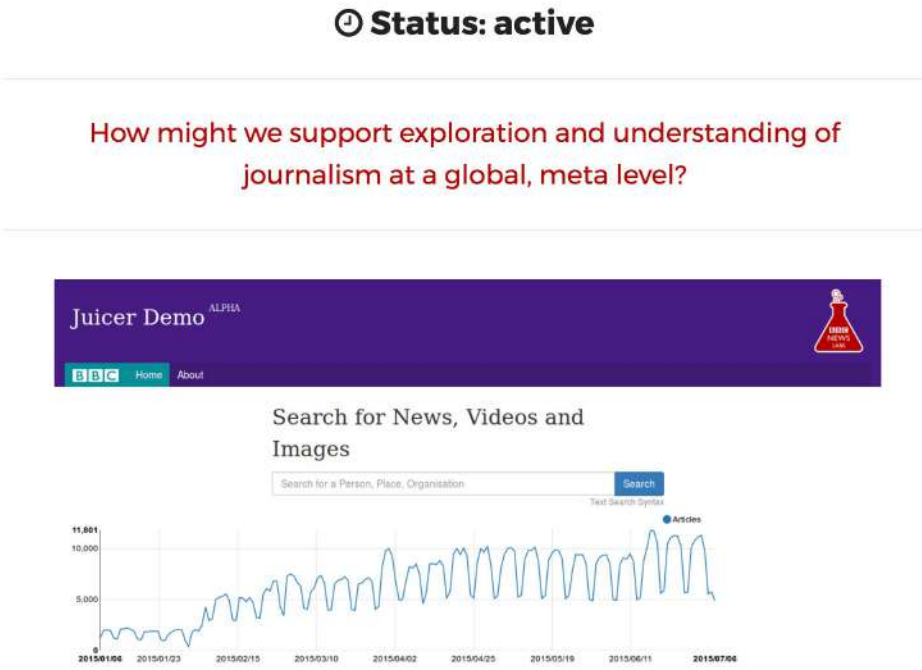
https://perspectiveapi.com/

BBC News Lab - Juicer

The BBC is an archive of a great amount of data, from daily news to feature films, videos, and archives.

Since 2012, the BBC News Lab has been using the **Juicer data-mining tool** to collect contents from BBC and other online sources through RSS feeds. Through NLP techniques is possible to assign semantic tags/categories to the stories and organizes them in one of the following categories: organizations, places, people and concepts (war, love)

Thus, if a journalist searches for the latest news about President Biden or articles related to it with AI systems, **Juicer** by navigating on the Web will quickly provide a list of related contents.



https://bbcnewslabs.co.uk/

Washington Post - Heliograf

For a few years, The WP has been experimenting the automatization writing of news (sometimes defined as «robotic journalism» or «automated journalism») using **Heliograf smart software**.

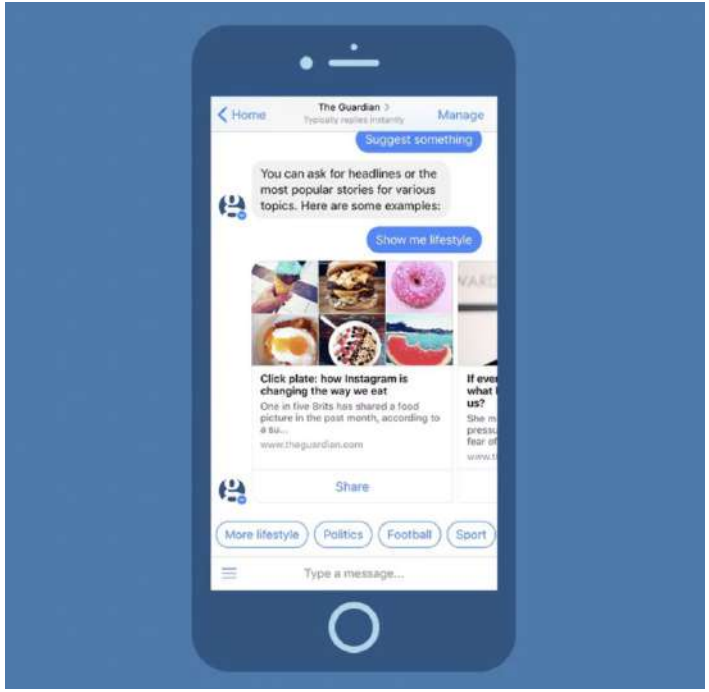
The bot made its debut in the summer 2016 by covering the Rio Olympic Games. Heliograf collected news, analyzing data on the emerged games.

During the Olympics, Heliograf was able to keep up with information about scores and medal counts in real-time, freeing up journalists and let them work on producing more creative contents.

The Guardian - Facebook Chatbot

In 2016, The Guardian launched its chatbot via Facebook. The chatbot allows users to choose between the US, UK and Australian versions of Guardian News and in order to save time in scrolling or searching for news to choose which type of news receiving and at what time of delivery (of 6:00 am – 7:00 am – 8:00 am).

In this way, the user gets, through Facebook, directly the news about the topic of interest (sports, technical – scientific, trend or fashion news, etc.).



Automation is an opportunity

➔ The machine will increasingly be able to replace some of the more **mechanical and repetitive journalistic tasks:** checking sources and spreading news.

➔ All the other **original content creation work** conducted by journalists will be «increased» by the ability to consciously access artificial intelligence and data science.

The newsroom of 2025 sees journalists and machines working together:

- Being Devoted only to quality content
- Saving costs
- Increasing the speed and efficiency of news publishing

First Step: Understanding the Domain

Some principles

The computational origin

In its early days, the computation was a “computational operation” that we might call solipsistic.

Machines operated in contexts of **finite worlds** (the game of chess) with fixed, codifiable and relatively contained rules. They were not supposed to know real-world scenarios.

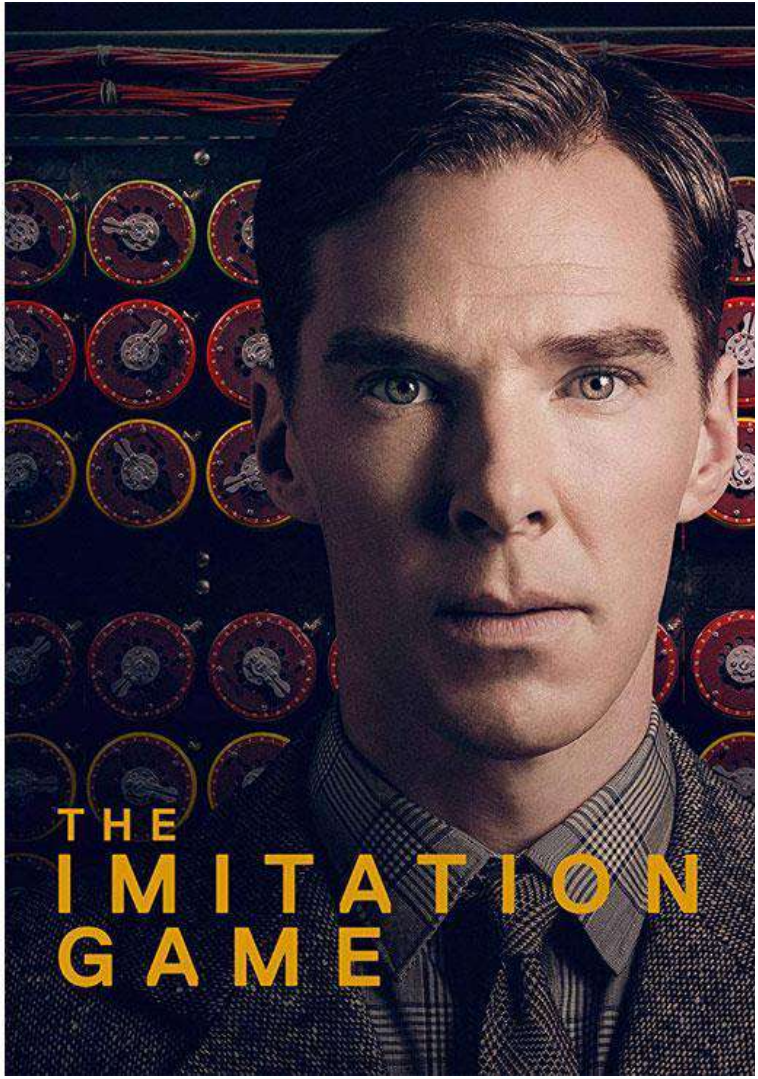
One of the most popular objection is **Ada Lovelace**: "The Analytical Engine has no pretensions to originate anything. It can do whatever we know how to order it to perform (cited by Hartree, p. 70).

English mathematician considered a pioneer in computer science, contributed significantly to the history of computational thinking and the birth of the first computing machines. She developed a particular algorithm referred to as the first ante litteram computer program. (1815-1852)

The Imitation Game

Alan Turing* is considered the father of modern computer science and artificial intelligence.
Turing sought to answer to the question: ***can machines think?***
In his attempt to provide answer, he proposed "**The imitation game**," which later became famous as **the Turing test**.

Turing proposed that a computer can be said to possess artificial intelligence if it can mimic human responses under specific conditions. The original Turing Test requires three terminals, each of which is physically separated from the other two. One terminal is operated by a computer, while the other two are operated by humans.
The test is repeated many times. If the questioner makes the correct determination in half of the test runs or less, the computer is considered to have artificial intelligence because the questioner regards it as "just as human" as the human respondent.



*__Alan__ Mathison __Turing__ (London, 23 June 1912 – Manchester, 7 June 1954) he was a British mathematician, logician, cryptographer and philosopher. Considered one of the fathers of computer science.

Alpha Go

Alpha Go, Deep Mind's supercomputer,
Alpha go competed against legendary Go player Mr Lee Sedol, the winner of 18 world titles, who is widely considered the greatest player of the past decade. AlphaGo's 4-1 victory in Seoul, South Korea, on March 2016 was watched by over 200 million people worldwide. This landmark achievement was a decade ahead of its time.

The surprising element is that the machine won not only thanks to its computing power but also thanks to **its creativity of playing as well as the power of imagination** and, consequently, of i**nnovation**.

Artificial intelligence algorithms can learn from experience and can suggest new products, services and applications.

Machine Learning – I Think

Machines can think and learn. They must be able to learn to be in the world autonomously and to do so, they must be able to produce their own «truth» about the world.

- ➔ **Machines must be able to learn automatically from experiences.**
- ➔ **What does experience is for machines? = data.**
- ➔ **Machine learning refers to the automated detection of meaningful patterns in data (Big Data).**

Big Data
numbers

Data are everywhere. More than **2.7 zettabytes of data** exist in today's digital universe and are expected to grow to **180 zettabytes by 2025**.

To better understand the amount of data that **1 zettabyte t**hat's the equivalent of **18 million times** the digital assets **stored** by the **Library** of **Congress in Washington**
1 zettabyte is approximately equal to **a thousand** Exabytes, a billion Terabytes, or a trillion Gigabytes.

McKinsey Global Institute: "Big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage and analyze."

Gartner: "Big Data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation".

catchy. DEEPINTODATA

Big Data
features

The essential features of Big Data can be summarized in 4V +1:

- **Volume,** the size of the data sets that need to be analyzed and processed, which are now frequently larger than terabytes and petabytes. This means that the data sets in Big Data are too large to process with a regular laptop or desktop processor.
- **Velocity,** refers to the speed with which data is generated. High velocity data is generated with such a pace that it requires distinct (distributed) processing techniques.
- **Variety**, Big Data comes from a great variety of sources and generally is one out of three types: structured, semi structured and unstructured data.
- **Veracity**, refers to the quality of the data that is being analyzed. High veracity data has many records that are valuable to analyze and that contribute in a meaningful way to the overall results.
- **Value**: the ability to turn data into value

In other words, big data refer to huge volume of data that cannot be processed effectively with traditional storage applications but need scalable space. Are used distributed architectures: clusters of computers connected to distribute very complex processing between the various computers, increasing the computing power of the system and/or ensuring greater availability of service

https://www.bigdataframework.org/the-four-vs-of-big-data/

Examples of Machine Learning

What does automated detection mean for meaningful patterns in Data?

For example, classification algorithms =
is to recognize objects and being able to separate them into categories.

- Classifying whether an email is a spam or not.

Machine Learning can be supervised (SL) by giving labels to provide input data or unsupervised: letting the algorithm finds patterns where we do not recognize them.

- Recognizing the kind of music among selected songs (Supervised).
- Predict the next word in a sentence (Unsupervised).

Examples of Deep Learning

When learning and finding patterns becomes complex. In the real world.

Deep Learning selects features of the information in the data by progressive abstractions. In the case of identifying an object in an image, we start from individual pixels and work our way up by degrees of abstraction to identify angels, contours, shapes and finally the object.

- The Startup **PathAI**, diagnostic imaging to detect cancer cells (Computer Vision)
- **The Vocal and Home Assistant** Amazon Alexa, and Google Home to recognize natural language (speech recognition and language processing)

Deep learning means learning automatically from experience thanks to artificial neural networks. It is **a** type **of machine learning** and **artificial** intelligence (AI) that imitates the way humans gain certain types **of** knowledge.

Second step: Use actively your knowledge. Luiss Data Lab, Master of Journalism and Catchy on automated journalism

Automated Journalism - Tools

Following there are the main activities and capabilities in Automated Journalism carried out by the three partner entities of IDMO, the Italian Digital Media Observatory**: Luiss Data Lab,** (Technology & Algorithm provider for experimental engines) and **Catchy** (Data provider and content provider with Luiss Data Lab).

- ***Zeta Luiss Skill (From Luiss Master of Journalism)***
- ***Catchy Summa***
- ***Social Data Intelligence***
- ***Fake news detection: Compass & Bot Detection***

Automated Journalism

Zeta Luiss Skill, Powered by Catchy Big Data

Voice

Data

Amazon

News

ZETA
Luiss Skill

Luiss University's Master of Journalism program, thanks to the technological support of Catchy, has revamped its digital touchpoints:

Zetaluiss.it is the new interactive and responsive website of the Master of Journalism.

Zeta Skill is a new skill for Amazon Alexa that allows users to directly get information about featured news on the Zeta Luiss Web Site.

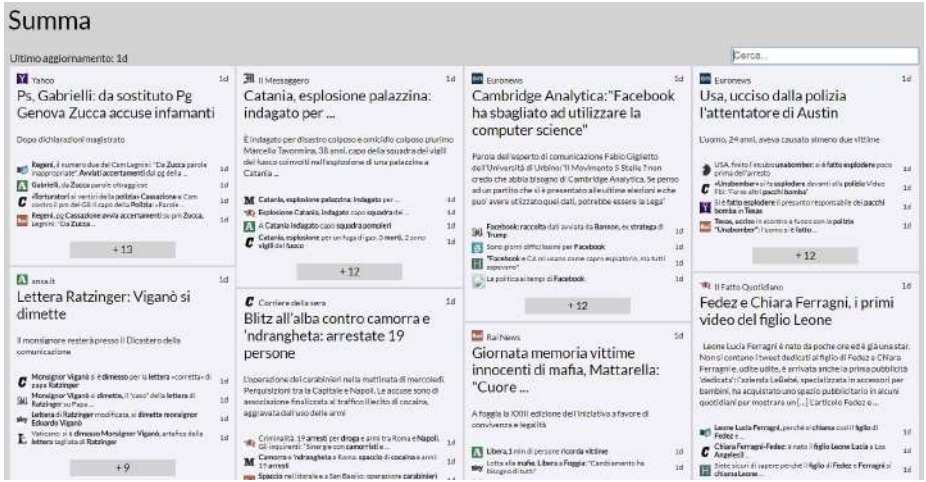
Zeta offers daily breaking news and in-depth coverage of current issues, as well as thematic focuses and interviews with experts and opinion leaders.

catchy

Automated Journalism

Catchy Summa

Catchy SUMMA

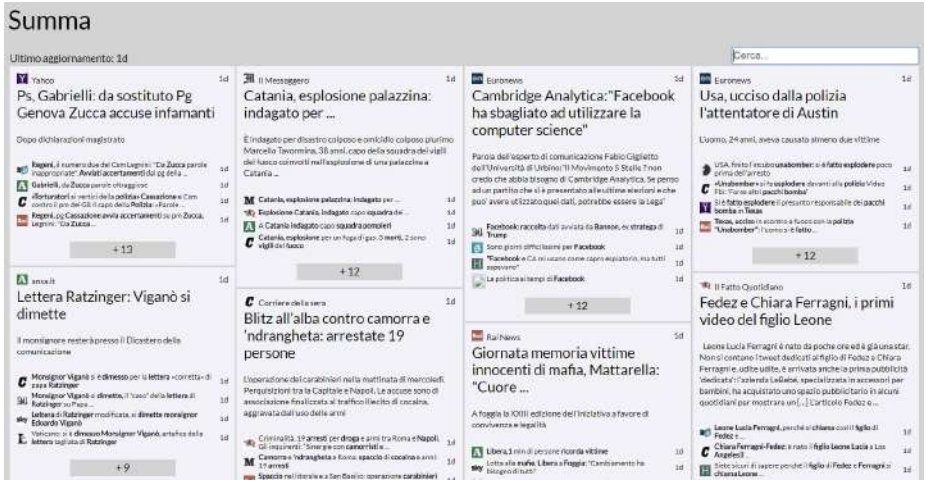


Summa is a **Digital Press Review** with near real time updating for all information disseminated from selected sources. Main features:

1) Collect all digital contents published from a group of selected sources
2) News Gathering related to the same topic from different sources
3) Creation of an automatic news summary.
4) Returning information on the type of information dissemination (number of articles, number of sources)
5) Returning information on the main keywords for each news item

catchy. DEEPINTODATA

Catchy SUMMA



Catchy Summa allows us to:

1) **Keep up to date** with the news published by sources of interest.
2) Know news **popularity trends** (which ones get the most articles? Who talks about them the most?).
3) Know the **latest trends and news** on a particular topic of interest.
4) Know what specialized magazines are **saying about us** using filters.

catchy. DEEPINTODATA

Catchy Summa User Journey

Drafting a Newspaper- Communication Area of a company – Communication agency

Agency

Catchy Summa

The communication agency needs to have real-time digital press review of the latest news

The agency selects a list of reliable national sources and news agencies

The agency logs onto Summa.Carchy.ai and gets all the latest news

The agency knows the most popular news items, the newspapers that published them, the number of articles obtained from each news item, and a brief summary of the news item

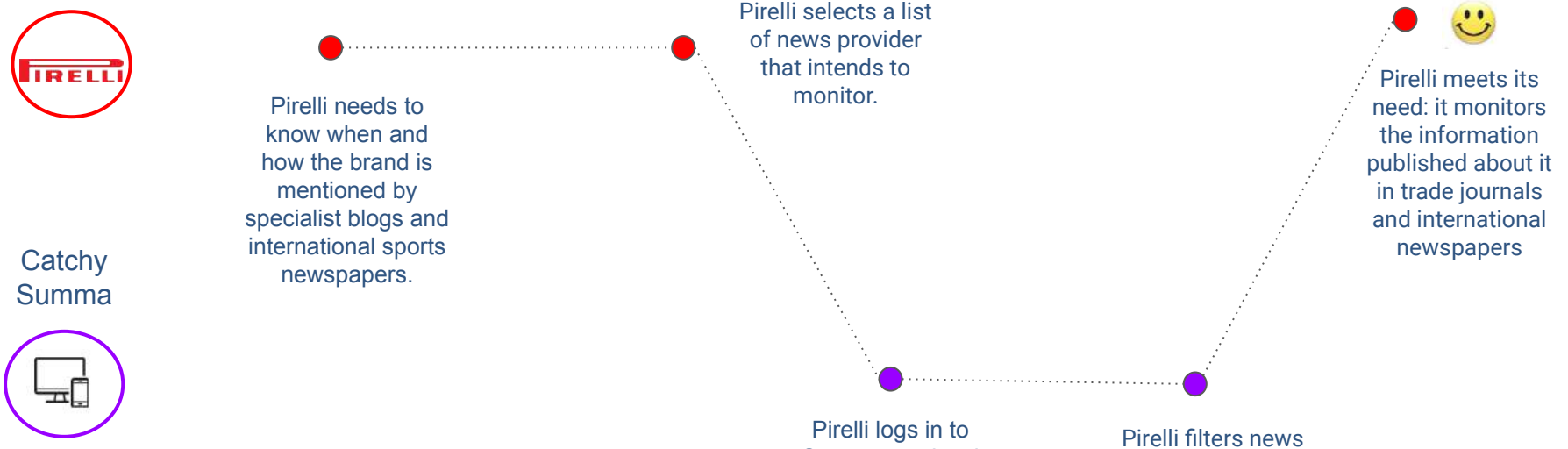
The news agency meets your need. It has a digital press review in near real time

Catchy Summa User Journey

Agency

Pirelli

Catchy Summa



Automated Journalism

Social Data Intelligence

Social Data Intelligence

Starting principles

- ➔ Data Science
- ➔ Data Intelligence
- ➔ Social Data Intelligence

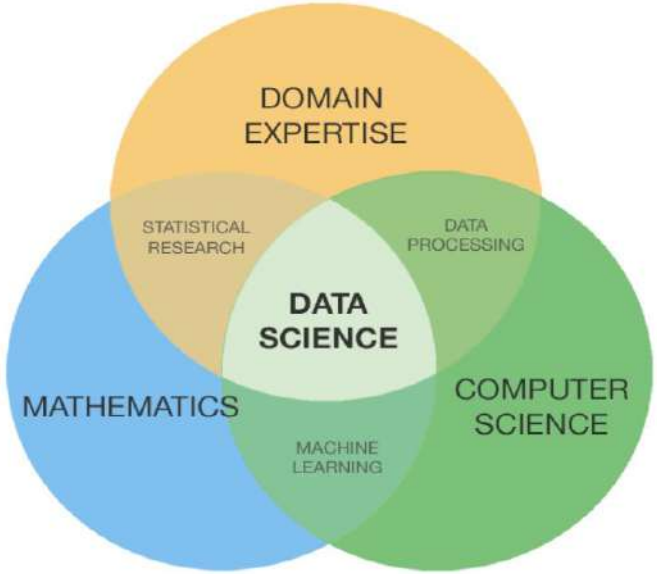
Data Science

From the characteristic “value” of Big Data

Data Science is an interdisciplinary field of study concerned with processes and systems for **extracting knowledge from data in various forms**,and from different sources: web, social networks, IoT devices, databases. Data Science inherits elements of Statistics, Data Mining, Machine Learning, Operations Research, Information Theory, Programming, and Big Data.

Descriptive analysis, the set of tools geared toward describing the current and past situation of business processes and/or functional areas.

Predictive analysis, advanced tools that perform data analysis to answer questions about what might happen in the future.
Prescriptive analytics, advanced tools that, together with data analysis, can propose operational/strategic solutions based on the analysis performed.



Source: Palmer, Shelly. Data Science for the C-Suite. New York: Digital Living Press, 2015. Print.

Data Intelligence

The term **Data Intelligence** refers to the strategic interpretation of the processing and analysis of data collected from heterogeneous sources.

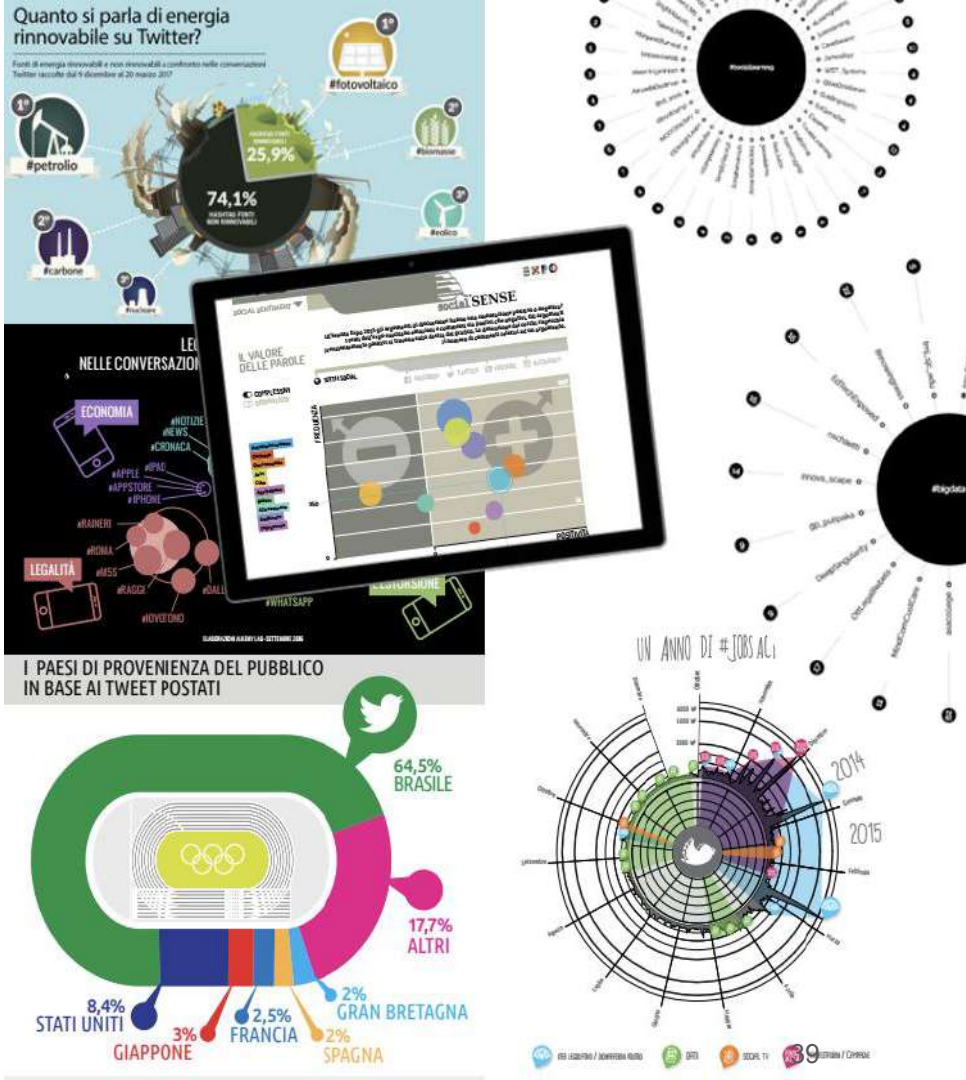
The objective of the analysis is to **render the information contained in data or data processing into a form that can be used** by companies/consumers and end-users in defining their lines of development.

Extracting knowledge and putting it at the service of a business.

Social Media Intelligence

Social Media Intelligence refers to the extraction, cleaning,processing algorithms, analysis and data storytelling from **social networks.** In other words: “using social tools and solutions to understand users and serve them better.”

Within the **Network Paradigms,** that make up our real and virtual society, the marketplace is increasingly connected, informed, and intelligent. **The challenge for today's Brands, Institutions, and Corporations** is to know how to move in these markets through a **clear and constant** presence in the exchanges of **digital information flows, where users inform themselves and guide their choices**. Information flows move mainly on social networks where citizens leave traces of their opinions, feelings

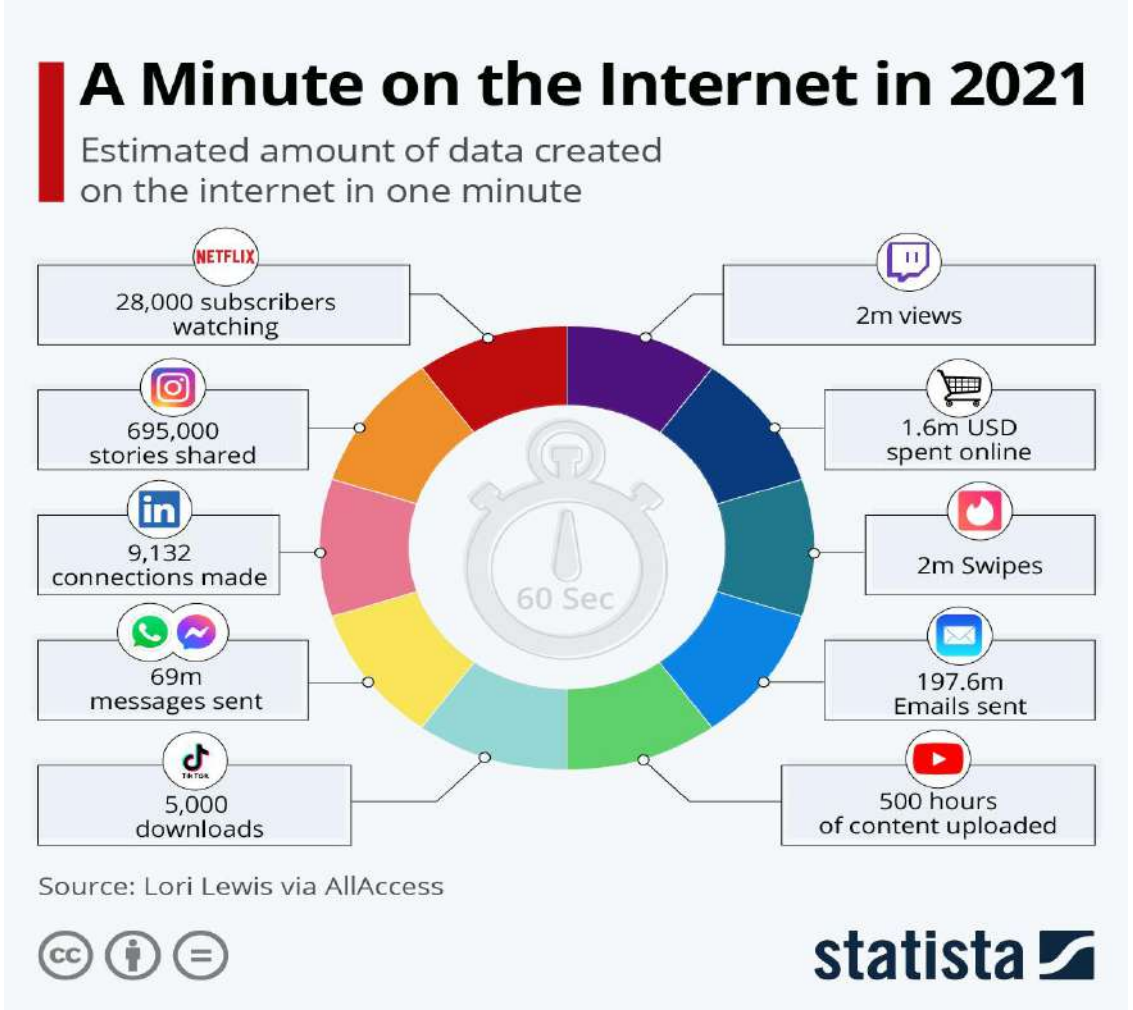


catchy. DEEPINTODATA

Social Data

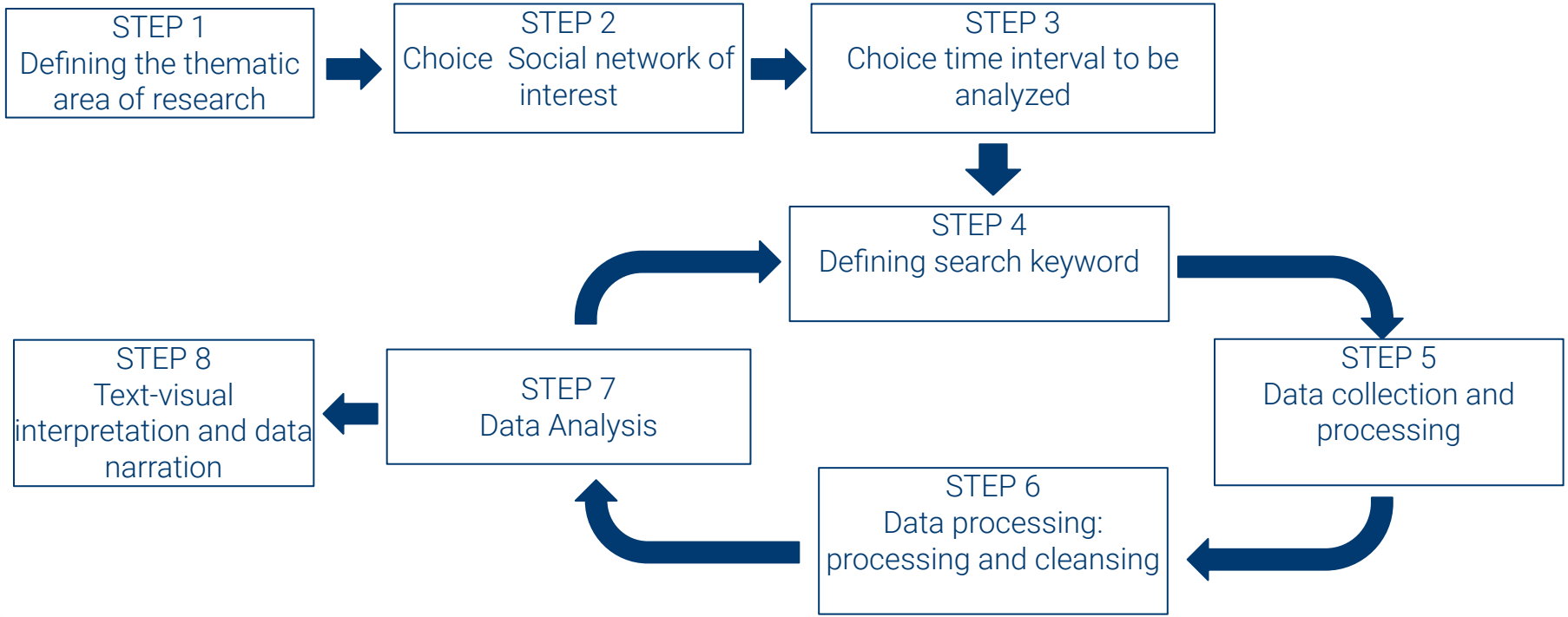
Billions of data are generated on the web each minute.

- 695.000 stories on Instagram per minute
- 2 millions of views on Twitch
- 500 hours of content uploaded on YouTube
- 69 millions of messages on Whatsapp



https://www.statista.com/chart/25443/estimated-amount-of-data-created-on-the-internet-in-one-minute/

Methodology



Results

ANALYSIS OF ACTIVITY, POPULARITY AND CENTRALITY

- Who posts the most? How much does he post?
- Who is the most popular? How much?
- Who is more communicative?
- Who is central in conversations?

SEMANTIC ANALYSIS

- What do they talk about? How much do they talk about it?
- What are the most used words?
- What are the most discussed themes?

SENTIMENT ANALYSIS

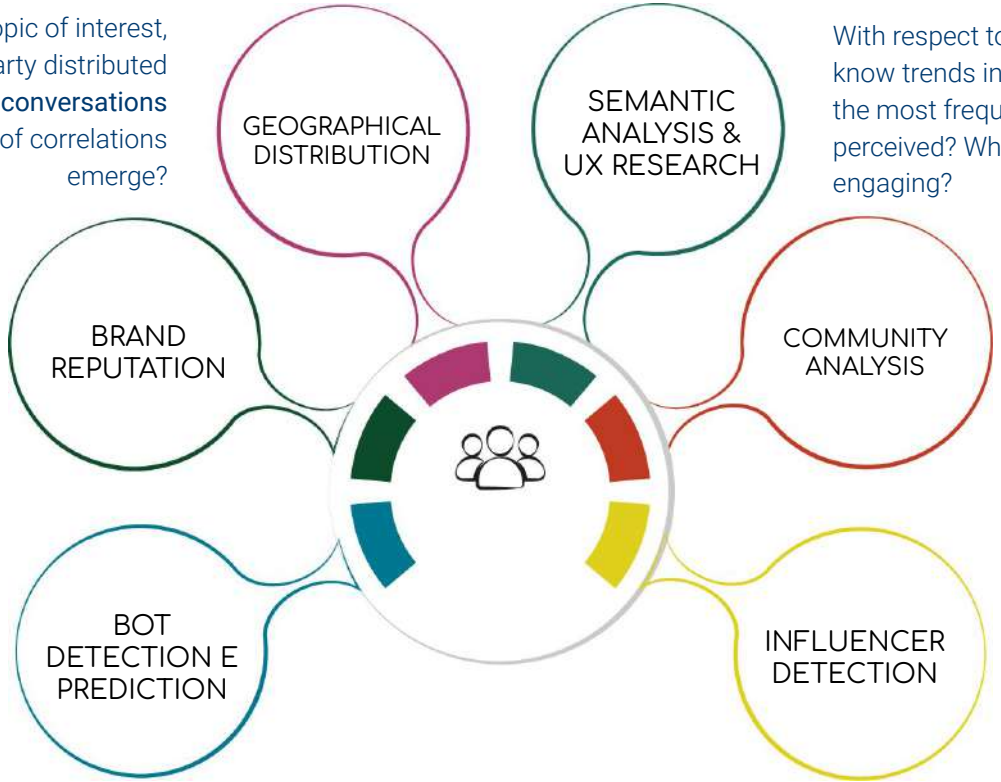
- How do people talk about the event?
- Do they talk about it in a positive, negativo or neutral way?

Social Media Intelligence: goals

How are conversations about a topic of interest, a brand, a political party distributed geographically? **Where are the conversations concentrated** and what kind of correlations emerge?

How is the **brand perceived externally**? Who is talking about it? How does the brand communicate and how can it reach a specific target audience? How does it move relative to competitors?

Interesting **predictive and analytical insights can be offered with respect to socio-political** issues: predicting the outcome of political elections, identifying BOTs that influence the spread of opinions.

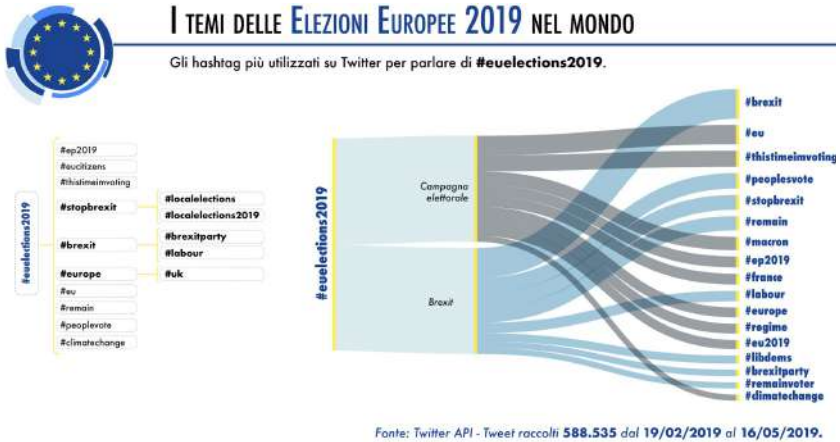


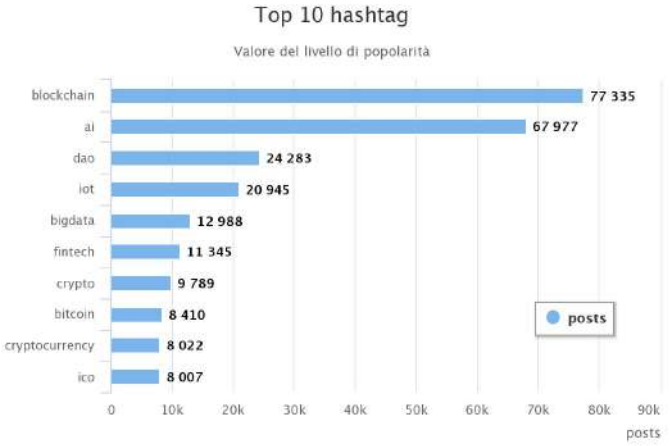
With respect to a given topic, it is possible to know trends in activity and popularity: what are the most frequently **used words**? How is a topic perceived? Which topics are the most engaging?

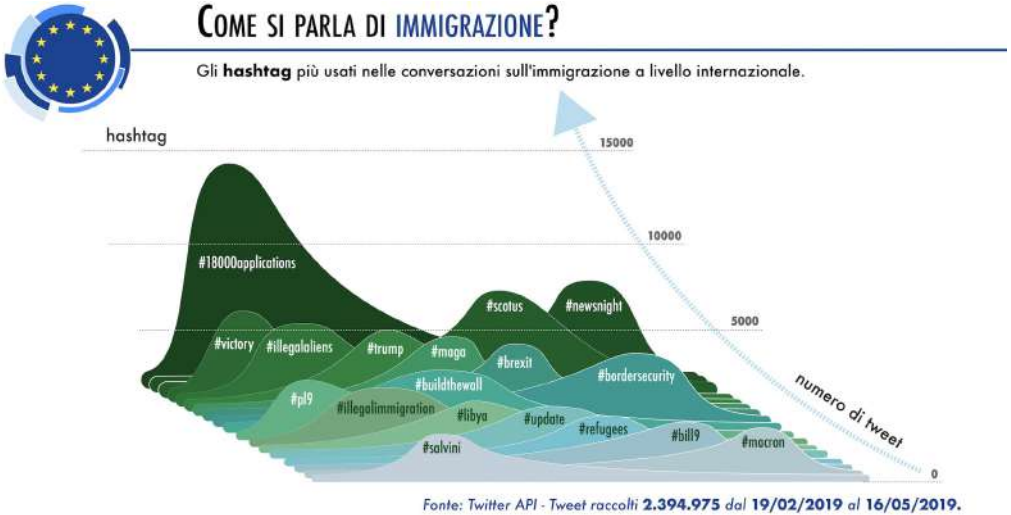
How do communities compose themselves with respect to a specific issue? What are the central nodes? Which are the peripheral nodes?

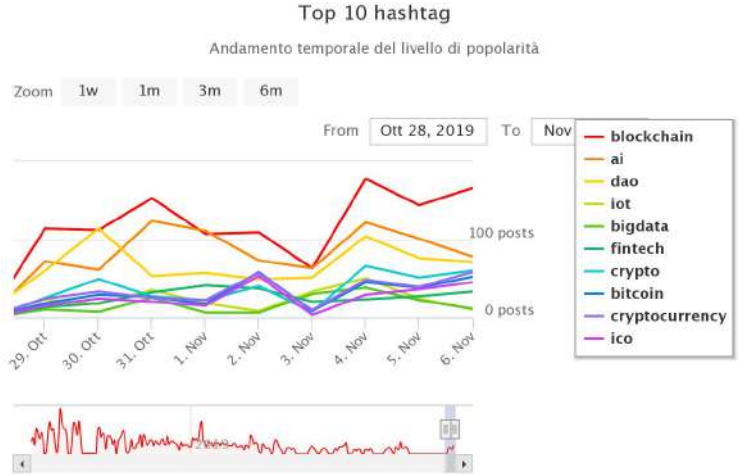
Who are the **influencers** with respect to an issue or brand? How can they get involved?

SEMANTIC ANALYSIS & UX RESEARCH





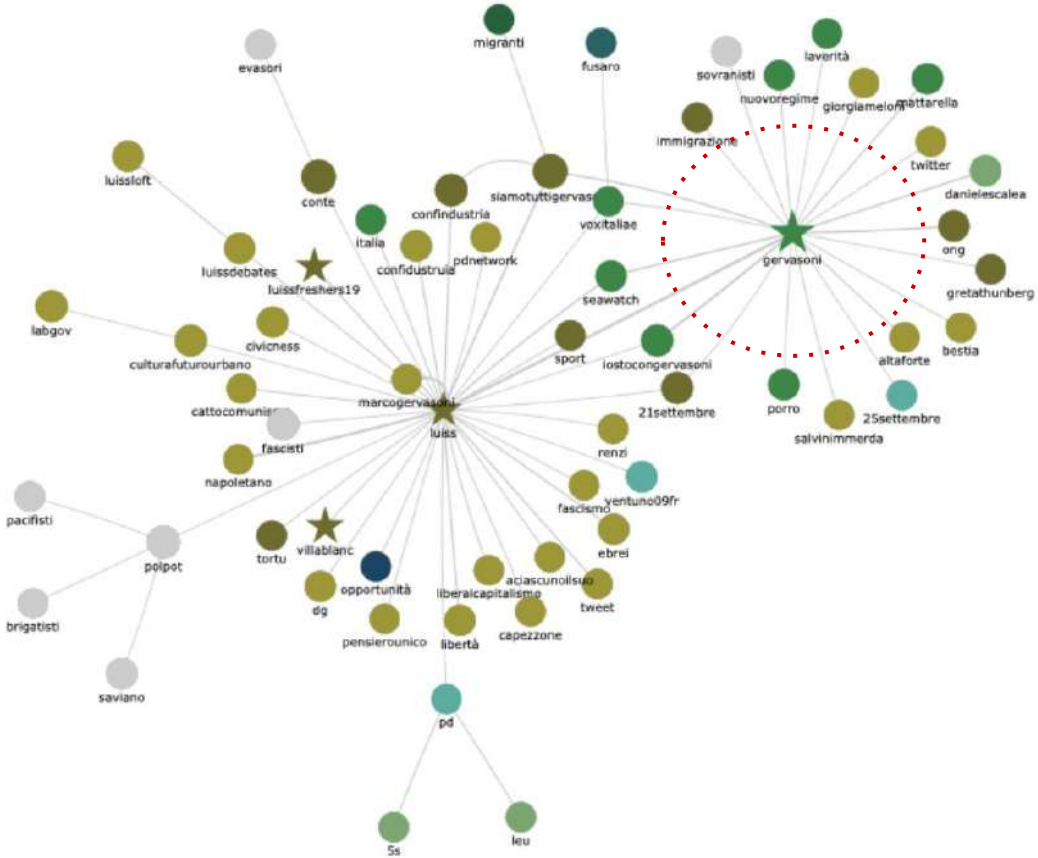




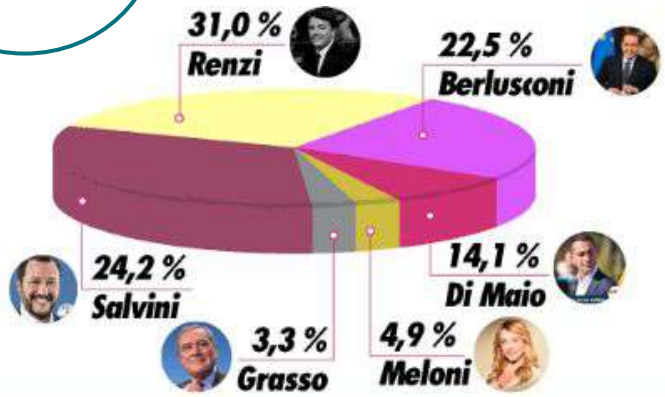
COMMUNITY' S ANALYSIS

Gucci

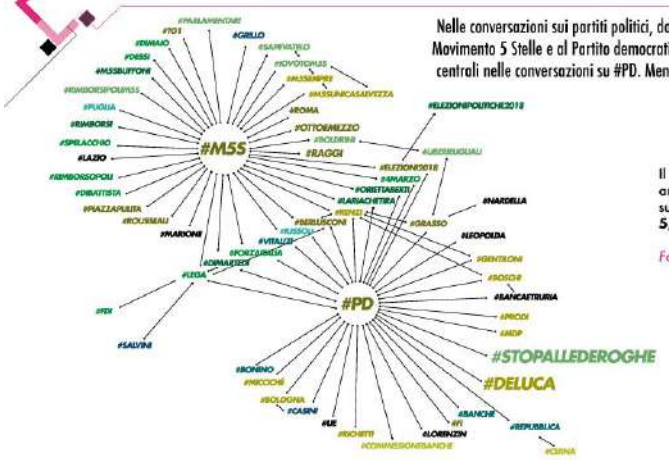
Luiss



ELECTORAL PREDICTION



LA RETE DELLE CONVERSAZIONI TWITTER SUI **PARTITI** POLITICI



Nelle conversazioni sui partiti politici, dominano le community che ruotano intorno al Movimento 5 Stelle e al Partito democratico. #Deluca e #Stopallederoghe sono hashtag centrali nelle conversazioni su #PD. Meno attive le community legate agli altri partiti.

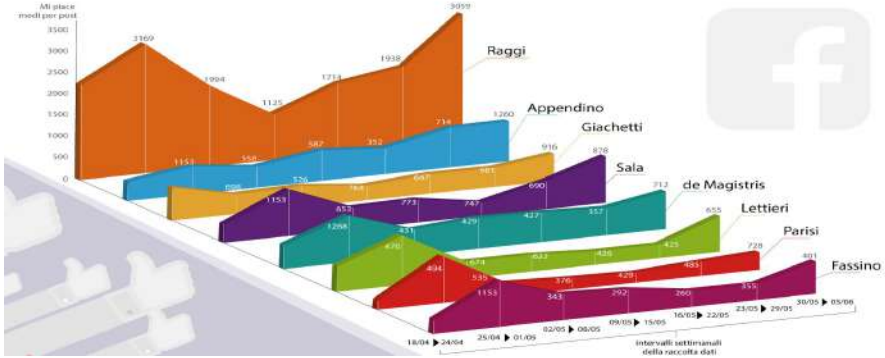
Il grafico mostra la complex network analysis delle conversazioni Twitter sui partiti politici. Periodo monitorato: **5/12/2017 - 19/02/2018.**

Fonte: Twitter API

Elezioni Politiche by

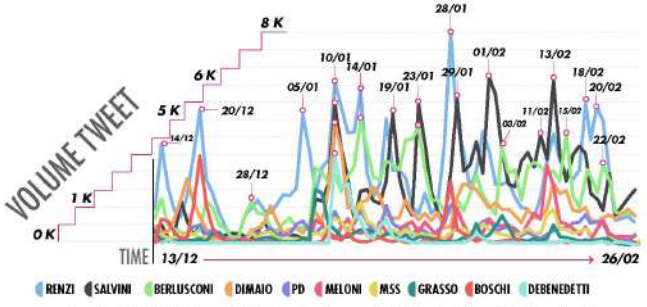
Come è cambiata nel corso delle settimane la popolarità su Facebook dei candidati al ballottaggio delle elezioni amministrative 2016?

L'attrattiva social degli sfidanti attraverso il numero medio di Mi piace guadagnato dai post pubblicati sulle pagine ufficiali tra il 18 aprile e il 5 giugno



LE CONVERSAZIONI TWITTER SUI **LEADER** NEL TEMPO

La popolarità dei leader su Twitter è caratterizzata da un andamento non lineare nel tempo, raggiungendo volumi molto alti ma privi di coda lunga in relazione ad eventi di cronaca - come i fatti di Macerata - o ad ospitate televisive.

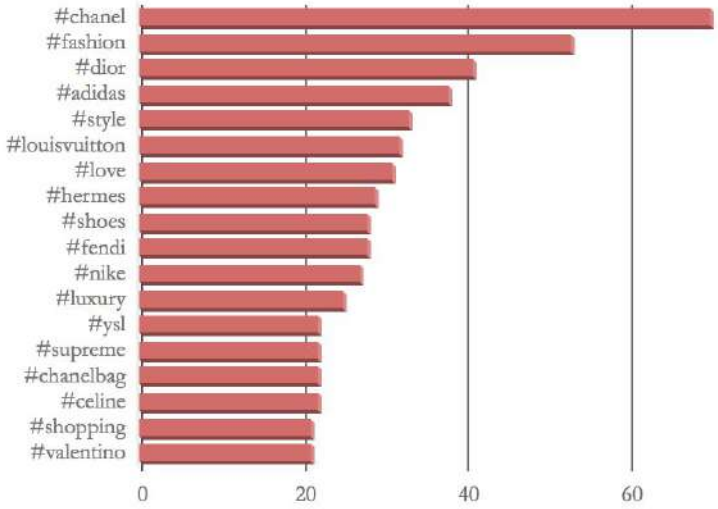


Il grafico highchart evidenzia l'andamento temporale dei dieci hashtag più utilizzati nelle conversazioni Twitter sui leader politici. Periodo monitorato: **13/12/2017 - 26/2/2018.** *Fonte: Twitter API*

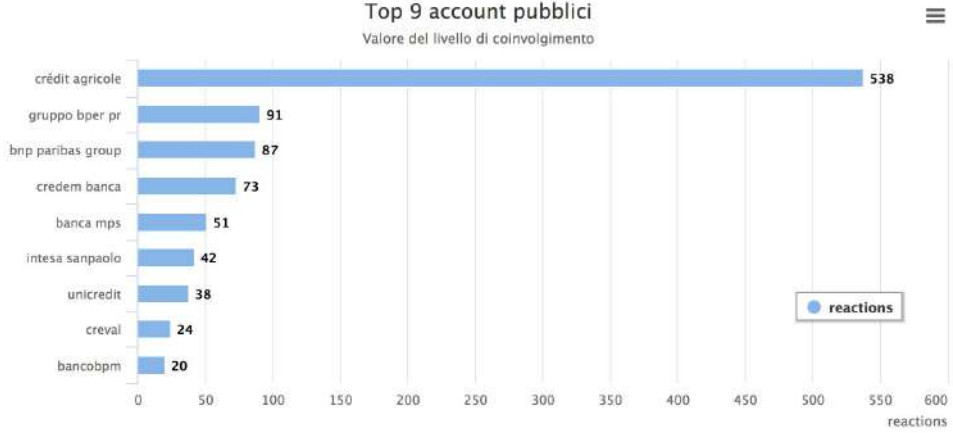
Elezioni Politiche by

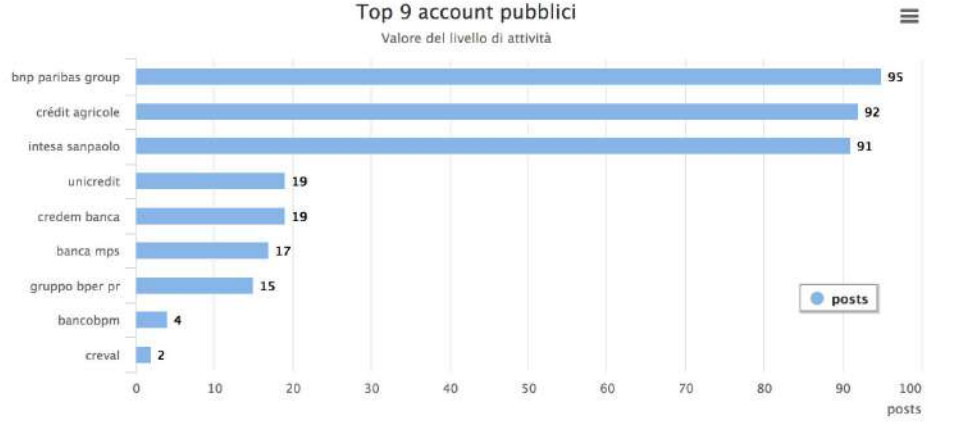
BRAND REPUTATION

Gucci

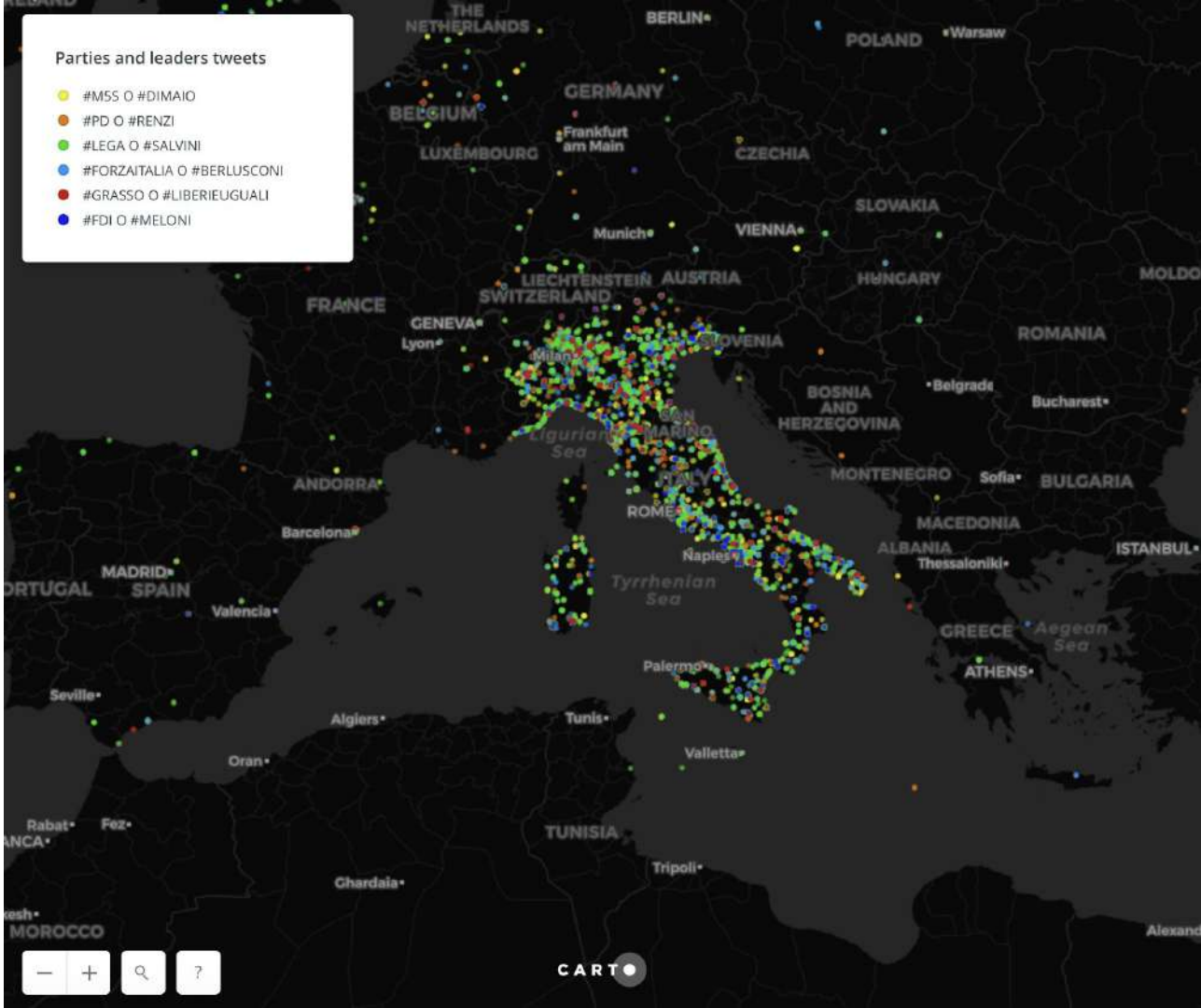


Bank



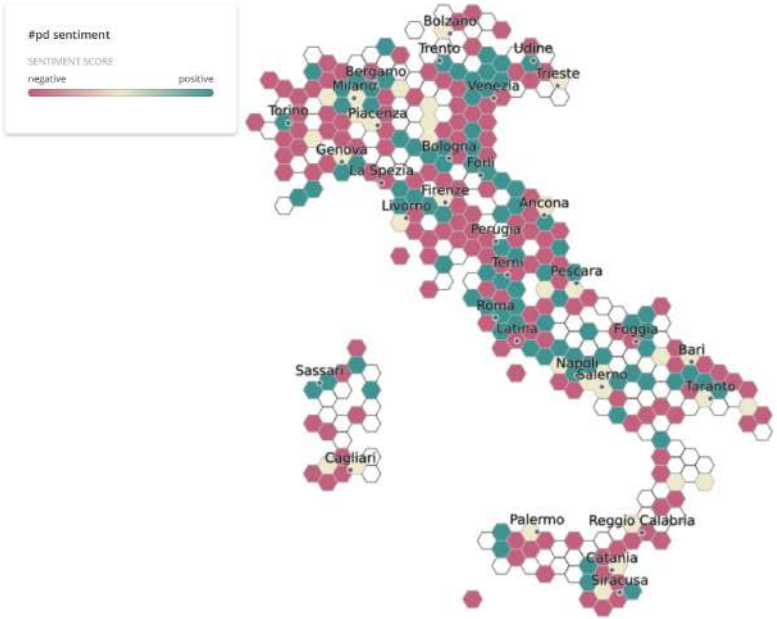


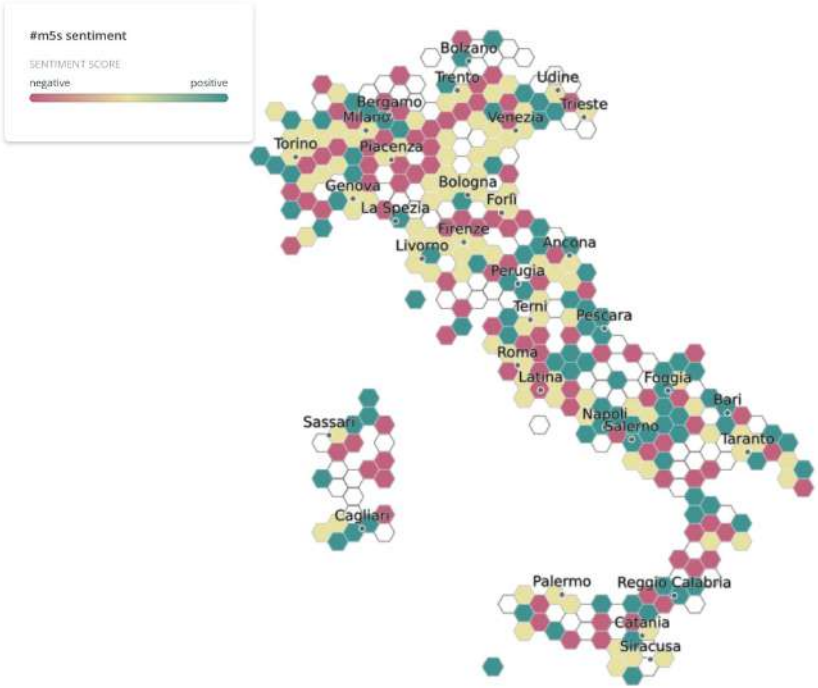
GEOGRAPHICAL DISTRIBUTION



catchy. DEEPINTODATA

GEOGRAPHICAL DISTRIBUTION SENTIMENT

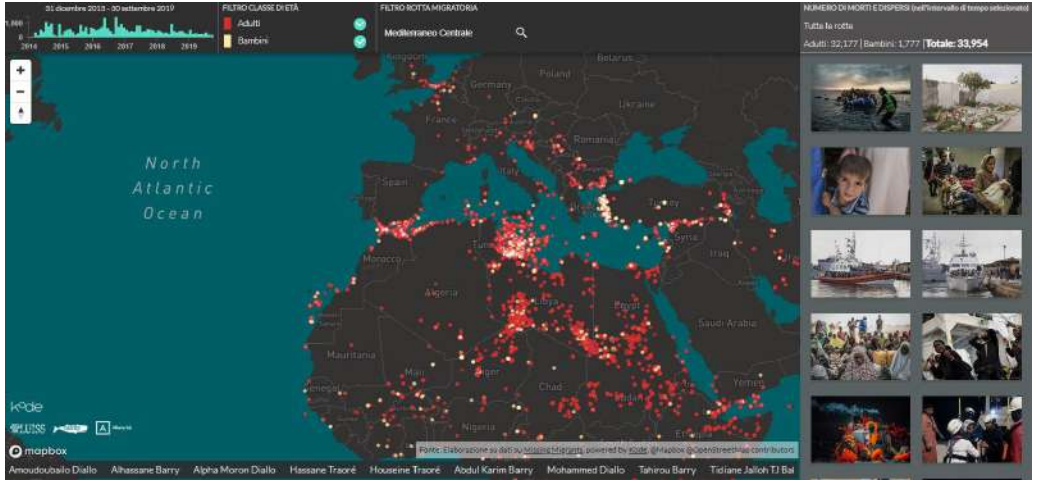


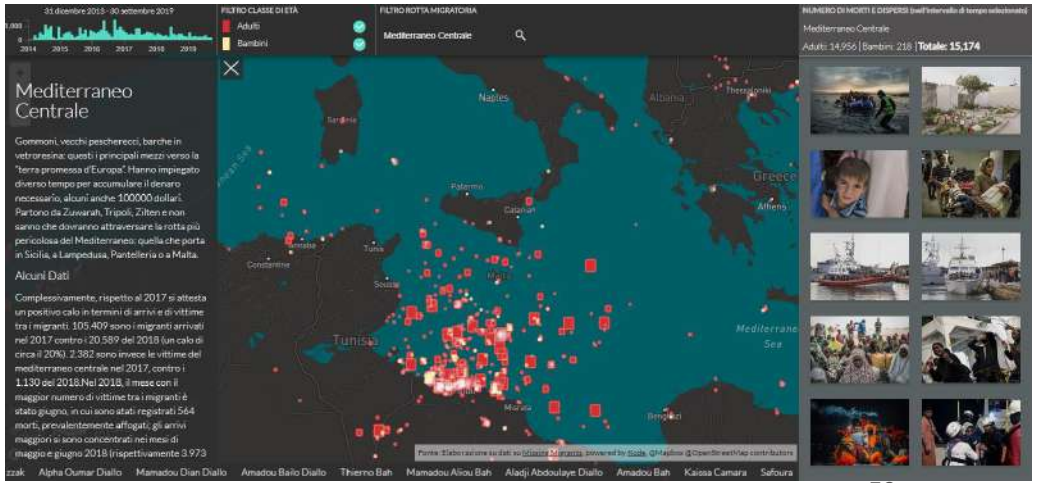


Geo Storytelling

Data and current events: how to tell reality in an objective and engaging way.

The Map of Migrants is a **Data & Geo Storytelling** project developed in collaboration with Catchy, Luiss Data Lab and Kode Srl, **intended to tell in visual and spatial form data, stories and faces of migrants involved in the tragedies of Mediterranean Sea, Ukraine war,** in recent years. To carry out the project, **open data from the Missing Migrants** portal on deaths at sea as of December 2013 were collected, analyzed and processed. The data were geographically distributed on a **Geo Intelligence map** that considers the main migration routes in the Mediterranean Sea. The data were accompanied by individual stories of strong journalistic impact, accompanied by emotional photos offered by **Oxfam Italy and Doctors Without Borders.**



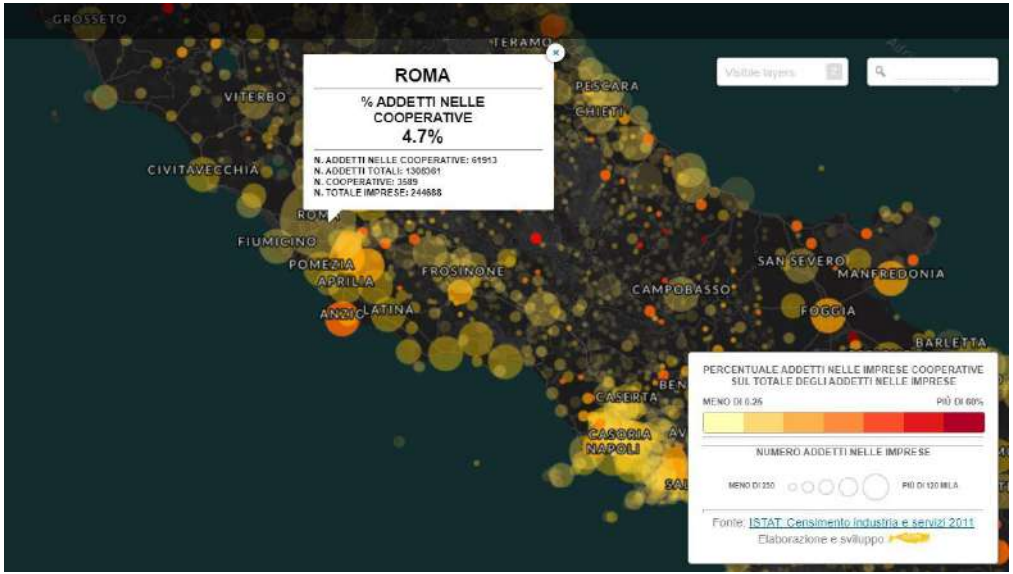


Geo Storytelling

Data and labor: how to tell the story of the distribution of labor on the Italian territory through open data.

Geo Intelligence's map, created in collaboration with startup Catchy, illustrates the spatial distribution of the percentage of employees in cooperative enterprises in Italy, compared to the total number of employees in all enterprises. In municipalities where the color tends toward red, a greater incidence of cooperatives in the labor market dynamics and contextually income generation of an area is observed.

The size of the bubble represents the number of employees in enterprises, the dark shade of color increases as the percentage of employees in cooperative enterprises increases.

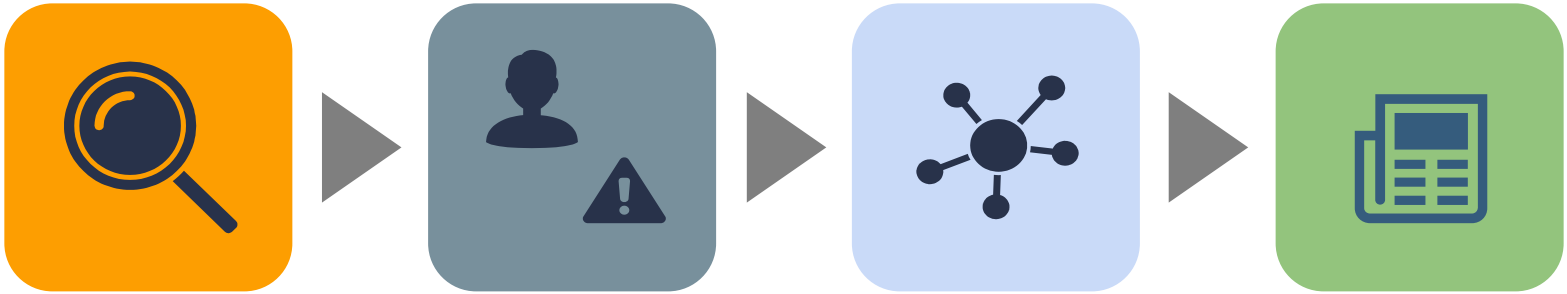


Automated Journalism

Compass & Bot Detection

Social Data Intelligence

Fake News Detection



Recognize and select fake news and analyze ways and themes by which it is spread on social media. Analyze the sources that spread fake news

Collecting and studying users who spread fake news. Detecting bots: automated accounts that create sounding boards especially in the political arena.

Devise network analyses that enable the identification of issues and users central to the spread of fake news.

Build report on the spread and recognition of disinformation as well as to predict and anticipate it in the future.

Compass
Fake News Detection

Compass

Fake News Detection

Compass is a platform that uses Artificial Intelligence for misinformation recognition.
It is based on a specialized engine for collecting digital news from a configurable list of sources (RSS feed, Facebook, Twitter, Youtube, Pinterest) and consolidates the collected data into a dedicated database.

The news contained in the Compass database is subjected to users rating through a **voting mechanism ("true," "false," "don't know") that triggers the evaluation by the algorithm certifying the degree of truthfulness of the news.**

The list of sources can be fed either through a dedicated content curation or through the features of the Compass Plugin. The plugIn allows users who install it **to choose any article they are reading on the web that they do not deem reliable, enter it as a news source in the platform**, vote on it, and allow other experts to express their preference in terms of reliability.

Compass's algorithmic engine is constantly being trained by contributions from communities specializing in the fields of journalism, institutional communication, politics, and corporate.

How do you recognize a social bot?

Social BOT
Markers

- The most obvious indicator that tell us if a particular account is automated is through its activity. The Oxford Internet Institute team suggests keeping an eye on those accounts that post more than 50 tweets per day; at the same time, however**, the Atlantic Council's Forensic Digital Research Laboratory indicates as suspicious a number greater than 72 and very suspicious if greater than 140.**
- **Amplification rate**: one of the main roles of bots is to act as a "sounding board" for specific accounts by retweeting, liking, or quoting related posts. Thus, a typical bot's history will consist of a long series of **retweets and news mentions, with few or no original posts.**
- **Source.** Of interest is the type of source application for tweets posted by accounts: in addition to the classic applications (Twitter for Android, Twitter for IoS or Twitter Web Client etc) bots often use sources that are not traditionally recognized.

Social BOT

Markers

- **Follower-Following**: need to keep in mind that social bots usually have **many followers but tend to follow few people** (since they are auto-tuned to follow and then stop following another account in a specific time frame).
- **Degree of anonymity** shown by the account. In general, **the less personal information** it provides the more likely it is to be a bot. Some bot producers try to disguise their anonymity by using **existing photos**: a good test of an account's veracity is therefore to reverse search the 'image of its avatar (search for the image on Google Image).
- If we instead analyze the ***handle*** (the name of the account preceded by @), we see that many potential bots have **simple alphanumeric sequences probably generated by an algorithm**; or some handles **look like names but do not match the name given by the account.**

Social BOT
Markers

- **Account creation date**: Knowing the exact date when the account was created (just a mouse-over on "joined" can in fact help us understand whether that account was created recently or close to some event (e.g., election campaigns).
- There are bots created “professionally” meaning that they tend to use images, bio of real people, and often links to other social accounts. To avoid spam detection**, they carefully choose their followers and spread them across continents.** Normally, real users do not have followers evenly distributed across continents, so when this situation occurs, it is probably a bot account. (*Followerwonk*).

Social BOT

Use Case

Photo by ev on Unsplash

Yellow Vest (Gilet Jaunes)

Analysis of a movement from its online data

Luiss Data Lab & Catchy with Luiss Master in Journalism.

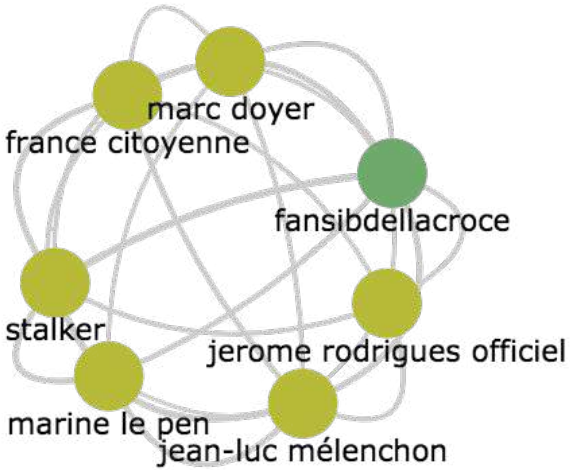
The analysis carried out by the Luiss Data Lab with Catchy processing and engines focused on Twitter data collected between **January and March 2019**, when occurred at least two relevant events for the French protest movement: the counter-manifestation of **"Red Scarfs"** caused a serious **eye injury** of one the high-profile member of the protest movement, Jerome Rodrigues. The same period occurred also the gradual demobilization of participants in the online debate around the **"Yellow Vests (Gilet Jaunes)**, simultaneously to the fall in the term of number of protesters taking part in demonstrations since March.

The research was focused on Twitter activity related to the French-language hashtag **#giletsjaunes** (and similar such as #giletjaune) recorded between **January 21, 2019, and March 12, 2019**, including four weekends of protest. **1,784,271 tweets** have been analyzed, from **144,579 accounts containing 30,356** different **hashtags.**

catchy. DEEPINTODATA

Social BOT

Use Case



Yellow Vest

Analysis of a movement from its online data

Luiss Data Lab & Catchy and Master of Journalism

As is always in the majority of analyses of political flow on social networks,, **suspicious accounts** have emerged in the yellow vests research, sometimes even among **influencers**. Analysis of the online activity flow during the period considered by the research shows a significant presence of such accounts.
In the graph, the account "**fansibdellacroce**" seems to come from another target community, while developing ties with all the key players in the movement. The account mainly makes **retweets** and mentions **influential** accounts.
Automatic analyses with well-known tools such as **botometer** or **botcheck were inconclusive, returning values between 40% and 60% accuracy.** Therefore, a deeper **(and manual) profile** analysis was conducted, **with the help of tools such as Truly Media and Truthnest,** coming to specific behaviors that the survey using algorithms could not define. The research presents an appendix of analysis of some of the most active suspicious accounts in the Twitter community that was created around the hashtag #giletsjaunes during the selected period.

catchy. DEEPINTODATA

Social BOT

Use Case

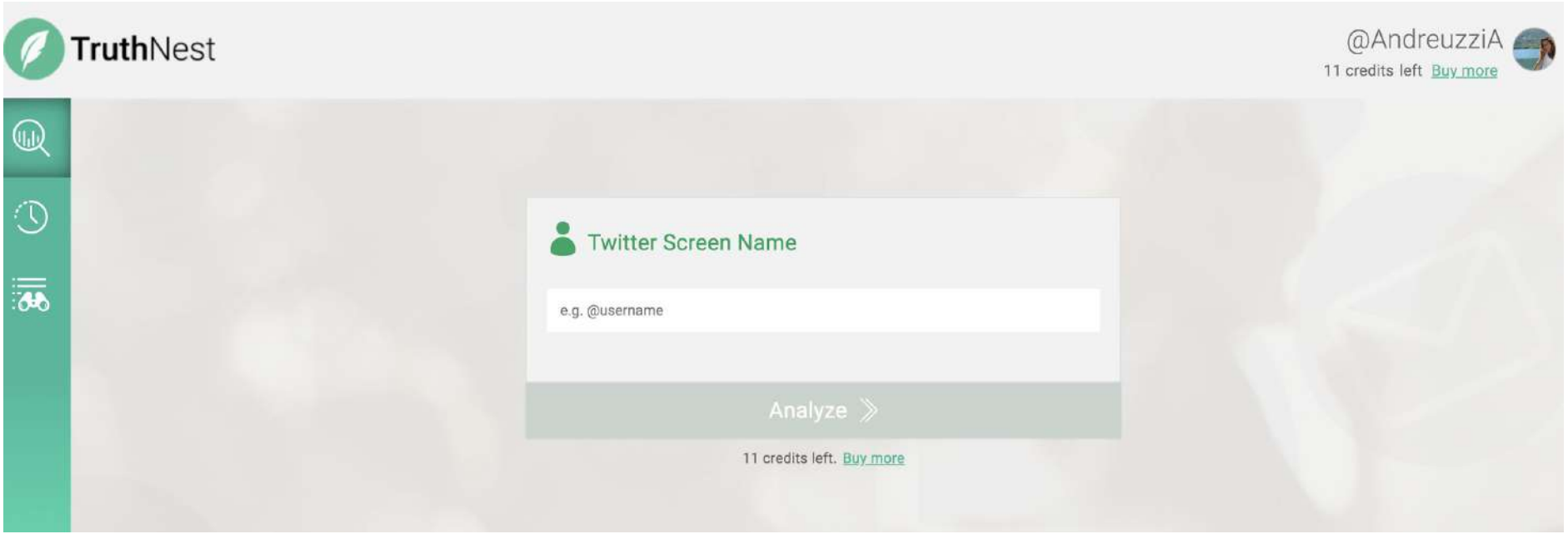
Yellow Vest- Analysis Account BOT

Analysis of a movement from its online data

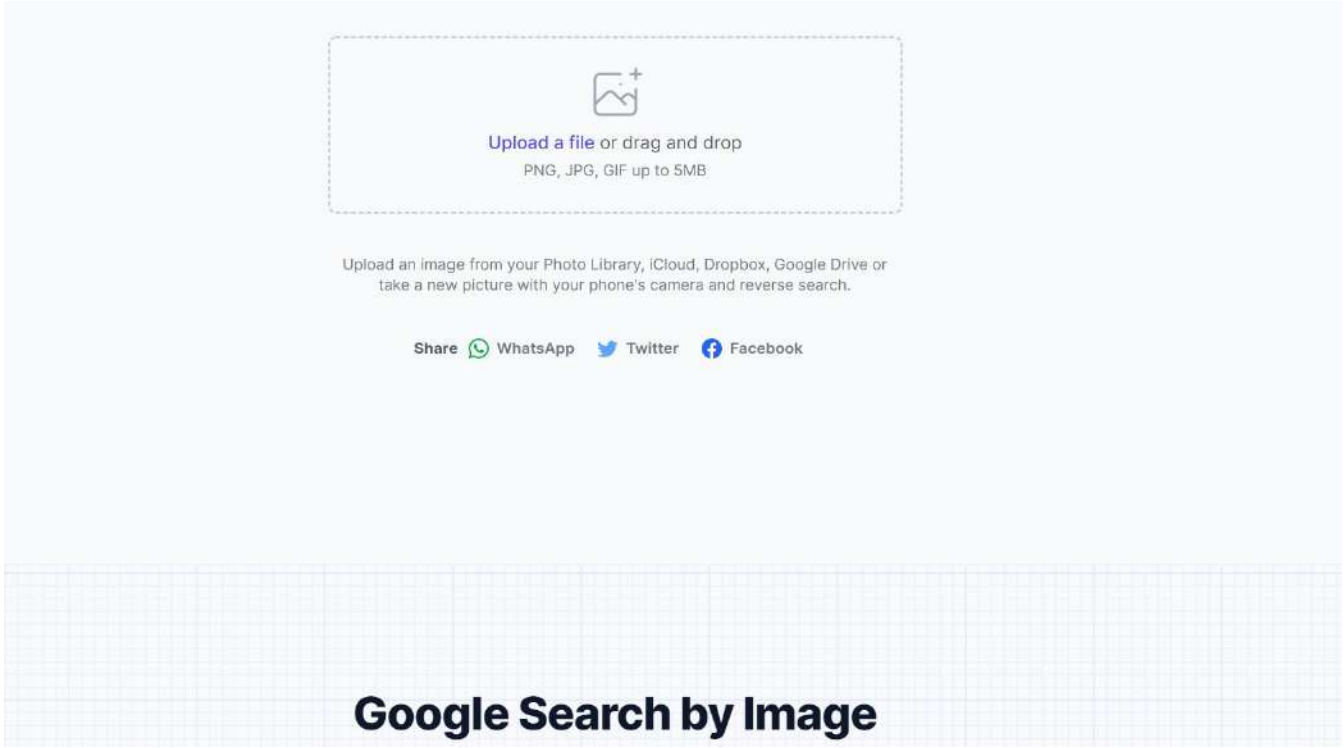
Luiss Data Lab & Catchy and Master of Journalism

- **The photo is by Christian Sterk, a young Dutch photographer, geolocated by the author in Bosnia.** An Open Source tool using the reverse image search engine allows dating it to at least 2016, since it was used in a blog post dated Dec. 14, 2016, published under the name of **Daria Sokolovska**, who described herself at the time as a student of the National University of Kyev, Ukraine, and working at a communications agency in the ukrainian capital.
- The account has been existing since **June 2015**. In the first few months, it seemed to care only about making itself known: **the tweets launch keep track of followers and greet new followers** with automatic content and almost always using English language through the **Crowdfire** app, or others like Unfollowerstat, for Twitter management.
- Then, suddenly, during **August 2016, tweet with stock photos and automated tweets, the @fdbfrancois account stops tweeting original content and just re-tweets more tweets.** The posts thinned out, until March 2018, when the account began to show a lot of activity, which was judged to be at risk of "automation" by analytics software such as Botometer or TruthNest.

Truthnest


TruthNest
@AndreuzziA
11 credits left Buy more
Twitter Screen Name
e.g. @username
Analyze
11 credits left. Buy more

Google Reverse Image Search



https://www.labnol.org/reverse/

LUISS

Data Lab

LUISS

Master in Giornalismo

Artificial Intelligence and Communication

Dr Elena Musi (Tenured Assistant professor)

PI UKRI project “Being Alone together: Developing Fake News immunity”

https://fakenewsimmunity.liverpool.ac.uk/

UNIVERSITY OF LIVERPOOL

UK Research and Innovation

Italian Digital Media Observatory

Outline

- What is Artificial Intelligence?
- What's the interface between AI and Communication?

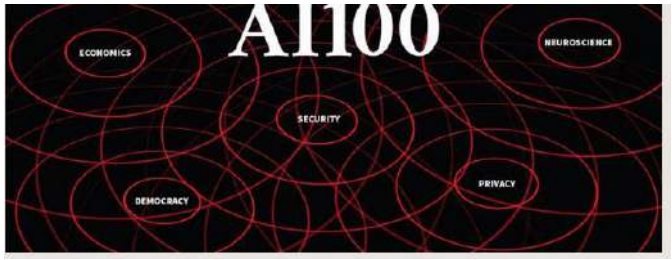
✔Computer Mediated Communication

✔Human Computer Interaction

What is Artificial Intelligence?

“Curiously, the lack of a precise, universally accepted definition of AI probably has helped the field to grow, blossom, and advance at an ever-accelerating pace. Practitioners, researchers, and developers of AI are instead guided by a rough sense of direction and an imperative to “get on with it.” ”

(One Hundred Year Study on Artificial Intelligence, or AI 100 report)



Are these all AI?

Defining Artificial Intelligence

Oxford English Dictionary:

The capacity of computers or other machines to exhibit or simulate intelligent behaviour; the field of study concerned with this.

Scientific field of inquiry

Particular artifact

"Artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment."

Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge, UK: Cambridge University Press, 2010)

Is then an electronic calculator AI?

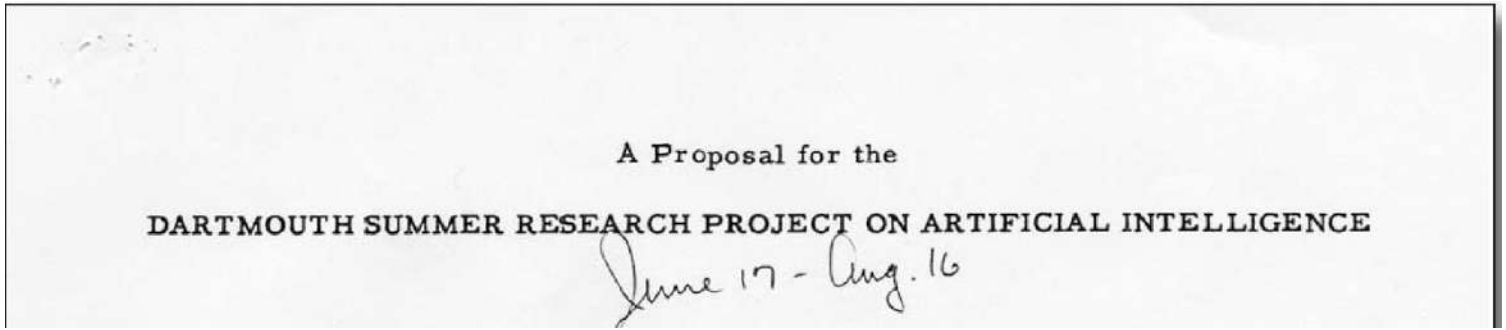
Defining Intelligence

Oxford English Dictionary:

The faculty of understanding; intellect. Also as a count noun: a mental manifestation of this faculty, a capacity to understand.

“AI people are fond of talking about intelligent machines, but when it comes down to it, there is little agreement on exactly what constitutes intelligence”

Schank, R., 1990. *What is AI anyway?* In: “The foundations of Artificial Intelligence: a sourcebook", Partridge, Derek and Yorick Wilks (eds), Cambridge University Press


A Proposal for the

DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

June 17 - Aug. 16

An attempt will be made to find out how machines

- use language
- form abstractions and concepts
- solve kind of problems now reserved for humans
- improve themselves

John Mccarthy: “during the next year and during the Summer Research Project on Artificial Intelligence I propose to study the relation of language to intelligence [...] the human mind apparently uses language as its means of handling complicated phenomena” ....

Defining “Artificial”

Oxford English Dictionary:

A. adj. **I.** Opposed to *natural*.

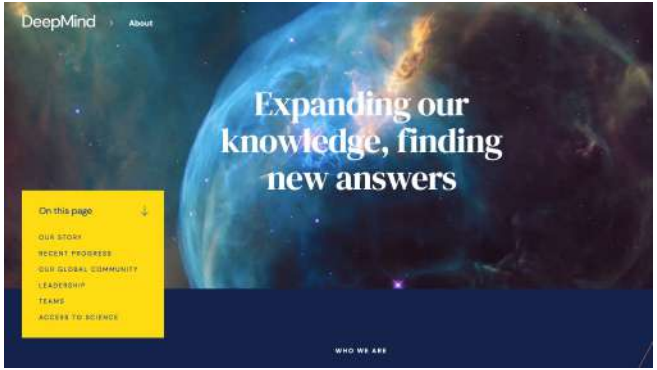
1. Of a thing: made or constructed by human skill, esp. in imitation of, or as a substitute for, something which is made or occurs naturally; man-made.

Artificial Intelligence is an imitation, a substitute for human Intelligence

Artificial Intelligence is intelligence produced by humans

General Artificial Intelligence

- Goal: replace human skills in an enhanced way
- Scope: breadth of the human intellect



Narrow Artificial Intelligence

- Goal: augment human skills
- Scope: specific tasks

The AI effect

AI as an evolving definition:

AI brings a new technology into the common fold people become accustomed to this technology it stops being considered AI newer technology emerges

- 1997: *IBM Deep Blue* defeated Gary Kasparov in chess
- 2011: *IBM Watson* defeated Ken Jennings at the US Quiz show *Jeopardy!*
- 2018: *IBM Project Debater* wins over Noa Ovadia, national debating champion

Project Debater – Speech by Crowd (**https://tinyurl.com/y5ff8xk9**)

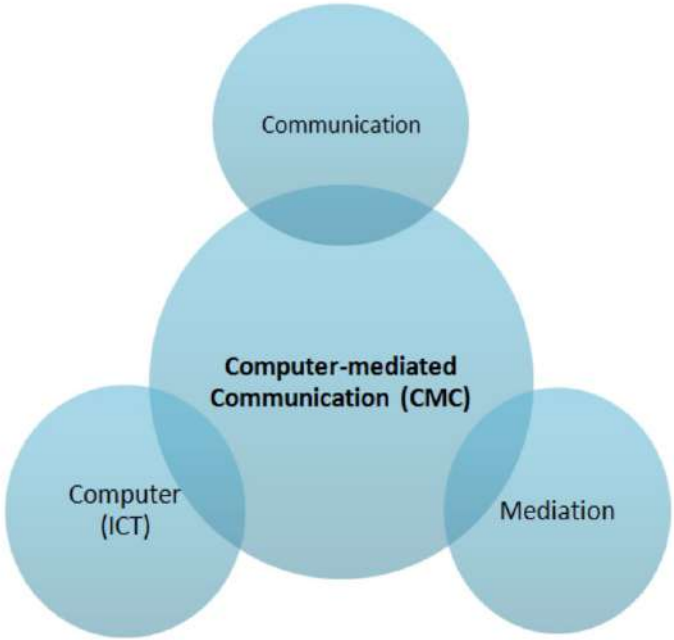
Computer Mediated Communication is an umbrella term which refers to **human communication via computers**

- TIME: synchronous vs. asynchronous

e.g. Skype, Video Conferences, tele
conversation, email, texting, faceboc
Reddit...

- SPACE: from (LANs) to the Intern

1 1 1 many
many many



Human-computer interaction (HCI) is a multidisciplinary field of study focusing on the design of computer technology and, in particular, the interaction between humans (the users) and computers. While initially concerned with computers, HCI has since expanded to cover almost all forms of information technology design.

Running apps

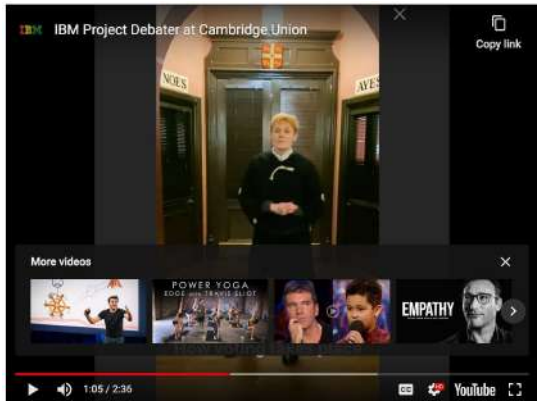
Nintendo

Augmented reality

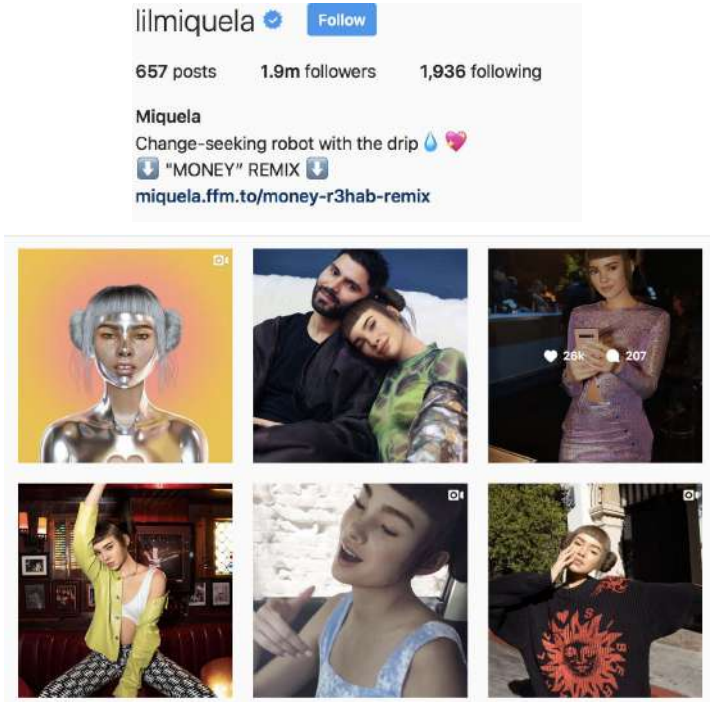
Siri

CMC and HCI: blurred borders

Augmenting Humans: IBM's Project Debater AI Helps Human Debaters Win



https://youtu.be/VRKwn1lkoz8



https://medium.com/@IBMResearch/augmenting-humans-ibm-s-project-debater-ai-gives-human-debating-teams-a-hand-at-cambridge-69a29bcd4eff

https://www.thecut.com/2018/05/lil-miquela-digital-avatar-instagram-influencer.html

Modes of CMC

"A mode is a genre of CMC that combines messaging protocols and the social and cultural practices that have evolved around their use"

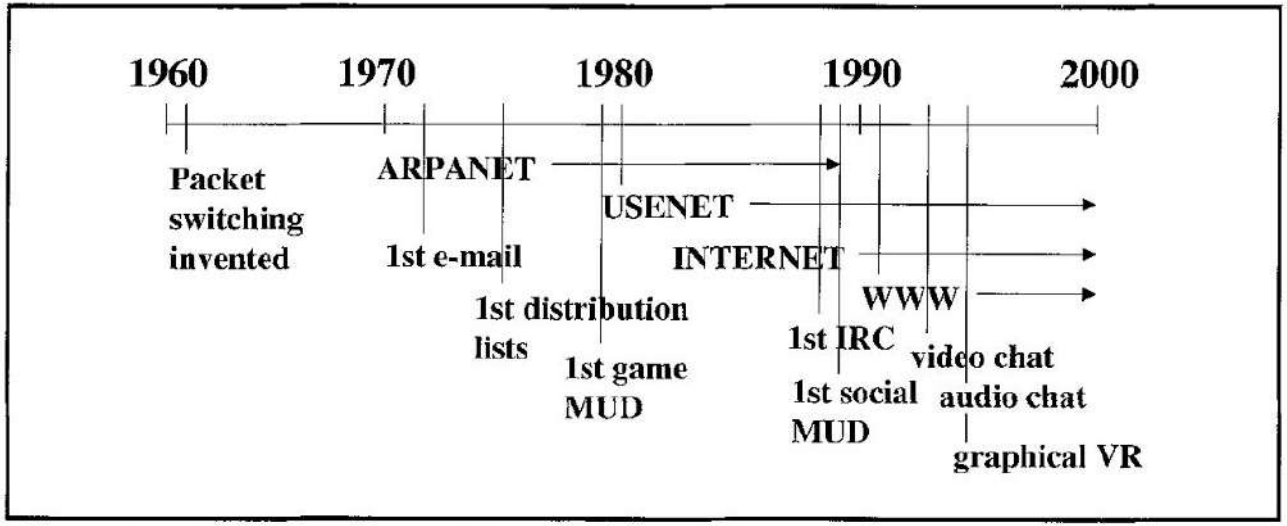


Figure 3.1 The co-evolution of the Internet and CMC

Herring, Susan C. "Computer-mediated communication on the Internet." *Annual review of information science and technology* 36, no. 1 (2002): 109-168.

Computer mediated Communication

Changes in

1. **How we express ourselves …**

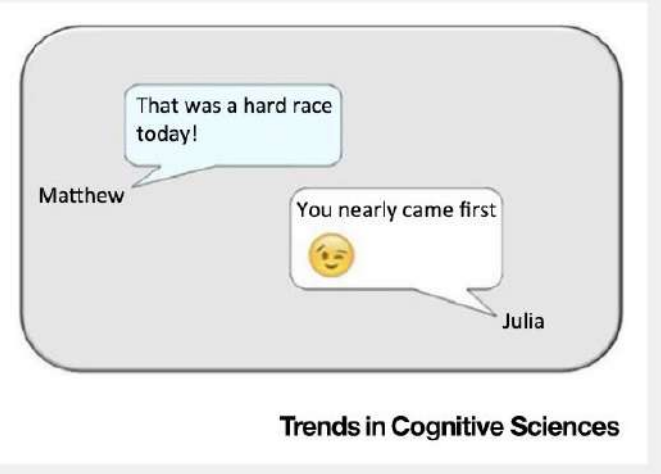
2. **How we relate to each other…**

3. **How we get informed**

!!! Zoom fatigue, hikikomoris

The communicative power of emoticons and emojis … ;)

- Disambiguate
- Express non verbal cues (e.g. turn-taking)
- Expression of users' personalities



The case of Lolspeak …

Special Internet language varieties

LoI: laughing out loud

Fiorentini, I., 2013. Zomg! Dis Iz A New Language": The Case Of Lolspeak. *Selected Papers from Sociolinguistics Summer School*, *4*, pp.90-108.

Zoom Fatigue

"Zoom fatigue is a unique kind of exhaustion that occurs wher
teleconferencing calls for an extended time pe

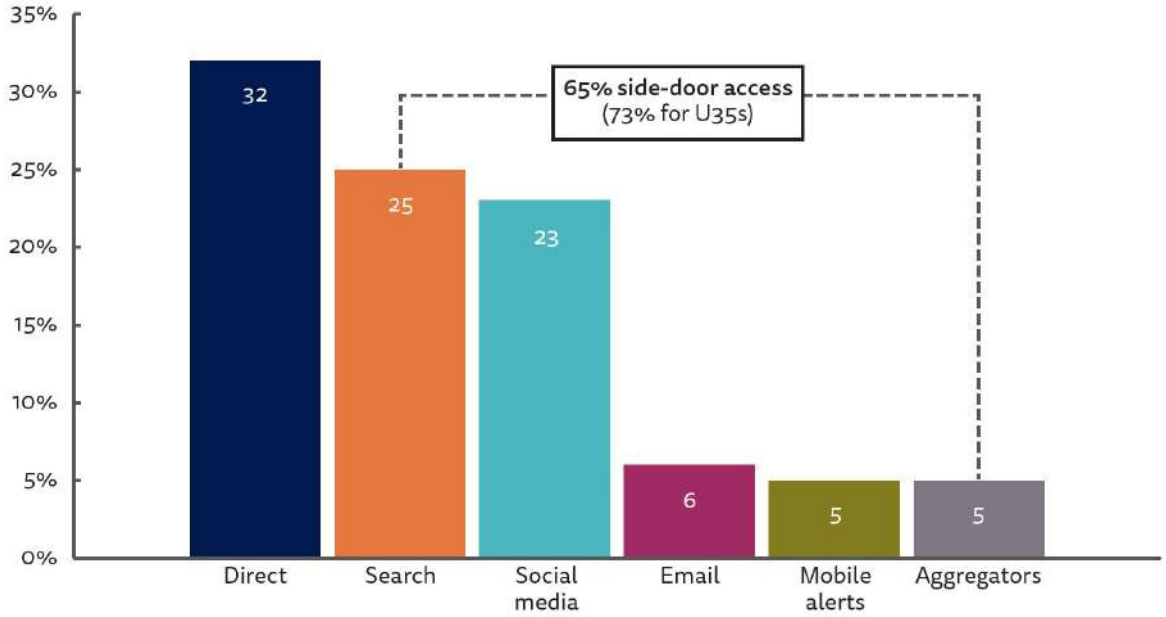
Savitri Dixon-Saxon, Ph.D., the vice provost for Walden University's College of Social and Behavioral Sciences

- lack of synchronicity
- no access to nonverbal cues provided by body language
- disconcerting framed heads of varying sizes
- talking at the mirror
- talking to "names instead of people"
- more than one **communication medium** at a time

"we get more biochemical bang for our buck during face-to-face contact because it offers a richer stream of social signals"

(Susan Pinker)

Digital media and news



Scrolling

Sharing without caring …

Disinformation

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D., and Nielsen, R. K. 2017. *Reuters Institute Digital News Report*. Reuters Institute for the Study of Journalism.

Human Computer Interaction

- Recommender systems
- From chatbots to spoken dialogue systems
- Social robots

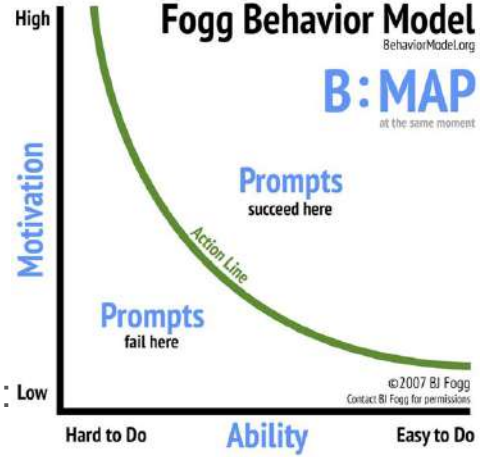
!!! More memory than humans…

Issue

The perlocutionary effect of RS changes across contexts…

E-commerce: Habit change:

Redundancy can prompt motivation



Redundancy kills motivation

A: “I always liked this eat always

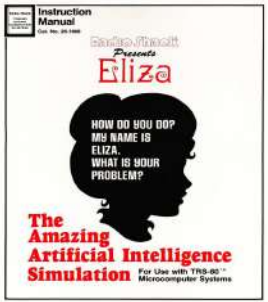
shape of dress…I will buy it again: healthy crap”

A: “I am tired to

the same

Let’s build a recommender system that promotes curiosity!

Bots



Bot is the short for *software robot:* chunk of software code designed to accomplish some particular routine task automatically and autonomously:

- **Googlebot**: web crawler that seeks out and ranks web pages

- **Spambots**: automatic generation of messages that overwhelm your inbox... marketing purposes or phishing campaigns...

- **Chat(ter)bots**: simulates human conversation either via text or text-to speech

Recommender systems

“Recommender systems propose ranked lists of items (that are subsets of a larger collection) according to their presumed relevance to individual users.” (Dietmar et al. 2012)

RS --- Standpoint: “you should buy/listen/eat x”

Arguments: “…”

Content-base

Collaborative filtering

Inductive and analogical argument schemes

Chatbots are software agents capable of **conversing** with users...

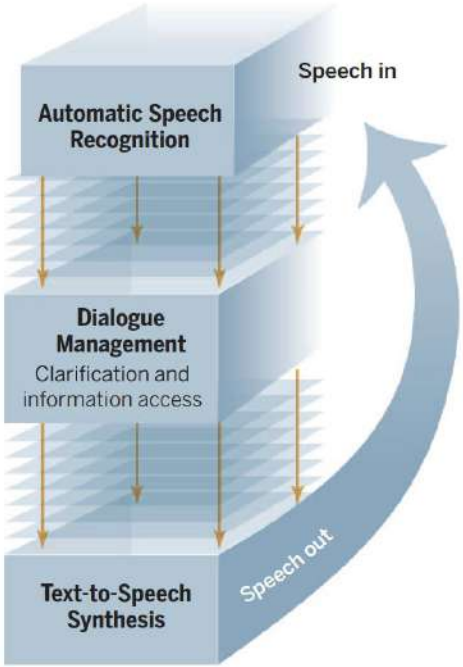
- 1966: ***ELIZA*** --- Joseph Weizenbaum: pattern matching and substitution methodology to simulate conversation.
- 1972: ***PARRY***--- Kenneth Colby --- The program imitated a patient with schizophrenia
- 1988: ***JABBERWACKY ---*** Rollo Carpenter ---- It aimed to simulate a natural human conversation in an entertaining
- 1992: ***DR. SBAITSO ---*** Creative Labs for MS-Dos --- first attempt to incorporate AI into a chatbot
- 1995: ***A.L.I.C.E ---*** Richard Wallace --- The program works with the XML schema known as artificial intelligence markup language (AIML), which helps specify conversation rules.
- 2001: ***SmartChild ---***The chatbot was available on AOL IM and MSN Messenger with the strength to carry out fun conversations with quick data access to other services.
- ***2012: Google Now –*** Google ***---*** It answers questions, performs actions through requests made to a set of web services and makes recommendations.
- ***2014: Cortana ---*** Microsoft ***---*** this program uses voice recognition and relevant alg and respond to voice commands.
- ***2014: Alexa ---*** Amazon ***---*** Intelligent personal assistant

Spoken Dialogue Systems

Also called.. smart speakers...digital assistants...vocal social agents... digital assistants... conversational agents ...

a few advancements... “automatic speech recognition (ASR) to identify what humans say; dialogue management (DM), to determine what the human wants; actions to obtain the information or perform the activity requested; and text-to-speech synthesis (TTS) to convey the information back to the human in spoken form” (Hirschber and Manning 2015: 262)

Structure of SDS



NLU: Natural Language Understanding

NLG: Natural Language Generation

Hirschberg, J. and Manning, C.D., 2015. Advances in natural language processing. *Science*, *349*(6245), pp.261-266.

Social robot

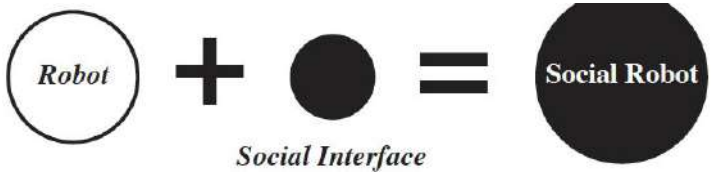
Etymology of robot: < Czech *robot* (1920 in *R.U.R.: Rossum's Universal Robots*, a play by Karel Čapek (1890–1938), Czech author) < *robota* forced labour, drudgery

Pinocchio

https://tinyurl.com/y2ys72fy

A robot is a physical object that functions in a an autonomous and situated way

a social robot is:



Hegel, F., Muhl, C., Wrede, B., Hielscher-Fastabend, M. and Sagerer, G., 2009, February. Understanding social robots. In *2009 Second International Conferences on Advances in Computer-Human Interactions* (pp. 169-174). IEEE

Cookies

A cookie is text file stored on your hard drive (more precisely in your browser folder) when you visit a website.

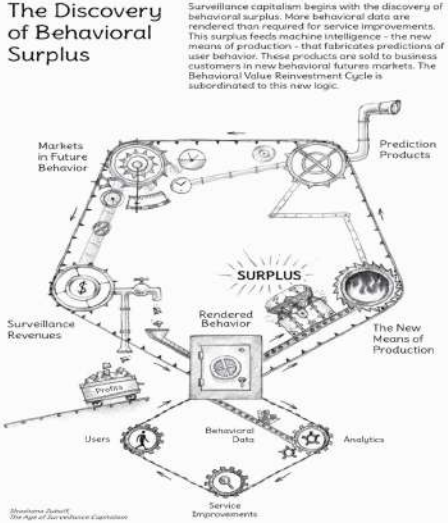
Three main types...:

1. **Session**
2. **Permanent**
3. **Third-party**

Under GDPR, “all EU member states must treat cookies and other technical identifiers as personal data.” □ Cookies banners

Surveillance capitalism?

- **Data collection is more likely to be automated involving machines rather than (or in addition to) involving humans.**
  - Facebook, Twitter, Amazon, Netflix, your bank card, mobile phone, most apps etc. etc.
  - All these collect and share data on your behaviours – purchase, messages, friends, health etc.
- **…data often resides with third parties. Data is available in real time and data collection can be continuous and offer information on the past, present and future**
  - All the above share your data with partners to offer adverts, suggest products, suggest content, and to analyse your data for other purposes
- **There may be only a short interval between the discovery of the information and the taking of action**
  - Watch how quick adverts match your web searches on Facebook…



Zuboff, S., 2019. *The age of surveillance capitalism: The fight for a human future at the new frontier of power: Barack Obama's books of 2019*. Profile books.

Me and My Big Data Resources: Related Toolkits and Guides

My and My Big Data: Related Toolkits and Guides · 28.3 KB · View full-size · Download

Data Management and Data Understanding

- **Data Detox Kit by Tactical Tech**

The Data Detox Kit is an 8-day step-by-step guide on how to reduce data traces online.

Let's reverse-engineer the process: AI as an educational tool …

https://www.liverpool.ac.uk/humanities-and-social-sciences/research/research-themes/centre-for-digital-humanities/projects/big-data/

Statistic for Journalism and Communication

KNOWING HOW TO READ DATA

Scuola Secondaria di secondo grado; Argomento: Introduzione – Leggere i dati (31.10.2014);
Pacchetto: S2.A.5

Istat

Statistics in daily life

Istat

INDEX

Istat

Statistic and daily life

Statistic is part of our daily life:

Tables, graphs, indicators, averages, …. Help us to represent, in a concise manner, the world we live in and make decision under uncertainty conditions.

They help to determine our vision of the world, to form a common sense

» Common sense:good sense and sound judgement in practical matters. (Oxford dictionary)

Statistic and daily life

Statistic is part of our lives, starting with the alarm clock in the morning. Everyone adjusts it based on their own experience, to get in time, without having to rush to not be late but without giving up sleep «more than necessary».

In short, everyone has «made a statistic» to set the alarm… even without realizing it…

Statistic and daily life

The **cookies we eat for breakfast** are the result of statistical studies of consumer taste.
Even their arrangement on supermarket shelves comes from statistical analysis of consumer behavior during the purchase.

Statistic is the basis of **television programming:** the time slots of programs and their eventual repetition are determined on audience data

How is the statistic perceived?

Statistic is often **perceived** based on **two opposing attitudes:**

- As **tool for manipulating reality**, whose variety of manifestations cannot be brought back within the «cage» of synthetic and simplistic representations.

secondo le statistiche d'adesso
risulta che te tocca un pollo all'anno
e, se nun entra nelle spese tue
t'entra nelle statistiche lo stesso
perché c'è un antro che se magna due
(Trilussa)

- As **absolute and undeniable truth**
  The datum is itself eloquent; the datum is the objective reality

This is because synthesis produces

As well as a GAIN

Summary measures (e.g., averages) provide summary information on numerous collectives

Average of books read in a year = 4

Students	Books read in a year
Giacomo	6
Mirella	6
Luca	0
Andrea	5
Valerio	3
Martina	6
Anna	2

… also a LOSS

Summary measures lose information about individual observed units

Metadata:

The information to understand the data

Istat

Metadata

- For a proper reading of statistical data, **meta-information is** necessary, which consists on information on the procedures followed for data collection and processing, on definitions and on used classifications, type of survey (total or sample), etc.
- Beyond the numbers and summaries that statistic produces, are fondamental the «additional» informations that tell «how» we arrived at such conclusions, for example:
  - How many cases have we observed? The whole population or only a part of it?
  - What computations? What calculations were performed?

One of the basic principles of public statistics

Statistical institutions should have the right to provide comments on misinterpretations and on improper uses

It is the IV of the basic principles of official statistics adopted by the UN and implies a dual right-duty:

1) Public statistics must oppose any kind of misinterpretation

2) Public statistics bodies are allowed to comment on misleading interpretations and misure of their statistical results

Istat

Examples of «improper» use of statistic

Nevertheless, statistical data are sometimes measured

- **To show a certain truth, when perhaps it is not supported by any data and indeed seems to be denied**

ES.: a pharmaceutical company wants to market a new mouthwash against sore throats and claims that laboratory analysis shows that 10 mg of the active substance kills
30,000 microbes

Is this correct? Maybe yes, but one might wonder:

a) What is described happens in a laboratory, but in a human throat? Would it be equally effective?

b) Once the substance is diluted, to prevent it from burning my throat tissues, is it still effective?

c) But are those really the microbes that cause sore throats?

Istat

Nevertheless, statistical data are sometimes used improperly

- **To «Make a sensation»**

Ex:

When, many years ago, John Hopkins University (USA) also began accepting women as students, someone thought to report the news that **33.3%** of female students had married a teacher

… but at that time there were only three women enrolled and one had married a professor

Nevertheless, statistical data are sometimes used improperly

- **Because of trivial mistakes**

The numbers that are «missing»

When statistics «condense» information and summarize the data of a collective, it is essential to know:

- Both **absolute values** (the dimensional aspect) of the phenomena;
- And **the percentages**, namely il «the contribution» each party brings to the collective as whole.
- Esemple:

Female students	Absolute Fréquences
They married a teacher	1
They did NOT marry a teacher	2
TOTAL	3

$$\frac{1}{3} \times 100 = 33{,}3\%$$

In a certain year there were 3 female students and one of them married a teacher

ABSOLUTE and PERCENTAGE value: information is correct when both are provided

Istat

If the character is «too focused»

Among **16** successful American **women,** selected by the Boston Chamber of Commerce, during the mid-20 century, almost **60 gained academic degrees and had 18 children**

… but there were two «special» women in the group, Virginia Gildersleeve, president of Barnard College, and Lillian M. Gilbreth, known, along with her husband, in the development of industrial technologies

The two together had one-third of all academic degrees, and 12 of the «18 children» were Mrs. Gilbreth's

Istat

What is the average number of children per woman?

Taking up the example of 16 successful women and their 18 children:

- If we keep all women with their respective offspring in the collective:

$\frac{18}{16}$ /1,12 Children per woman

- If we exclude Mrs. Gilbreth with her 12 children from the group

$\frac{6}{15}$ 0,4 Children per woman

Not even «half» child per woman!!!

Istat

What would you think if I told you that…

Table 1.1.1 – activities performed on an average weekly day by the population aged 15 years and older by type of activity and some characteristics

Years 2002-2003 - *(average generic duration in hours and minutes and percentage share of time over 24 hours)*

FEATURES	Sleeping, eating and other personal care		Paid work		Education and training	
	Mg	% sulle 24 ore	Mg	% sulle 24 ore	Mg	% sulle 24 ore

… In 2002-2003 the Italian population aged 15 years and older was working – gainfully employed – on average 2 hours and 36 minutes per day?

Fonte: Istat, 2007, *L'uso del tempo. Indagine multiscopo sulle famiglie "Uso del tempo" - Anni 2002-2003,* Istat, Roma (Informazioni, n.2).

Istat

Observe the results more carefully

Table 1.1.10 – Weekday (Mon – Fri) activities carried out by the population aged 15 and older by type of activity and some characteristics

Years 2002-2003 *(average specific duration in hours and minutes and frequency of participation in percent)*

FEATURES	Sleeping, eating and other personal care		Paid work		Education and training	
	Ms	%	Ms	%	Ms	%
TOTAL	11:43	100,0	7:41	41,5	6:04	8,2
DEMOGRAPHIC CHARACTERISTICS						
AGED CLASSES						
15-24	11:31	100,0	7:28	26,4	6:33	50,3
25-44	10:58	100,0	7:48	66,6	4:49	4,3
45-64	11:23	100,0	7:38	45,5	2:35	0,9
65 and more	13:30	100,0	6:31	3,6	1:48	0,2

Fonte: ISTITUTO NAZIONALE DI STATISTICA, 2007a. *L'uso del tempo. Indagine multiscopo sulle famiglie "Uso del tempo" - Anni 2002-2003*. Roma: Istat. (Informazioni, n.2). Disponibile su <http://www.istat.it/dati/catalogo/20070301_00/> [25 ottobre 2010]

Istat

How can these differences be explained?

It he seemingly **contrasting** data in two tables compared can be explained by taking the following factors into consideration:

- **Weekly reporting day**
  - Every day of the week
  - From Monday to Friday

- **«Type» of media used**:
  - **Generic** average
  - **Specific** average

Istat

Weekly average day and generic average

Table 1.1.1 – activities performed on an average weekly day by the population aged 15 years and older by type of activity and some characteristics

Average daily duration of paid work activity = 2h 36 min

- It refers to the **average weekly day,** «to the construction of which all days of the week contribute» including Saturdays and Sundays

- **The average is generic**

Istat

Weekday and specific average

Table 1.1.10 - weekday (Mon. – Fri.) activities carried out by the population aged 15 and older by type of activity and some characteristics

Average daily duration of paid work activity

= 7h 41 min

- It refers only to **weekdays,** i.e., it considers only days from Monday to Friday

- **The average is specific**

Generic average and specific average

- **Generic average**

  « In calculating generic averages, durations refer to the total population (…). For example, the average generic duration of an activity indicated the average time spent on that activity **by the entire population, including those who did not do it».**

- **Specific average**

  « This indicator is calculated **only on** the **overall** population **that actually performed an activity**»

Istat

“Misleading” graphs

Graphical representations

They are **statistical tools** that enable:

- Interpret **the information** gathered about the observed phenomenon **more quickly**,
- To **immediately** grasp some of its **characteristics**

This does **not** mean that graphs can **replace** the numbers in tables: they should only provide additional but useful support for statistical analysis

Graphical or tabular representation?

Some advantages the graphs have over the tables they accompany are:

- **Immediate visualization of the trend of the phenomenon** (e.g., is it increasing or is it decreasing?) and **the structure of** the distribution (e.g., are more males or females?), which allows **a comprehensive description of** the data
  - **Synthesis and thus possibility, in a small space, to compare multiple distributions** (curves, breaks, etc.)
  - **More popularized form for statistical data than what allowed by the tabular form**

Istat

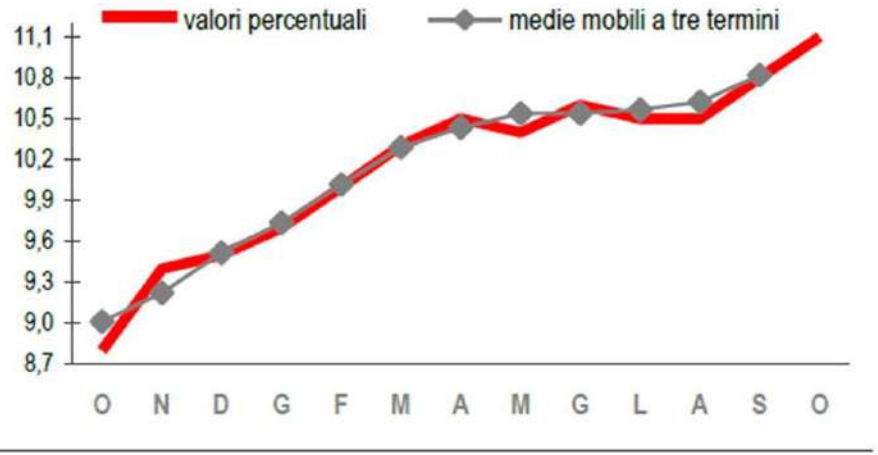
In any case, it is good to remember that …

... for a graphical representation to be useful and effective it must contain all the information necessary for understanding the data represented in it, namely:

- The **title**, which should indicate the **subject, place** and **time** to which the data refer
- The **character with the respective modes** (es.: "males" e "females" for the cariable "sex"), according to which statistical units are classified
- The **unit of measurement** used to graduate the axes
- The **source** of the data

But even when a «well done» graphis …

TASSO DI DISOCCUPAZIONE. Ottobre 2011- ottobre 2012,
dati destagionalizzati, valori percentuali



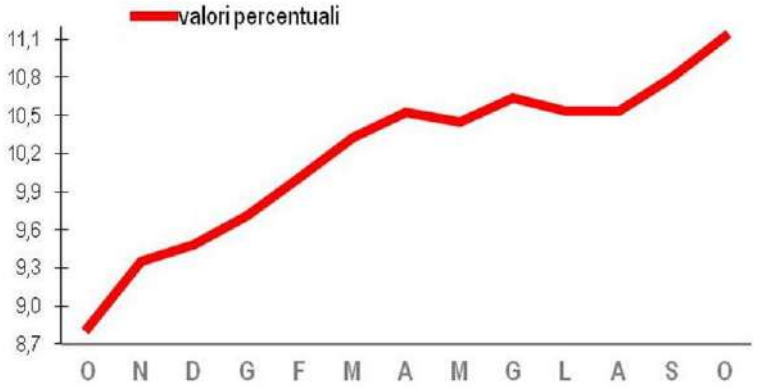
… if you do not look at it closely you can get the following out of it.

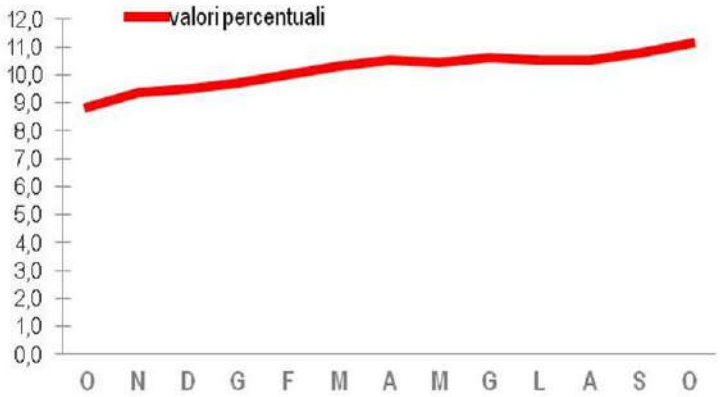
A distorted perception of the phenomenon!

Istat

The phenomenon is increasing: how much?

The two graphs **both** show **the rise** in the unemployment rate from October 2011 to October 2012





The only difference between the two is in the starting value of the y-axis: the top graph (the one in the publication) is missing the part between the oring (zero) and the value 8,7

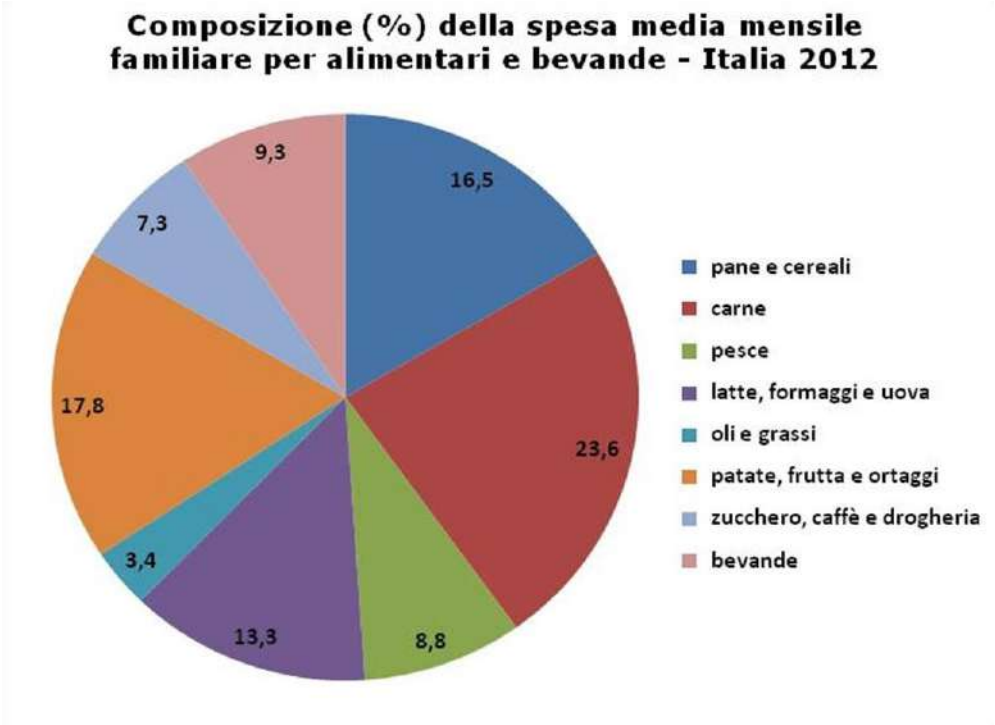
Istat

The table «below»

Year	Month	Unemployment rate (%)
2011	October	8,8
	November	9,4
	December	9,5
2012	January	9,7
	February	10,0
	March	10,3
	April	10,5
	May	10,5
	June	10,6
	July	10,5
	August	10,5
	September	10,8
	October	11,1

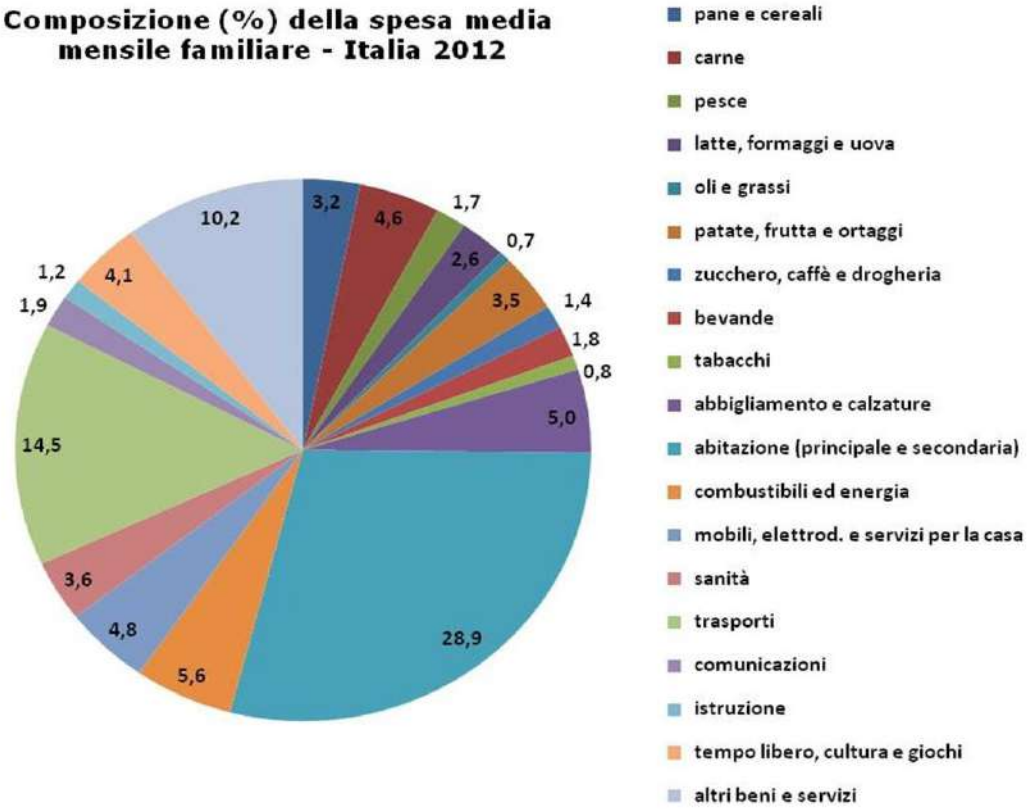
Istat

Pie charts are useful for representing the contribution of each part to the formation of the total



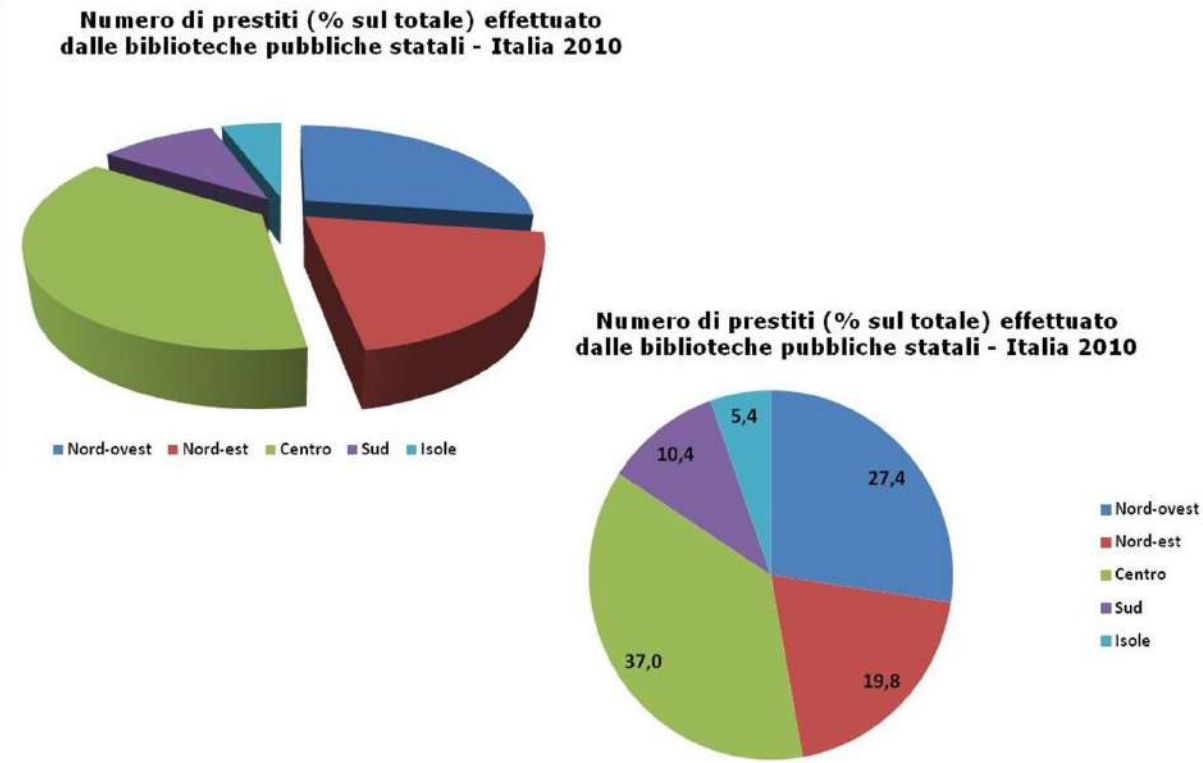
Istat

Too many categories!



Istat

The «exploded cakes» can mislead the perception of the phenomenon!

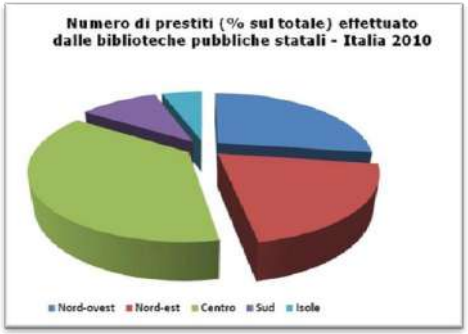


Conclusions…

Students	Books read in a year
Giacomo	6
Mirella	6
Luca	0
Andrea	5
Valerio	3
Martina	6
Anna	2

Average number of book read in a year = 4

… in order for statistics and its tools to prove useful and be fruitfully employed, it is good to learn to give numbers «a second look» !!!



Istat

LUISS

Data Lab

MediaFutures

DATALAB@LUISS.IT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951962