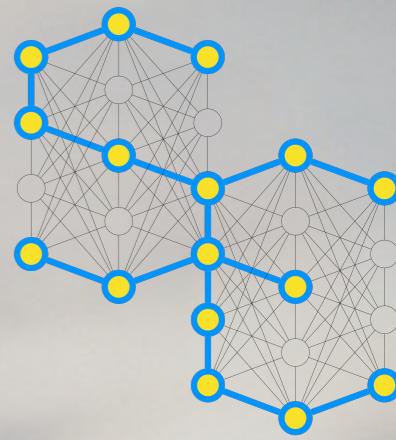


**800**  
ANNI  
1222-2022



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



# Overall Organisation

## Search Engines

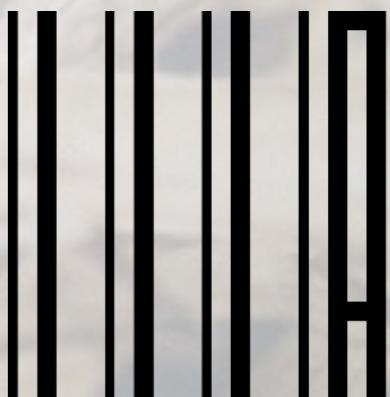
Master Degree in Computer Engineering

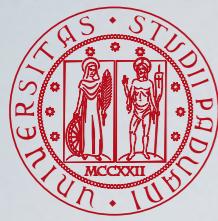
Master Degree in Data Science

Academic Year 2023/2024

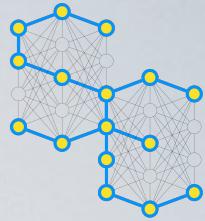
**Nicola Ferro**

Intelligent Interactive Information Access (IIIA) Hub  
Department of Information Engineering  
University of Padua



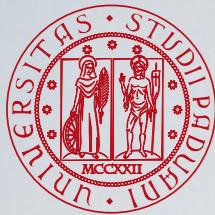


# Outline

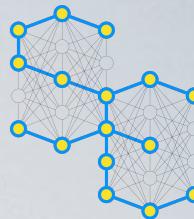


- Objectives
- Contents
- Organisation
- Exams

# Objectives and Contents



# Search Engines



Google search engine

search engine

Tutti Immagini Notizie Libri Video Altro Impostazioni Strumenti

Circa 1.640.000.000 risultati (0,57 secondi)

[www.searchenginejournal.com > ... Traduci questa pagina](#)

**17 Great Search Engines You Can Use Instead of Google**  
5 apr 2020 — 1. Bing. Bing Search Engine. As of January 2020, Microsoft sites handled a quarter of all search queries in the ...  
[Meet the 7 Most Popular... · The 11 Best Image Search... · SEO Tool Free Trials](#)

Le persone hanno chiesto anche

What are the top 5 search engines?  
What are the top 20 search engines?  
What is search engine and example?  
What are the 10 most popular search engines?

[www.searchenginejournal.com > ... Traduci questa pagina](#)

**The 7 Most Popular Search Engines in the World - SEO 101**  
7 gen 2018 — Meet the 7 Most Popular Search Engines in the World. Here are seven of the most popular search engines, a brief exploration of the SEO pros ...  
[en.wikipedia.org/wiki/Search\\_engine](#)

**Search engine - Wikipedia**  
A search engine is a software system that is designed to carry out web searches (Internet searches), which means to search the World Wide Web in a systematic ...  
[List of search engines · Search engine optimization · Search engine technology](#)

[searchengineland.com > ... Traduci questa pagina](#)

**Search Engine Land - News On Search Engines, Search ...**  
News On Search Engines, Search Engine Optimization (SEO) & Search Engine Marketing (SEM)

[www.searchenginewatch.com > ... Traduci questa pagina](#)

**Search Engine Watch: Homepage**  
Get the Latest daily news and insights about search engine marketing, SEO and paid search.  
[Sign up ...](#)

[www.searchenginewatch.com > ... Traduci questa pagina](#)

Motore di ricerca per internet (Search engine)

Generi di software

Tradotto dall'inglese - Un motore di ricerca è un sistema software progettato per effettuare ricerche sul web, il che significa cercare nel World Wide Web in modo sistematico particolari informazioni specificate in una query di ricerca web testuale. [Wikipedia \(inglese\)](#)

Vedi la descrizione originale

Ricerche correlate

Ricerca scientifica Web directory Scuola Design responsivo Search engine marketing

Vedi risultati per

Motore di ricerca Nell'ambito delle tecnologie di Internet, un motore di ricerca è un sistema automatico ...

Ottimizzazione per i motori di ricerca Con il termine ottimizzazione per i motori di ricerca si intendono tutte quelle attività ...

Microsoft Bing search engine

ALL WORK IMAGES VIDEOS MAPS NEWS SHOPPING

119.000.000 Results Date Language Region

**Onion Search Engine Anonymous - Anti-Censorship private search**  
[https://onionsearchengine.com](#)  
Ad Onion Search Engine Anonymous Anti-Censorship private search, no trace, no cookies, privacy is safe, no tracing, easy access to deepweb more services for protect your privacy

**Google**  
[www.google.com](#)  
Google allows users to search the Web for images, news, products, video, and other content.  
[Search · Gmail · Google Maps · Translate · Google News · Drive](#)

**DuckDuckGo — Privacy, simplified.**  
[https://duckduckgo.com](#)  
Other search engines track your searches even when you're in private browsing mode. We don't track you — period. Dismiss forever! Back to search. Already a fan? Invite friends to the Duck Side! Share DuckDuckGo and help friends take their privacy back! Tweet #ComeToTheDuckSide ...  
[HTML · APP](#)

**Yahoo Search - Web Search**  
[https://search.yahoo.com](#)  
The search engine that helps you find exactly what you're looking for. Find the most relevant information, video, images, and answers from all across the Web.

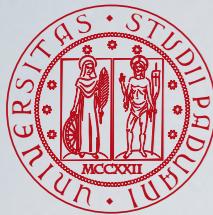
**The Best Search Engines of 2021 - Lifewire**  
[https://www.lifewire.com/best-search-engines-2483352](#)  
11/12/2020 — The Best Search Engines of 2021 Google Search. Favors fresh content. Ranks blogs and services. Accessible on any device. Collects information on users. Duck Duck Go Search. Does not track or store user information. Fast searches. Optional one ...

**Videos of Search engine**  
[bing.com/videos](#)

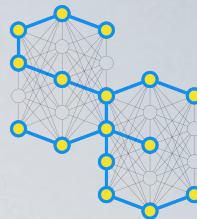
THE ORIGINAL SEARCH ENGINE 1:36 What Was The First Search Engine Before Google? | Mashable 2,8K views · 5 months ago YouTube > Mashable

How do Search Engines Work? 1:25 GoDaddy guides 506 views · gen 15, 2020 YouTube > GoDaddy

NON-PRIVATE SEARCH ENGINES 1:19 Stop paying with your privacy. How private search engines protect your perso... 5K views · 7 months ago YouTube > Startpage



# Search Engines



# Яндекс

**Yandex** search engine  Search

**Web** Images Video News Translate Disk Mail Ads

**17 Great Search Engines You Can Use Instead of Google**  
searchenginejournal.com > ...search-engines/271409/ This search engine is perfect if you need music for a video, an image for a blog post, or anything else without worrying about an angry artist coming after you for ripping off their...

**K 22 Alternative Search Engines To Use in 2021**  
kinsta.com > blog/alternative-search-engines/ Google isn't the only search engine out there. Check out the best alternative search engines for encrypted data, more privacy, and cool search features!

**W List of search engines - Wikipedia**  
en.wikipedia.org > List of search engines This is a list of search engines, including web search engines, selection-based search engines, metasearch engines, desktop search tools... Read more >

**L The Best Search Engines of 2021**  
lifewire.com > best-search-engines-2483352/ The many different search engines on the web are optimized to different things in different context. Pick the engine that meets your needs, not your habits.

**B 12 BEST Private Search Engines: Anonymous search - NO...**  
guru99.com > private-search-engines-anonymous-no... You should opt for a private search engine. Following is a curated list of Top 12 handpicked Private Search Engines with popular features and latest links.

**S 20 Of the Best Search Engines in the World Right Now 2020**  
zerodollartips.com > best-search-engines...the-world/ Different search engines come with different features and functions. If you are looking ... List of top 20 most popular search engines in the World 2021 | Top search engines...

**D DuckDuckGo — Максимум конфиденциальности, минимум...**  
duckduckgo.com > On other search engines, ads are based on profiles compiled about you using your personal information like search, browsing, and purchase history.

**Все настройки**  
Customize DuckDuckGo to your liking! For example, change the fonts and...

**Вакансии**  
Join the team at DuckDuckGo! Find out about our company culture, hiring...

**Узнать Больше**  
Search thousands of sites directly from DuckDuckGo.

**Top 8 Best Search Engines (of 2021) | RapidAPI**  
rapidapi.com > blog/best-search-engines/ Another outstanding search engine is Duckduckgo. Unlike the other search engines,...

# Baidu 百度

**Baidu** search engine  百度一下

百度为您找到相关结果约32,900,000个

**The Best Search Engine List On The Internet!**  
查看此网页的中文翻译, 请点击 翻译此页  
This search engine list is a real time saver! The web's best search engines, COMPLETE WITH SEARCH BOXES, on one page. Once a search term is ...  
www.20search.com/ 百度快照

**Search Engine Land - News On Search Engines, Search En...**  
查看此网页的中文翻译, 请点击 翻译此页  
News On Search Engines, Search Engine Optimization (SEO) & Search Engine Marketing (SEM)  
searchengineland.com/ 百度快照

**search engine - 百度百科**  
简介: search engine一般指搜索引擎。所谓搜索引擎, 就是根据用户需求与一定算法, 运用特定策略从互联网检索出制定信息反馈给用户的一门检索技术。搜索引擎依托于多种技术, 如网络爬虫技术、检索排序技术、网页处理技术、大数据处理技术、自然语言处理技术等, 为信...  
定义 发展历程 工作原理 分类 主要特点 体系结构 更多 >  
baike.baidu.com/

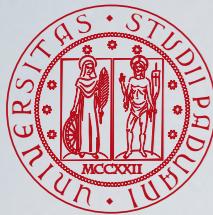
**search engine - 百度翻译**

**search engine**  
英['sɜːtʃ endʒɪn] 美['sɜːrf endʒɪn]  
n. (计算机)搜索引擎;  
[例句] Design of the Book Search Engine System Based on Deep Web  
基于deep web的图书搜索引擎系统设计  
[其他] 复数: search engines  
进行更多翻译

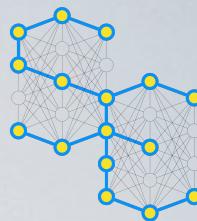
**Top 15 Best Search Engines | August 2020**  
Here are the top 15 Best Search Engines ranked by a combination of co...  
ntinually updated traffic statistics.  
www.ebizmba.com/articles/searc.../ 百度快照 - 翻译此页

**其他人在搜**  
search engines翻译 Google Everything Search 5 Best search engine appetizer  
telegramdownload database entrepreneur provider是什么软件 Time官网

**Search Engine**  
A leader since 1996, search engine submission, search engine optimization, free search engine s...



# Information Retrieval: Web Search



# Яндекс

**Yandex** search engine  Search

**Web** Images Video News Translate Disk Mail Ads

**17 Great Search Engines You Can Use Instead of Google**  
searchenginejournal.com > ...search-engines/271409/ This search engine is perfect if you need music for a video, an image for a blog post, or anything else without worrying about an angry artist coming after you for ripping off their...

**K 22 Alternative Search Engines To Use in 2021**  
kinsta.com > blog/alternative-search-engines/ Google isn't the only search engine out there. Check out the best alternative search engines for encrypted data, more privacy, and cool search features!

**W List of search engines - Wikipedia**  
en.wikipedia.org > List of search engines This is a list of search engines, including web search engines, selection-based search engines, metasearch engines, desktop search tools... Read more >

**L The Best Search Engines of 2021**  
lifewire.com > best-search-engines-2483352/ The many different search engines on the web are optimized to different things in different context. Pick the engine that meets your needs, not your habits.

**B 12 BEST Private Search Engines: Anonymous search - NO...**  
guru99.com > private-search-engines-anonymous-no... You should opt for a private search engine. Following is a curated list of Top 12 handpicked Private Search Engines with popular features and latest links.

**S 20 Of the Best Search Engines in the World Right Now 2020**  
zerodollartips.com > best-search-engines...the-world/ Different search engines come with different features and functions. If you are looking ... List of top 20 most popular search engines in the World 2021 | Top search engines...

**D DuckDuckGo — Максимум конфиденциальности, минимум...**  
duckduckgo.com > On other search engines, ads are based on profiles compiled about you using your personal information like search, browsing, and purchase history.

**Все настройки**  
Customize DuckDuckGo to your liking! For example, change the fonts and...

**Вакансии**  
Join the team at DuckDuckGo! Find out about our company culture, hiring...

**Узнать Больше**  
Search thousands of sites directly from DuckDuckGo.

**Top 8 Best Search Engines (of 2021) | RapidAPI**  
rapidapi.com > blog/best-search-engines/ Another outstanding search engine is Duckduckgo. Unlike the other search engines,...

# Baidu 百度

**Baidu** search engine  百度一下

百度为您找到相关结果约32,900,000个  搜索工具

**The Best Search Engine List On The Internet!**  
查看此网页的中文翻译, 请点击 翻译此页  
This search engine list is a real time saver! The web's best search engines, COMPLETE WITH SEARCH BOXES, on one page. Once a search term is ...  
www.20search.com/ 百度快照

**Search Engine Land - News On Search Engines, Search En...**  
查看此网页的中文翻译, 请点击 翻译此页  
News On Search Engines, Search Engine Optimization (SEO) & Search Engine Marketing (SEM)  
searchengineland.com/ 百度快照

**search engine - 百度百科**  
简介: search engine一般指搜索引擎。所谓搜索引擎, 就是根据用户需求与一定算法, 运用特定策略从互联网检索出制定信息反馈给用户的一门检索技术。搜索引擎依托于多种技术, 如网络爬虫技术、检索排序技术、网页处理技术、大数据处理技术、自然语言处理技术等, 为信...  
定义 发展历程 工作原理 分类 主要特点 体系结构 更多 >  
baike.baidu.com/

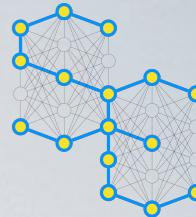
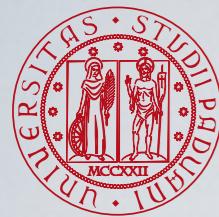
**search engine - 百度翻译**

**search engine**  
英['sɜːt̬ endʒɪn] 美['sɜːr̬ endʒɪn]  
n. (计算机)搜索引擎;  
[例句] Design of the Book Search Engine System Based on Deep Web  
基于deep web的图书搜索引擎系统设计  
[其他] 复数: search engines  
 扫码下载百度翻译APP

**Top 15 Best Search Engines | August 2020**  
Here are the top 15 Best Search Engines ranked by a combination of co...  
ntinually updated traffic statistics.  
www.ebizmba.com/articles/searc.../ 百度快照 - 翻译此页

**其他人在搜**  
search engines 翻译 Google Everything Search 5 Best search engine appetizer  
telegramdownload database entrepreneur provider是什么软件 Time官网

**Search Engine**  
A leader since 1996, search engine submission, search engine optimization, free search engine s...



# Is It Just Search?

Query Completion  
[recommendation,  
natural language processing]

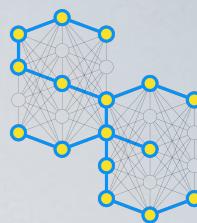
The screenshot shows a Google search interface. In the search bar, the text 'search|' is typed. A dropdown menu displays several suggestions:

- search console
- search
- Searching  
Film (2018)
- Search and Destroy  
Film (2020)
- search and destroy manga
- search engine
- search traduzione
- searchquarry
- searching for bobby fischer  
In cerca di Bobby Fischer — Film (1993)
- search engine optimization

Below the suggestions are two buttons: 'Cerca con Google' and 'Mi sento fortunato'. At the bottom right of the suggestions box is a link 'Segnala previsioni inappropriate Ulteriori informazioni'.



# Is It Just Search? SERP Features



Google

search engine

X | ⚡ | 🔎

Tutti

Immagini

Notizie

Libri

Video

Altro

Impostazioni

Strumenti

Circa 1.640.000.000 risultati (0,57 secondi)

[www.searchenginejournal.com > ...](#) Traduci questa pagina

**17 Great Search Engines You Can Use Instead of Google**

5 apr 2020 — 1. Bing. Bing Search Engine. As of January 2020, Microsoft sites handled a quarter of all search queries in the ...

[Meet the 7 Most Popular...](#) · [The 11 Best Image Search...](#) · [SEO Tool Free Trials](#)

Le persone hanno chiesto anche

What are the top 5 search engines?

What are the top 20 search engines?

What is search engine and example?

What are the 10 most popular search engines?

Feedback

[www.searchenginejournal.com > ...](#) Traduci questa pagina

**The 7 Most Popular Search Engines in the World - SEO 101**

7 gen 2018 — Meet the 7 Most Popular Search Engines in the World. Here are seven of the most popular search engines, a brief exploration of the SEO pros ...

[en.wikipedia.org > wiki > Search...](#) Traduci questa pagina

**Search engine - Wikipedia**

A search engine is a software system that is designed to carry out web searches (Internet searches), which means to search the World Wide Web in a systematic ...

[List of search engines](#) · [Search engine optimization](#) · [Search engine technology](#)

[searchengineland.com](#) Traduci questa pagina

**Search Engine Land - News On Search Engines, Search ...**

News On Search Engines, Search Engine Optimization (SEO) & Search Engine Marketing (SEM)

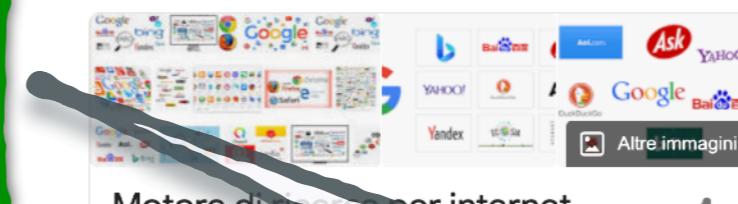
[www.searchenginewatch.com](#) Traduci questa pagina

**Search Engine Watch: Homepage**

Get the Latest daily news and insights about search engine marketing, SEO and paid search.

Sign up ...

[www.searchenginewatch.com](#) Traduci questa pagina



Motore di ricerca per internet  
(Search engine)

Genere di software

Tradotto dall'inglese - Un motore di ricerca è un programma informatico progettato per effettuare ricerche nel World Wide Web in modo sistematico e automatico, fornendo risultati specifici in una query di ricerca.

Vedi la descrizione originale

Ricerche correlate

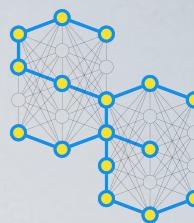


**Organic Search Results**  
[information retrieval]

**unici puri risultati di  
information retrieval**



# Is It Just Search? SERP Features



Google

search engine



Tutti Immagini Notizie Libri Video Altro Impostazioni Strumenti

Circa 1.640.000.000 risultati (0,57 secondi)

[www.searchenginejournal.com](http://www.searchenginejournal.com) > ... ▾ Traduci questa pagina

## 17 Great Search Engines You Can Use Instead of Google

5 apr 2020 — 1. Bing. Bing Search Engine. As of January 2020, Microsoft sites handled a quarter of all search queries in the ...

[Meet the 7 Most Popular...](#) · [The 11 Best Image Search...](#) · [SEO Tool Free Trials](#)

Le persone hanno chiesto anche

What are the top 5 search engines?

What are the top 20 search engines?

What is search engine and example?

What are the 10 most popular search engines?

[www.searchenginejournal.com](http://www.searchenginejournal.com) > ... ▾ Traduci questa pagina

## The 7 Most Popular Search Engines in the World - SEO 101

7 gen 2018 — Meet the 7 Most Popular Search Engines in the World. Here are seven of the most popular search engines, a brief exploration of the SEO pros ...

[en.wikipedia.org](http://en.wikipedia.org) > wiki > Search... ▾ Traduci questa pagina

## Search engine - Wikipedia

A search engine is a software system that is designed to carry out web searches (Internet searches), which means to search the World Wide Web in a systematic ...

[inf of search engines](#) · [Search engines](#) · [Search engine optimization](#) · [Search engine technology](#)

## Query Suggestion:

## Drill-down Queries

## [recommendation]

[www.searchenginewatch.com](http://www.searchenginewatch.com) > ... ▾ Traduci questa pagina

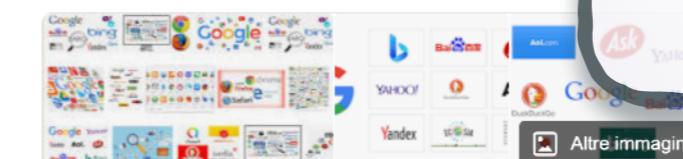
## Search Engine Watch: Homepage

Get the Latest daily news and insights about search engine marketing, SEO and paid search.

Sign up ...

[www.searchenginewatch.com](http://www.searchenginewatch.com) > ... ▾ Traduci questa pagina

Query Suggestion:  
Other Related Queries  
[recommendation]



## Motore di ricerca per internet (Search engine)

Genere di software

Tradotto dall'inglese - Un motore di ricerca è un sistema software progettato per effettuare ricerche sul web, il che significa cercare nel World Wide Web in modo sistematico particolari informazioni specificate in una query di ricerca web testuale. [Wikipedia \(inglese\)](#)

Vedi la descrizione originale ▾

## Ricerche correlate



Ricerca  
scientifica



Web  
directory



Scuola



Design  
responsivo



Search  
engine  
marketing

Visualizza altri 10 elementi

Feedback

## Vedi risultati per

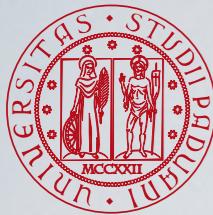
### Motore di ricerca

Nell'ambito delle tecnologie di Internet, un motore di ricerca è un sistema automatico ...

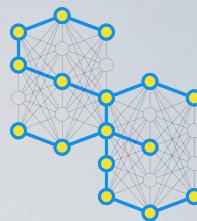


### Ottimizzazione per i motori di ricerca

Con il termine ottimizzazione per i motori di ricerca si intendono tutte quelle attività ...



# Is It Just Search? SERP Features



Google

search engine

X | 🎤 | 🔎

Tutti Immagini Notizie Libri Video Altro Impostazioni Strumenti

Circa 1.640.000.000 risultati (0,57 secondi)

[www.searchenginejournal.com](#) > ... ▾ Traduci questa pagina

**17 Great Search Engines You Can Use Instead of Google**

5 apr 2020 — 1. Bing. Bing Search Engine. As of January 2020, Microsoft sites handled a quarter of all search queries in the ...

[Meet the 7 Most Popular...](#) · [The 11 Best Image Search...](#) · [SEO Tool Free Trials](#)

Le persone hanno chiesto anche

What are the top 5 search engines?

What are the top 20 search engines?

What is search engine and example?

What are the 10 most popular search engines?

Feedback

[www.searchenginejournal.com](#) > ... ▾ Traduci questa pagina

**The 7 Most Popular Search Engines in the World - SEO 101**

7 gen 2018 — Meet the 7 Most Popular Search Engines in the World. Here are seven of the most popular search engines, a brief exploration of the SEO pros ...

[en.wikipedia.org](#) > wiki > Search... ▾ Traduci questa pagina

**Search engine - Wikipedia**

A search engine is a software system that is designed to carry out web searches (Internet searches), which means to search the World Wide Web in a systematic ...

[List of search engines](#) · [Search engine optimization](#) · [Search engine technology](#)

[searchengineland.com](#) ▾ Traduci questa pagina

**Search Engine Land - News On Search Engines, Search ...**

News On Search Engines, Search Engine Optimization (SEO) & Search Engine Marketing (SEM)

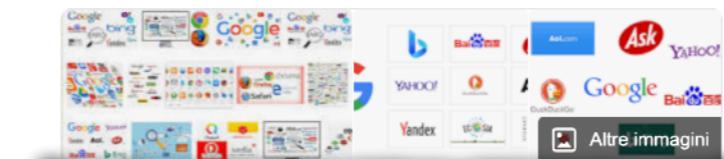
[www.searchenginewatch.com](#) ▾ Traduci questa pagina

**Search Engine Watch: Homepage**

Get the Latest daily news and insights about search engine marketing, SEO and paid search.

Sign up ...

[www.searchenginewatch.com](#) ▾ Traduci questa pagina



Motore di ricerca per internet  
(Search engine)

Genere di software

Tradotto dall'inglese - Un motore di ricerca è un sistema software progettato per effettuare ricerche sul web, il che significa cercare nel World Wide Web in modo sistematico particolari informazioni specificate in una query di ricerca web testuale. [Wikipedia \(inglese\)](#)

Vedi la descrizione originale ▾

Ricerche correlate



Ricerca scientifica



Web directory



Scuola



Design responsivo



Search engine marketing

Visualizza altri 10 elementi

Vedi risultati per



Motore di ricerca

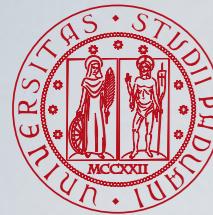
Nell'ambito delle tecnologie di Internet, un motore di ricerca è un sistema automatico ...



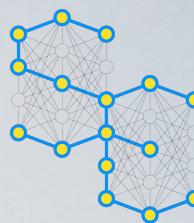
Ottimizzazione per i motori di ricerca

Con il termine ottimizzazione per i motori di ricerca si intendono tutte quelle attività ...

Knowledge Panel  
[natural language processing,  
semantics]



# Is It Just Search? Clicks, Interaction, Personalization



Google search engine

Circa 1.500.000.000 risultati (0,50 secondi)

[www.searchenginejournal.com](http://www.searchenginejournal.com) › ... ▾ Traduci questa pagina

**17 Great Search Engines You Can Use Instead of Google**

5 apr 2020 — 1. Bing. Bing Search Engine. As of January 2020, Microsoft site handled a quarter of all search queries in the ...

Meet the 7 Most Popular... · The 11 Best Image Search... · DuckDuckGo SEO...

[en.wikipedia.org](http://en.wikipedia.org) › wiki › Search... ▾ Traduci questa pagina

**Search engine - Wikipedia**

A search engine is a software system that is designed to carry out web searches (Internet searches), which means to search the World Wide Web in a systematic ...

List of search engines · Search engine optimization · Search engine technology

Le persone hanno chiesto anche

What are the top 10 search engines?

What are the top 20 search engines?

What is search engine and example?

What are the 3 most popular search engines?

Feedback

searchengineland.com ▾ Traduci questa pagina

**Search Engine Land - News On Search Engines, Search ...**

News On Search Engines, Search Engine Optimization (SEO) & Search Engine Marketing (SEM)

[programmablesearchengine.google.com](http://programmablesearchengine.google.com) › ... ▾ Traduci questa pagina

**Programmable Search Engine by Google**

Help people find what they need on your website. Add a customizable search box to your web pages and show fast, relevant results powered by Google.

[www.searchenginewatch.com](http://www.searchenginewatch.com) ▾ Traduci questa pagina

**Search Engine Watch: Homepage**

Get the Latest daily news and insights about search engine marketing, SEO and paid search. Sign up ...

[www.reliablesoft.net](http://www.reliablesoft.net) › top-10-sea... ▾ Traduci questa pagina

**Top 10 Search Engines In The World (2021 Update)**

The Top 10 Most Popular Search Engines In The World · Google · Microsoft Bing · Yahoo ·

Motore di ricerca per internet (Search engine)

Genere di software

Tradotto dall'inglese - Un motore di ricerca è un sistema software progettato per effettuare ricerche sul web, il che significa cercare nel World Wide Web in modo sistematico particolari informazioni specificate in una query di ricerca web testuale.

[Wikipedia \(inglese\)](#)

Vedi la descrizione originale ▾

Ricerche correlate Visualizza altri 10 elementi

Ricerca scientifica Web directory Scuola Design responsivo

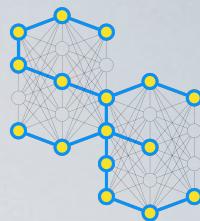
Vedi risultati per

Motore di ricerca Nell'ambito delle tecnologie di Internet, un motore di ricerca è ...

Ottimizzazione per i motori di ricerca Con il termine ottimizzazione per i motori di ricerca si intende ...



# Is It Just Search? Artificial Intelligence & Machine Learning



Google search results for "search engine":

Cerca: search engine

Tutti Immagini Notizie Libri Video Altro Impostazioni Strumenti

Circa 1.500.000.000 risultati (0,50 secondi)

[www.searchenginejournal.com](#) > ... ▾ Traduci questa pagina

**17 Great Search Engines You Can Use Instead of Google**

5 apr 2020 — 1. Bing. Bing Search Engine. As of January 2020, Microsoft sites handled a quarter of all search queries in the ...

Meet the 7 Most Popular... · The 11 Best Image Search... · DuckDuckGo SEO

[cn.wikipedia.org](#) > wiki > Search... ▾ Traduci questa pagina

**Search engine - Wikipedia**

A search engine is a software system that is designed to carry out web searches (Internet searches), which means to search the World Wide Web in a systematic ...

List of search engines · Search engine optimization · Search engine technology

Le persone hanno chiesto anche

- What are the top 10 search engines?
- What are the top 20 search engines?
- What is search engine and example?
- What are the 3 most popular search engines?

[searchengineland.com](#) ▾ Traduci questa pagina

**Search Engine Land - News On Search Engines, Search...**

News On Search Engines, Search Engine Optimization (SEO) & Search Engine Marketing (SEM)

[programmablesearchengine.google.com](#) ▾ Traduci questa pagina

**Programmable Search Engine by Google**

Help people find what they need on your website. Add a customizable search box to your web pages and show fast, relevant results powered by Google.

[www.searchenginewatch.com](#) ▾ Traduci questa pagina

**Search Engine Watch: Homepage**

Get the Latest daily news and insights about search engine marketing, SEO and paid search. Sign up ...

[www.reliablesoft.net](#) ▾ top-10-sea... ▾ Traduci questa pagina

**Top 10 Search Engines In The World (2021 Update)**

The Top 10 Most Popular Search Engines In The World · Google · Microsoft Bing · Yahoo ·

Motore di ricerca per internet (Search engine)  
Genere di software  
Tradotto dall'inglese - Un motore di ricerca è un sistema software progettato per effettuare ricerche sul web, il che significa cercare nel World Wide Web in modo sistematico particolari informazioni specificate in una query di ricerca web testuale.  
Wikipedia (inglese)  
Vedi la descrizione originale ▾

Ricerche correlate Visualizza altri 10 elementi

Feedback

Ricerca scientifica Web directory Scuola Design responsivo Feedback

Vedi risultati per Motore di ricerca Nell'ambito delle tecnologie di Internet, un motore di ricerca è ...  
Ottimizzazione per i motori di ricerca Con il termine ottimizzazione per i motori di ricerca si intende ...



# Is Search Solved?

jews

All Images • Videos Maps News Shopping

Jewish Scroll

Dirty Jew

Ugly Jews

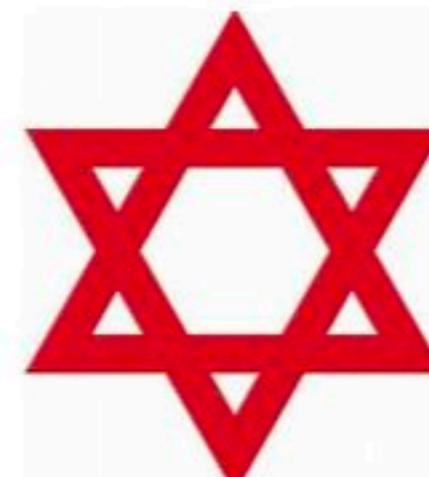
WW2 Jews

Evil Jew

Jew Hat

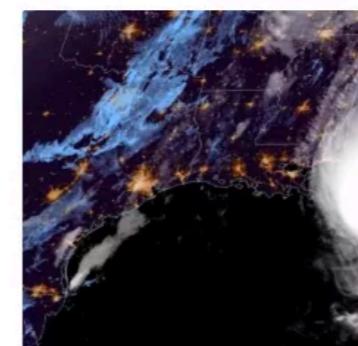
Jew Flag

The Word Jews



# YAHOO!

Mail News Finance



black people ar

black people are stupid

black people are animals

black people are racist

black people are t

black people are criminals

black people are guing

black people are ugly

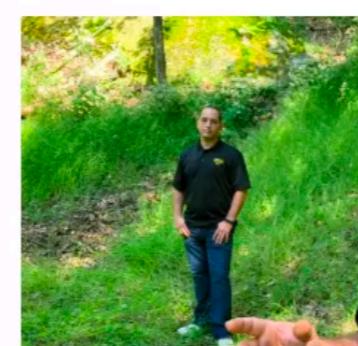
black people are the real jews

black people are lazy

black people are monkeys

# YAHOO!

Mail News Finance



Jews ar

jews are evil

jews are racist

jews around the world

jews are not people they are animals

jews are a race

jews are black

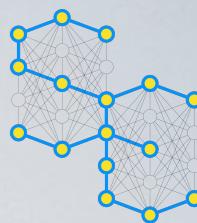
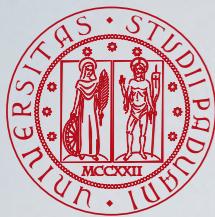
jews are the devil

jews are white

jews are aliens

jews are liars

<https://www.theverge.com/2018/10/10/17959328/bing-yahoo-offensive-search-suggestions-racism-antisemitism>



# Is Search Solved?

informazioni proposte da Google ricavate scansionando header di siti

Google search results for "covid-19":

- Notizie principali:**
  - Malattia da coronavirus (COVID-19)
  - Panoramica
  - Statistiche
  - Informazioni san...
  - afrontare
  - Notizie
  - Condividi
- Statistiche:**
  - Variazione giornaliera
  - Nuovi casi Italia Dall'inizio
  - 26 feb 2021 Nuovi casi: 20.488 Media 7 gg: 15.434
  - Ciascun giorno mostra i nuovi casi segnalati dal giorno precedente · Statistiche aggiornate meno di 1 giorno fa · Fonte: JHU CSSE COVID-19 Data · Informazioni su questi dati
- Notizie locali:**
  - I numeri del coronavirus a Brescia e in Lombardia
  - BRESCIA TODAY Coronavirus, allarme in paese: 100 positivi e 16 persone in ospedale
  - Il Fatto Quotidiano Covid a Brescia, il sindaco impone ulteriori restrizioni per il weekend: "Costretti a...

Google search results for "covid-19 weekend: "Costretti a...":

- Informazioni sanitarie:**
  - Sintomi Prevenzione Cure
  - La malattia da coronavirus (COVID-19) colpisce in vari modi. La maggior parte delle persone presenta sintomi lievi o moderati e guarisce senza necessità di ricovero in ospedale.
  - Sintomi più comuni:
    - febbre
    - tosse secca
    - spossatezza
  - Sintomi meno comuni:
    - indolenzimento e dolori muscolari
    - mal di gola
    - diarrea
    - congiuntivite
    - mal di testa
    - perdita del gusto o dell'olfatto
    - eruzione cutanea o scolorimento delle dita di piedi o mani
  - Ulteriori informazioni su who.int
  - Traduzione di SDI Inc. Solo a scopo informativo. Rivolgiti all'autorità sanitaria locale per consigli.
- Risorse di Google:**
  - Gli strumenti e le risorse di Google ti aiutano a rimanere sempre al corrente e al passo con le notizie
  - Risorse sul COVID-19
- Migliori risultati:**
  - www.salute.gov.it > dettaglioContenutiNuovoCoronavirus
  - Covid-19 - Situazione in Italia - Ministero della Salute**

Il monitoraggio dell'epidemia dei casi di Covid-19 in Italia viene effettuato attraverso due flussi di dati: il flusso dei dati aggregati inviati dalle Regioni coordinato da ...

Monitoraggi Covid-19 · Covid-19, ministro Speranza... · Notizie · Pubblicazioni
  - www.salute.gov.it > dettaglioFaqNuovoCoronavirus
  - Che cos'è il nuovo coronavirus - Ministero della Salute**

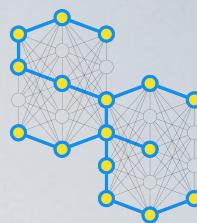
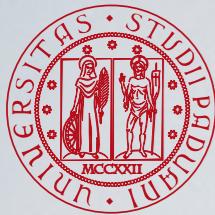
Cosa è SARS-CoV-2? Cos'è COVID-19?
  - it.wikipedia.org > wiki > COVID-19
  - COVID-19 - Wikipedia**

La COVID-19, acronimo dell'inglese CoronaVirus Disease 19, conosciuta anche come malattia respiratoria acuta da SARS-CoV-2 o malattia da coronavirus ...

Mortalità mondiale: 0,5 - 1% ICD-10: U07.1

Eziologia: Infezione da SARS-CoV-2

COVID-19 in gravidanza · Pandemia di COVID-19 · Vaccino anti COVID-19 · Brasile
- Autorità locali e sanitarie su Twitter:**
  - Dipartimento Prot...
  - Ministero della S...
  - Ministero della S...



# Is Search Solved?

meccanismi intrinseci di search engines -> spesso cercano SERENDIPITY: anche risultati inaspettati, anziché solo risultati migliori non va bene per ogni problema (e.g. problemi di salute)

HARVARD KENNEDY SCHOOL | SHORENSTEIN CENTER ON MEDIA, POLITICS, AND PUBLIC POLICY

CONTACT [Twitter](#) [Facebook](#) [Search](#)

HARVARD KENNEDY SCHOOL

**Misinformation Review**

EXPLORE the review

SUBMIT an essay

REVIEW for Misinfo

ABOUT the journal

MAY 11, 2020

SHARE [Facebook](#) [Twitter](#) [Email](#) [Download PDF](#)

PEER REVIEWED

## How search engines disseminate information about COVID-19 and why they should do better

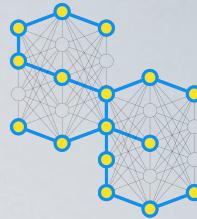
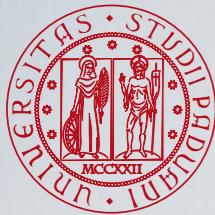
*Access to accurate and up-to-date information is essential for individual and collective decision making, especially at times of emergency. On February 26, 2020, two weeks before the World Health Organization (WHO) officially declared the COVID-19's emergency a "pandemic," we systematically collected and analyzed search results for the term "coronavirus" in three languages from six search engines. We found that different search engines prioritize specific categories of information sources, such as government-related websites or alternative media. We also observed that source ranking within the same search engine is subjected to randomization, which can result in unequal access to information among users.*

BY MYKOLA MAKHORTYKH  
University of Bern, Institute of Communication and Media Studies

ALEKSANDRA URMAN  
University of Bern, Institute of Communication and Media Studies

ROBERTO ULLOA  
GESIS – Leibniz Institute for the Social Sciences

<https://misinforeview.hks.harvard.edu/article/how-search-engines-disseminate-information-about-covid-19-and-why-they-should-do-better/>



# Is Search Solved?

query completion  
dipende da training set

RUTGERS THE STATE UNIVERSITY OF NEW JERSEY 🔍 ≡

RUTGERS TODAY Explore Topics ▾

Research & Innovation

## Online Autocompletes Are More Likely to Yield COVID-19 Misinformation in Spanish than in English

By Megan Schumann

Date June 16, 2020

Media Contact Megan Schumann 848-445-1907 [megan.schumann@rutgers.edu](mailto:megan.schumann@rutgers.edu)

Share 



Rutgers researchers find Spanish autocompletes are more likely to yield harmful, negative results

Online autocomplete results for COVID-19 related information are more likely to yield misleading results if the user types in Spanish than in English, according to a [new Rutgers report](#).

The difference may harm Spanish speakers by connecting them with misinformation about handwashing, sanitizers, masks or the disease itself, according to lead author [Vivek Singh](#), an assistant professor at Rutgers-New Brunswick's [School of Communication and Information](#).

<https://www.rutgers.edu/news/online-autocompletes-are-more-likely-yield-covid-19-misinformation-spanish-english>



# Is Search Solved?

MIT  
Technology  
Review

Artificial intelligence / Machine learning

## We read the paper that forced Timnit Gebru out of Google. Here's what it says.

The company's star ethics researcher highlighted the risks of large language models, which are key to Google's business.

by Karen Hao

December 4, 2020



COURTESY OF TIMNIT GEBRU

On the evening of Wednesday, December 2, Timnit Gebru, the co-lead of Google's ethical AI team, announced via Twitter that the company had forced her out.

Gebru, a widely respected leader in AI ethics research, is known for coauthoring a groundbreaking paper that showed facial recognition to be less accurate at identifying women and people of color, which means its use can end up discriminating against them. She also cofounded the Black in AI affinity group, and champions diversity in the tech industry. The team she helped build at Google is one of the most diverse in AI and includes many leading experts in their own right. Peers in the field envied it for producing critical work that often challenged mainstream AI practices.

### Popular

01.

The coming war on the hidden algorithms that trap people in poverty

02.

Logging in to get kicked out: Inside America's virtual eviction crisis

03.

This is the most precise 3D map of the Milky Way ever made

MIT Technology Review

Topics Magazine Newsletters Events

Sign in

Subscribe



### Massive data, inscrutable models

Large language models are also trained on exponentially increasing amounts of text. This means researchers have sought to collect all the data they can from the internet, so there's a risk that racist, sexist, and otherwise abusive language ends up in the training data.

An AI model taught to view racist language as normal is obviously bad. The researchers, though, point out a couple of more subtle problems. One is that shifts in language play an important role in social change; the MeToo and Black Lives Matter movements, for example, have tried to establish a new anti-sexist and anti-racist vocabulary. An AI model trained on vast swaths of the internet won't be attuned to the nuances of this vocabulary and won't produce or interpret language in line with these new cultural norms.

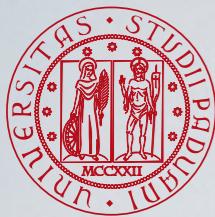
It will also fail to capture the language and the norms of countries and peoples that have less access to the internet and thus a smaller linguistic footprint online. The result is that AI-generated language will be homogenized, reflecting the practices of the richest countries and communities.

Moreover, because the training data sets are so large, it's hard to audit them to check for these embedded biases. "A methodology that relies on datasets too large to document is therefore inherently risky," the researchers conclude. "While documentation allows for potential accountability, [...] undocumented training data perpetuates harm without recourse."

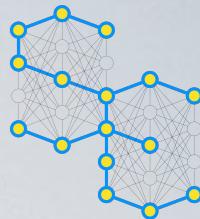
### Research opportunity costs

The researchers summarize the third challenge as the risk of "misdirected research effort." Though most AI researchers acknowledge that large language models don't actually understand language and are merely excellent at manipulating it, Big Tech can make money from models that

<https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/>



# Is Search Solved? CORD-19



AI2 Allen Institute for AI

Research ▾

## CORD-19: COVID-19 Open Research Dataset

Semantic Scholar • 2020

CORD-19 is a free resource of tens of thousands of scholarly articles about COVID-19, SARS-CoV-2, and related coronaviruses for use by the global research community.

Download

Read Paper

View Website

License: Semantic Scholar Dataset License

**AI2**

AI FOR THE COMMON GOOD

Email us: ai2-info@allenai.org

Call us: 206.548.5600

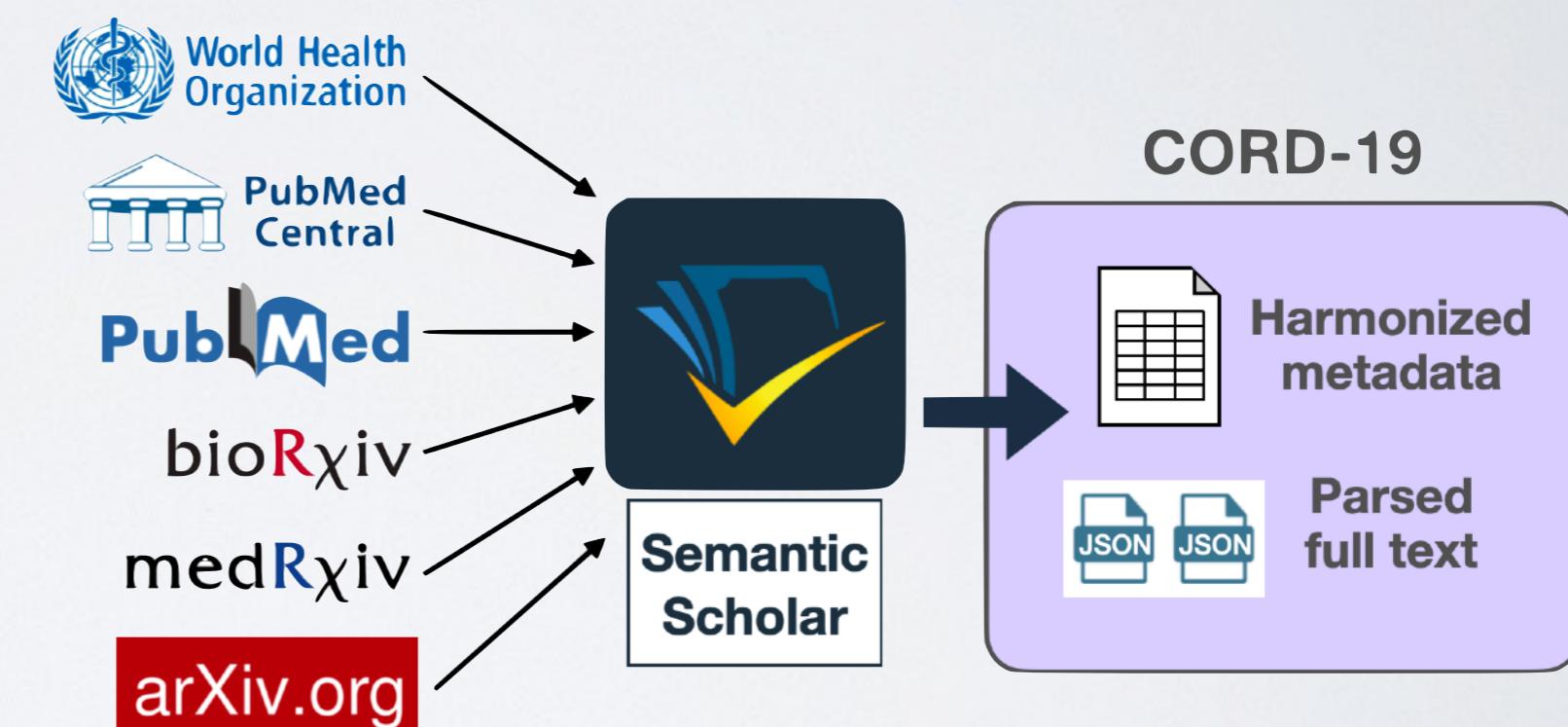
Follow us: @allen\_ai

Subscribe to the AI2 Newsletter

Privacy Policy | Terms and Conditions

© The Allen Institute for Artificial Intelligence - All Rights Reserved.

<https://allenai.org/data/cord-19>



Wang, L. L., Lo, K., Chandrasekhar, Y., Reas, R., Yang, J., Burdick, D., Eide, D., Funk, K., Katsis, Y., Kinney, R., Li, Y., Liu, Z., Merrill, W., Mooney, P., Murdick, D., Rishi, D., Sheehan, J., Shen, Z., Stilson, B., Wade, A., Wang, K., Ru Wang, N. X., Wilhelm, C., Xie, B., Raymond, D., Weld, D. S., Etzioni, O., and Kohlmeier, S. (2020). CORD-19: The COVID-19 Open Research Dataset. *arXiv.org, Digital Libraries (cs.DL)*, arXiv:2004.10706.

<https://arxiv.org/abs/2004.10706v4>



# Is Search Solved? Google on CORD-19



## Google AI Blog

The latest news from Google AI

### An NLU-Powered Tool to Explore COVID-19 Scientific Literature

Monday, May 4, 2020

Posted by Keith Hall, Research Scientist, Natural Language Understanding, Google Research

Due to the COVID-19 pandemic, scientists and researchers around the world are publishing an immense amount of new research in order to understand and combat the disease. While the volume of research is very encouraging, it can be difficult for scientists and researchers to keep up with the rapid pace of new publications. Traditional search engines can be excellent resources for finding real-time information on general COVID-19 questions like "How many COVID-19 cases are there in the United States?", but can struggle with understanding the meaning behind research-driven queries. Furthermore, searching through the existing corpus of COVID-19 scientific literature with traditional keyword-based approaches can make it difficult to pinpoint relevant evidence for complex queries.

To help address this problem, we are launching the [COVID-19 Research Explorer](#), a semantic search interface on top of the [COVID-19 Open Research Dataset \(CORD-19\)](#), which includes more than 50,000 journal articles and preprints. We have designed the tool with the goal of helping scientists and researchers efficiently pore through articles for answers or evidence to COVID-19-related questions.

When the user asks an initial question, the tool not only returns a set of papers (like in a traditional search) but also highlights snippets from the paper that are possible answers to

<https://covid19-research-explorer.appspot.com/>

COVID-19 Research Explorer

Send Feedback

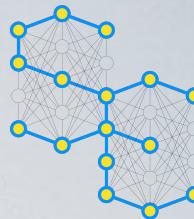
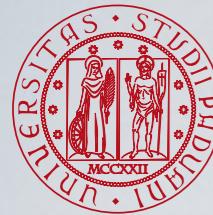
Get answers to complex scientific questions related to COVID-19

Based on 100,000+ scholarly articles from CORD-19

Ask a question like "What are the rapid molecular diagnostic"

Learn more

<https://ai.googleblog.com/2020/05/an-nlu-powered-tool-to-explore-covid-19.html>



# Is Search Solved?

<https://www.theguardian.com/technology/2021/dec/29/amazons-alexa-child-penny-live-plug>

**Support the Guardian**  
Available for everyone, funded by readers  
[Contribute →](#) [Subscribe →](#)

Search jobs

News Opinion Sport Culture Lifestyle More ▾  
World UK Coronavirus Climate crisis Environment Science Global development Football Tech Business Obituaries

Amazon Alexa This article is more than 1 month old

## Amazon's Alexa device tells 10-year-old to touch a penny to a live plug socket

The child had asked the Echo smart speaker for a challenge, prompting her mother to post the response on Twitter

**Maya Yang**  
Thu 29 Dec 2021 00.09 GMT

A mother tweeted her outrage after her Echo smart speaker told her young daughter to touch a penny to a live plug socket. Photograph: Mike Blake/Reuters

Virtual assistants can set timers for people, play music, control smart home devices, respond to voice commands and set up reminders. As of Sunday, they have also proven their ability to challenge children to lethal dares.

Alexa, Amazon's virtual assistant, recently advised a 10-year-old girl to touch a penny to a live plug socket after she asked the Echo smart speaker for a challenge.

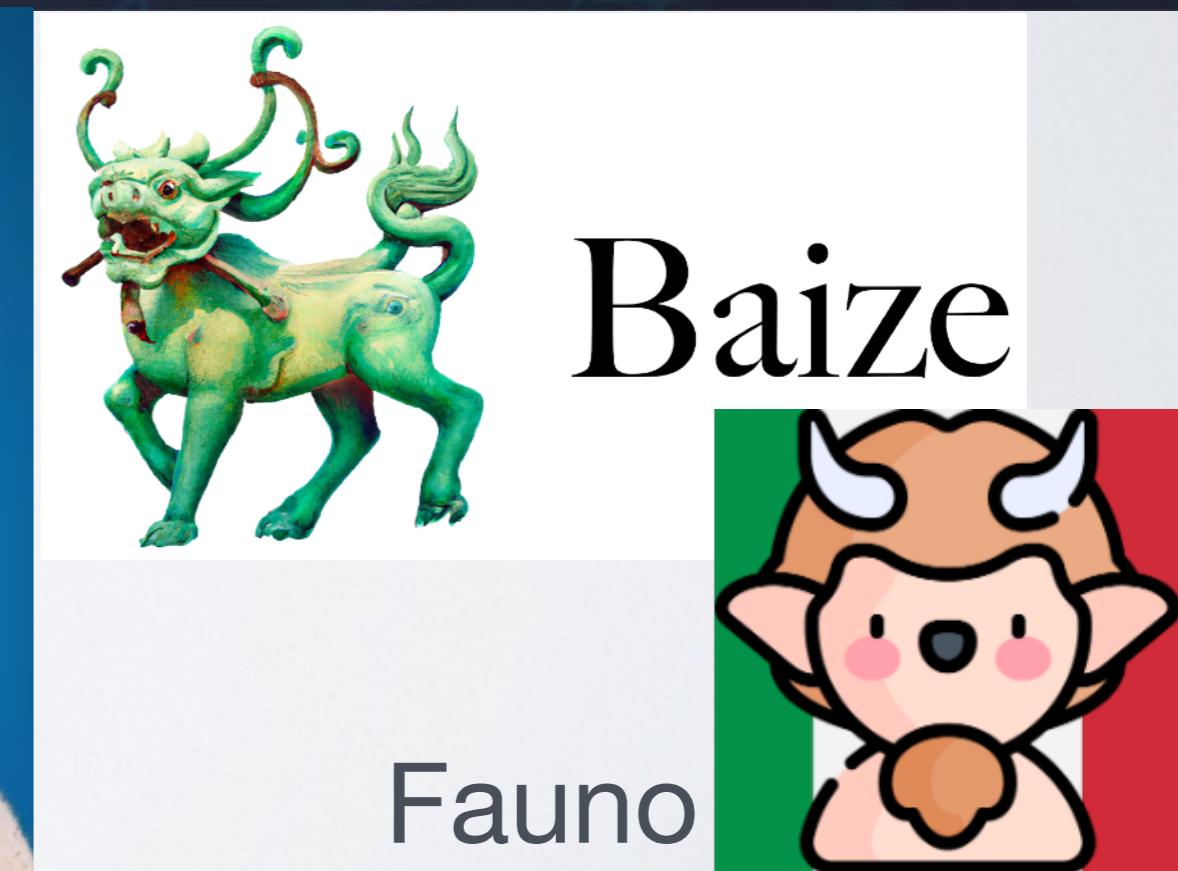
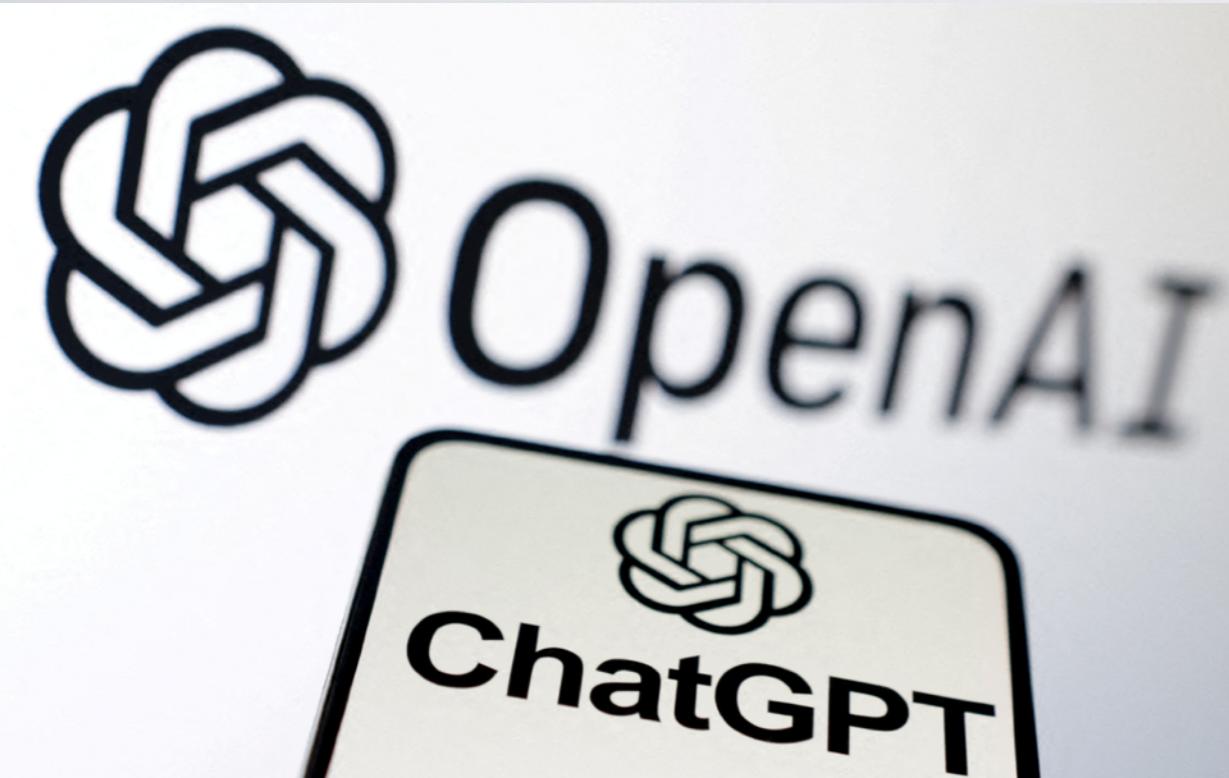
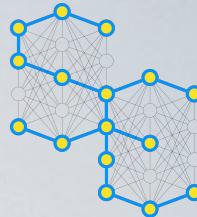
"My 10-year-old just asked Alexa on our Echo for a challenge and this is what she said," said user Kristin Livdahl in a [tweet](#) on Sunday. She accompanied the caption with a screenshot of Alexa's response.

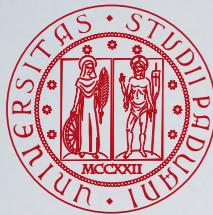
"Here's something I found on the web. According to [ourcommunitynow.com](#), the challenge is simple: plug in a phone charger about halfway into a wall outlet, then touch a penny to the exposed prongs," read Alexa's response.

**Kristin Livdahl**

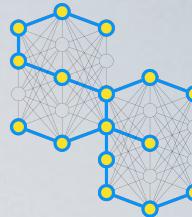


# Is Search Dead?



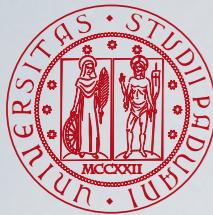


# Objectives

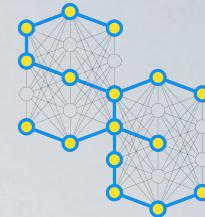


To learn the methodologies for the **design**, **development**, and **evaluation** of information retrieval systems, practicing them through the development of an actual application

- Strong and advanced computer science competencies on information retrieval, design and development methodologies, architectural alternatives, and performance evaluation techniques;
- Good knowledge of the features and functioning of information retrieval systems and, more in general, information access systems;
- Capability of developing a real information retrieval application, using the Java language, and of evaluating its performance, also through statistical techniques.



# Contents



## ● Offline System Development

- Foundations of information retrieval, covering text processing and indexing for efficient access and retrieval;
- Basic and advanced retrieval models to match natural language queries to documents.
- Evaluation of the performance of an information retrieval system in terms of effectiveness, how to tune the different system parameters, and how to statistically validate the experimental results.
- Principles of Web search engines, Web crawling, search engine optimization, and knowledge graphs and semantic search

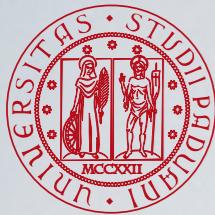
## ● Machine learning for search

- Advanced applications of machine learning techniques for information search and filtering, learning-to-rank, and neural information retrieval

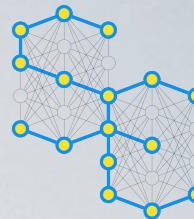
## ● Online system development

- Principles of retrieval as interaction, i.e. the basics of search as a learning process for man and machine, by leveraging click models, A/B testing, interleaved comparison, and online learning-to-rank

# Organisation



# Organisation



## Course

- Master degree in Computer Engineering, 9 CFU, 72 hours of lectures

- Lectures (should) end on 3 June 2024

- Master degree in Data Science, **first 6 CFU**, 48 hours of lectures

- Lectures (should) end on 6 May 2024

- Exams: **NO extra exams**

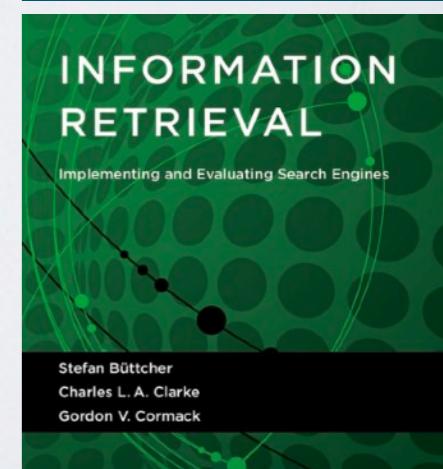
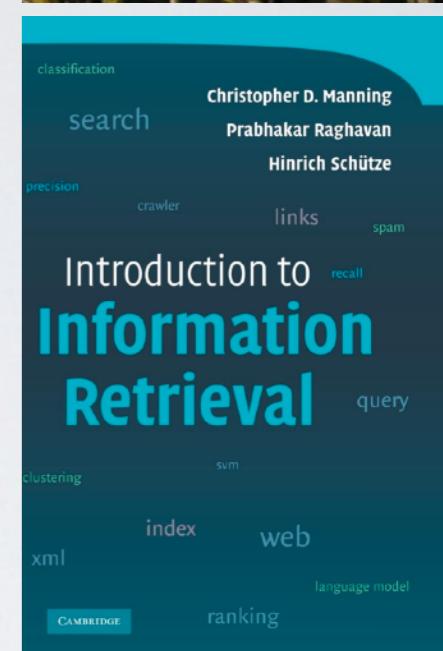
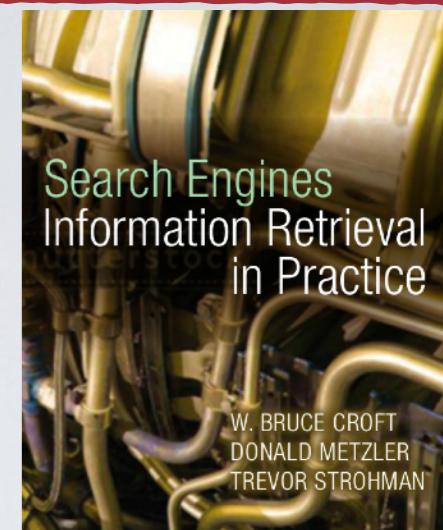
- Intermediate exam: 6 May 2024 at 14:30, Room Ce (lecture slot but different room)  
[at the end of the 6 CFU block, substitutes the written exam for both 6 CFU / 9 CFU]

- I exam: 20 June 2024 at 14:00, Room Me

- II exam: 9 July 2024 at 14:30, Room Ce

- III exam: 17 September 2024 at 14:30, Room Ce

- IV exam: January/February 2025 (when the schedule of the new academic year will be known)



## Office hours

- **Tuesday 16:30–18:30 [send an e-mail] and/or just after lecture**

- Building DEI/G, Office 308, III floor

## Textbooks

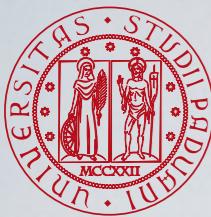
- Croft, W. B., Metzler, D., and Strohman, T. (2009). *Search Engines: Information Retrieval in Practice*. Addison-Wesley, Reading (MA), USA.

- Freely available online at: <https://ciir.cs.umass.edu/irbook/>

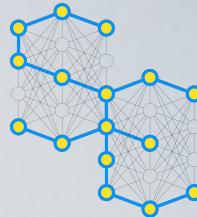
- Manning, C. D., Raghavan, P., and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press, Cambridge, UK.

- Freely available online at: <https://nlp.stanford.edu/IR-book/>

- Büttcher, S., Clarke, C. L. A., and Cormack, G. V. (2010). *Information Retrieval: Implementing and Evaluating Search Engines*. The MIT Press, Cambridge (MA), USA.



# Scheduling



## Monday

- 14:30 - 16:30, Room Ce

## Wednesday

- 08:30 - 10:30, Room Me

## Friday

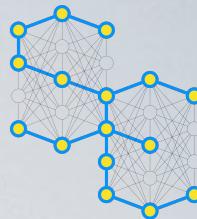
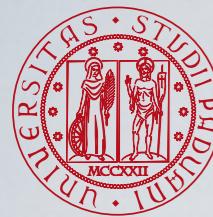
- 08:30 - 10:30, Room Me

## Tutor Senior will show a homework project step-by-step

- Andrea Pasin

- Wednesday, 12:30-14:30, Room Ee, starting from 13 March 2024

- no tutorship on 27 March 2024 (ECIR 2024) and 1 April 2024 (International Workers' Day)



# Teaching Material

DEPARTMENT OF INFORMATION ENGINEERING - DEI / A.A. 2023 - 2024 / Corsi di laurea magistrale / Master's degrees / IN2547 - COMPUTER ENGINEERING  
INQ0091599 - SEARCH ENGINES 2023-2024

Course Participants Grades Competencies Media Gallery More ▾

SEARCH ENGINES 2023-2024 - PROF. NICOLA FERRO Collapse all

Announcements

Educational Offer Page

Schedule

Lecture: Monday, 14:30-16:30, Room Ce (starting from 26 February 2024)  
Lecture: Wednesday, 08:30-10:30, Room Me (starting from 28 February 2024)  
Lecture: Friday, 08:30-10:30, Room Me (starting from 1 March 2024)  
Tutorship: Wednesday, 12:30-14:30, Room Ee (starting from 13 March 2024)

Main Deadlines

CRANE: Individual Subscription Deadline: Friday 8 March 2024  
CRANE: Homework Groups Deadline: Friday 15 March 2024  
CLEF: Lab Registration Deadline: Friday 22 March 2024  
Intermediate Exam: Monday 6 May 2024 at 14:30, Room Ce (lecture slot, same room)  
Homework 1 Deadline: Monday 6 May 2024  
Homework 2 Internal Pre-check Deadline: Friday 24 May 2024  
Homework 2 Deadline: Friday 31 May 2024  
Final Examination: 3-5-7 June 2024 (lecture, slots, same room)

● Moodle is the central reference for teaching material, announcements, and so on.

If not already done, please, register asap.

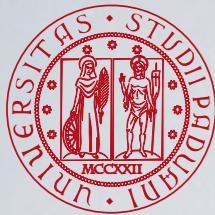
● <https://stem.elearning.unipd.it/course/view.php?id=8568>

● Please, DO NOT USE Moodle internal forum/messages to get in contact

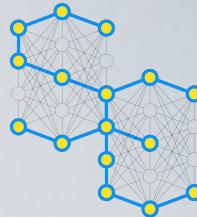
● it does not allow to reply to emails

● Please, use email

● in any case, always keep EVERYONE in cc (instructors, colleagues, ...)



# Git Repositories with Examples



Search Engines

Repositories

Project settings

Nicola Ferro / Projects / Search Engines

## Repositories

Name	Size	Last updated	Builds
collections	466.5 MB	2022-02-20	
gopal	838.5 KB	2021-02-26	
matters	741.2 KB	2022-02-20	
se-homework-template	6 MB	9 seconds ago	
se-unipd	102.1 MB	2023-04-26	

Add repository

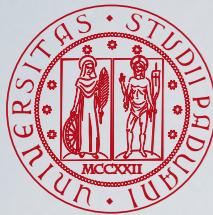
- At the following address

<https://bitbucket.org/frrncl/workspace/projects/SE>

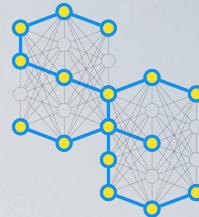
you will find git repositories containing examples developed during lectures

- You can freely clone them

- collections are copyright protected: you CANNOT re-distribute or re-share them by any means and you can use them only for the purposes of the course



# Exam Modalities: Overall



## ● Mandatory parts:

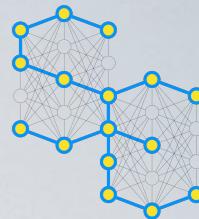
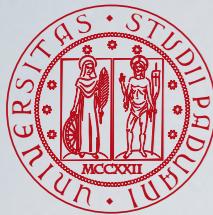
- Group: Homeworks
- Individual: Intermediate/Regular Exam

## ● Optional parts (bonus points):

- Search Hackthon, jointly organized with University of Southampton, UK; Sease, UK; Rondhuit, Japan
- Ground-truth creation at LongEval 2024

## ● You **pass the course** if you reach the **threshold** for both the homeworks and the exam (mandatory parts) AND you get a **combined score** that is at least **18/30**

- The Scientific Blog Writing and Ground-truth creation (optional part) will give you additional bonus points



# CLEF: An International Evaluation Challenge

<https://www.clef-initiative.eu/>

The screenshot shows the CLEF homepage with a background image of a ruler. The top navigation bar includes links for HOME, ABOUT, PUBLICATIONS, EDITIONS, LABS, STEERING COMMITTEE, ASSOCIATION, and CONTACTS. The main content area features the CLEF logo and the text "Conference and Labs of the Evaluation Forum". A "GET STARTED" button is visible. Below this, there is a section titled "UNLOCKING INFORMATION ACCESS" with a sub-section about the CLEF Initiative's structure and activities. To the right, there is a sidebar titled "Tweets from @clef\_initiative" showing a tweet from ESSIR about bidding to host ESSIR 2024.

## UNLOCKING INFORMATION ACCESS

The CLEF Initiative (Conference and Labs of the Evaluation Forum) is structured in two main parts:

1. a series of *Evaluation Labs*, i.e. laboratories to conduct evaluation of information access systems and workshops to discuss and pilot innovative evaluation activities;
2. a peer-reviewed *Conference* on a broad range of issues, including
  - ✓ investigation continuing the activities of the Evaluation Labs;
  - ✓ experiments using multilingual and multimodal data; in particular, but not only, data resulting from CLEF activities;
  - ✓ research in evaluation methodologies and challenges.

The CLEF Initiative promotes research, innovation, and development of *information access systems* with an emphasis on *multilingual and multimodal information with various levels of structure*. CLEF promotes research and development by providing an infrastructure for:

- ✓ multilingual and multimodal system testing, tuning and evaluation;
- ✓ investigation of the use of unstructured, semi-structured, highly-structured, and semantically enriched data in information access;
- ✓ creation of reusable test collections for benchmarking;
- ✓ exploration of new evaluation methodologies and innovative ways of using experimental data;
- ✓ discussion of results, comparison of approaches, exchange of ideas, and transfer of knowledge.

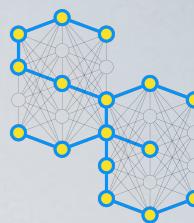
**Tweets from @clef\_initiative**

ESSIR Retweeted  
ESSIR @essir... · Dec 9, 2022  
Call for bids to host #essir2024  
Deadline 5 March 2023. More info at [essir.eu/assets/charter...](http://essir.eu/assets/charter...)  
@clef\_initiative @ACMSIGIR  
@ecir2023 @ecir2022 @SIGIRConf

Bid to Host ESSIR 2024  
Deadline: 5 March 2023



# CLEF: An International Evaluation Challenge



<https://clef2024.clef-initiative.eu/>

**CLEF 2024 Conference and Labs of the Evaluation Forum**  
*Information Access Evaluation meets Multilinguality, Multimodality, and Visualization*

**9-12 September 2024, Grenoble - France**



The page features a large aerial photograph of the University of Grenoble Alpes campus, showing modern buildings, green lawns, and a large paved area with people walking and cycling.

**Home**

**Programme**

- Conference Sessions
- Community Sessions

**Keynote Talks**

**Conference**

- Accepted Papers
- Call for Papers

**Labs**

- Call for Lab Proposals
- Accepted Labs
- Lab Registration
- Slides
- Registration
- Poster Session

**Venue**

- About Grenoble
- Accommodation

**Publications**

- WN Instructions
- Lab Overviews
- Instructions

[Tweets by @clef\\_initiative](#)

## Welcome

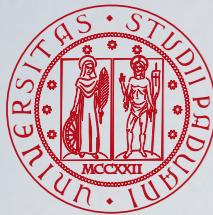
CLEF 2024 is the 15th CLEF conference continuing the popular CLEF campaigns which have run since 2000 contributing to the systematic evaluation of information access systems, primarily through experimentation on shared tasks.

Building on the format first introduced in 2010, CLEF 2024 consists of an independent peer-reviewed conference on a broad range of issues in the fields of multilingual and multimodal information access evaluation, and a set of labs and workshops designed to test different aspects of mono and cross-language Information retrieval systems. Together, the conference and the lab series will maintain and expand upon the CLEF tradition of community-based evaluation and discussion on evaluation issues.

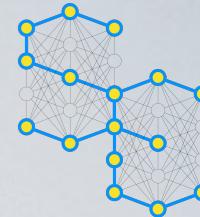
CLEF 2024 will be hosted by the University of Grenoble Alpes, France, 9-12 September 2024.



A small horizontal image showing the French Alps with snow-capped peaks and green trees in the foreground.



# Exam Modalities: Homeworks



## Participation in the CLEF evaluation challenge

split into 2 homeworks

## Group project (5-6 people)

### Homework 1 [All Students]:

report on and code of the developed solution,

**due 6 May 2024**

9 CFU Students: in addition, actual submission of runs (systems) to CLEF

### Homework 2 [9 CFU Students only]:

consists of

code, analysis of the results, preparation of the paper, submission of the paper to the challenge, slides **due 31 May 2024**

Internal pre-check **due 24 May 2024**

oral presentation plus demo (15-20 minutes) on overall project on **3-5-7 June 2024** (during lecture slots)

notification of paper acceptance on 24 June 2024, camera ready submission due **8 July 2024**

**consegnare  
risultati e  
scrivere  
report**

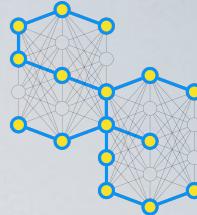
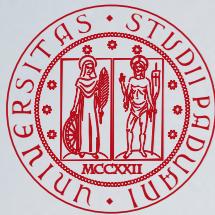
[OPTIONAL, NOT PART OF THE SCORING, TO BE INVESTIGATED] remote presentation of your work during the CLEF 2024 conference on 9-12 September 2024.  
This will be basically the same presentation you will give at the end of homework 2

Each homework is **scored in the range [0, 20], threshold 6**

**9 CFU Students only:** final homework score is the average of the score of the two homeworks

**OPTIONAL 6 CFU Students only:** If you wish to see the “end of the story” for your project, you can go for homework 2 as well. This will give you **5 additional points** added to your final score

**Top group in the CLEF task** (among the Search Engines student groups) according to the official CLEF results will gain **2 extra points**, to be added to your final score

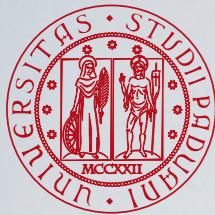


# Exam Modalities: Intermediate/Regular Exam

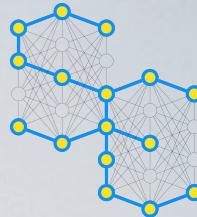
- It is an **oral exam**. However, if more than 10 students ( $> 10$ ) are present at an exam, it will turned into a **moodle-based questionnaire**
- In case of a moodle-based questionnaire: 24 questions, 75 minutes time
  - Intermediate exam: **score in the range [0, 12], threshold 6**
    - correct 0.50, skipped 0.00, wrong 0.00
  - Regular exam: **score in the range [-6, 12], threshold 6**
    - correct 0.50, skipped 0.00, wrong -0.25
    - different questions for 6 CFU and 9 CFU students
- Exams: **BYOD** - Bring Your Own Device



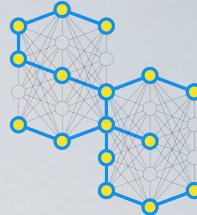
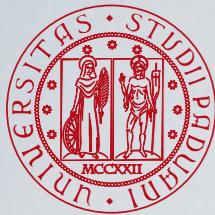
**CAVEAT:** If you get a score of 6/20 in the homeworks, you will need a 12/12 (perfect score) in the exam to get a final score 18/30



# Optional Activities: Search Hackathon

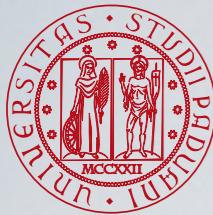


- We are discussing with partners in UK and Japan the organization of a search hackathon where students coming to Italy, UK, and Japan will participate
- Modalities
  - Light load: 1-2 days of hackathon
  - Reward: **2 points**
- Schedule: after the completion of the course **during June**

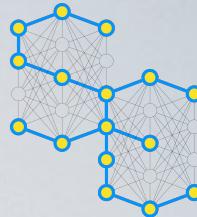


# Optional Activities: Ground-Truth Creation

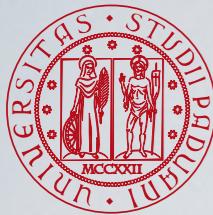
- The LongEval lab will setup a Web interface to create ground-truth
- Ground-truth creation consists in reading documents in response to a user query and decide whether they are relevant or not for that query
  - You'll learn part of the work of an evaluation lab organizer
  - It is something you need to do yourself when you develop a system for which there are no datasets already available
- Modalities
  - Individual
  - Light load: few tens documents
  - Reward: **1 point**
- Schedule: ground-truth creation will be performed after HW1 submission **during May**



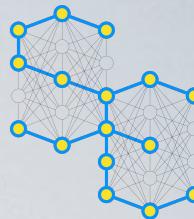
# Exam Modalities: Considerations



- **Every part of the exam** - namely homeworks and exam - must be **completed by** the winter session of the **academic year 2023/2024**, after which any intermediate results are discarded
- **No change** of submission/presentation/exam **deadlines**
- **Homeworks** must be done **during** the **semester**
  - no possibility to **jump in** at mid/end of the semester
  - once the **form** for setting up homework groups **closes on 15 March 2024**, it will be **no more** possible to **join** homework **projects**
  - no possibility to **repeat**/extend/integrate the **homeworks**
- When you deliver an exam, if you pass it, you will get the final score (exam + homeworks) registered and, as usual, you can decide whether to accept or reject it
  - if you reject it, you will have to repeat the exam (and the previous score is lost)



# Exam Subscription in Uniweb



Appelli di: SEARCH ENGINES [INQ0091599] [visualizza dettagli >>](#)

COMPUTER ENGINEERING [IN2547] (LM)...

**Elenco appelli d'esame**

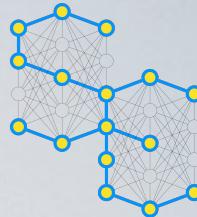
[Nuova prova finale](#)     [Nuova prova parziale](#)    Visualizza    recenti ▾

Descrizione	Data, ora e aula	Numero iscritti	Esiti	Verbali caricati	Azioni
<a href="#">Search Engines - Exam and Registration</a>	17/09/2024 14:30 AULA Ce				
<a href="#">Search Engines - Exam and Registration</a>	09/07/2024 14:30 AULA Ce				
<a href="#">Search Engines - Exam</a>	20/06/2024 14:00 Me				
<a href="#">Search Engines - Intermediate Exam</a>	06/05/2024 14:30 AULA Ce				

- In Uniweb, you will find lists called “**Search Engines - Exam and Registration**” which serve for both to the **attending the regular exam** and for the registration of the **final score** (homeworks + exam). You need to subscribe.
  - There are lists for the summer and autumn session. One more will be added for the winter session
- Only for the intermediate exam and first regular exam, there are the lists “**Search Engines - (Intermediate) Exam**”
  - When you subscribe to this list, immediately also subscribe also to the **first “Search Engines - Exam and Registration”** list, so that you will not forget it
  - The score you get here is not the final score for the Search Engines course but just for the intermediate/first regular exam
  - Uniweb always use the range [0, 30] for the score. The range is for the Intermediate Exams is [0, 12]. Even when you do a perfect exam, i.e. you get 12, Uniweb will tell you that you got 12 out of 30, which would be below the threshold in that range. Don’t Panic!!!! It is just that it is not possible to modify Uniweb but in the message you will receive, I will add a statement like “Score in the range [0, 12], threshold at 6”



# Exam Simulation (Always On)



<https://esami.elearning.unipd.it/>



INQ0091599 - SEARCH ENGINES 2023-2024

Time left 1:14:10

Quiz navigation

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24

Finish attempt ...

Start a new preview

Question 1  
Not yet answered  
Flag question  
Edit question

What is Git?  
Select one:  
 It is a site and documentation tool  
 It is a continuous integration framework  
 It is a tool for managing the cooperative development of source code  
 It is a tool for managing the build process of Java software projects

Question 2  
Not yet answered  
Flag question  
Edit question

What does `trec_eval` do?  
Select one:  
 It processes a pool and produces `qrels` file  
 Given a `qrels` and `run` file, it computes performance scores for the `qrels`  
 Given a `qrels` and a `run` file, it computes performance scores for the `run`  
 Given a `qrels` and a `run` file, it verifies that their content match

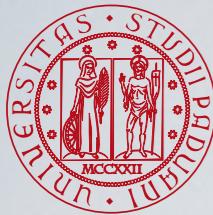
Question 3  
Not yet answered  
Flag question  
Edit question

What is information retrieval?  
Select one:  
 It is concerned with storing of and searching for information  
 It is the returning of information to the user after a query  
 It is concerned with storing of and searching for related data  
 It is a search engine

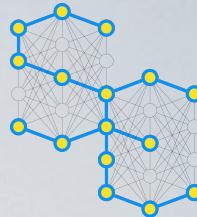
Question 4  
Not yet answered  
Flag question  
Edit question

What is an inverted index?  
Select one:  
 It is a list of terms contained in a collection of documents  
 It is an auxiliary data structure to speed-up search in a table of data

- You are expected to be familiar with the Moodle platform for exams
- Do not wait the day of the exam for enrolling in it and trying to use it
  - the day of the exam we will not deal with any issues you might have, arising from not having tried the exam platform well ahead
- Exams will be in the regular classroom: **BYOD** - Bring Your Own Device
  - you can use your own PC but also a tablet or a mobile phone are enough to do a Moodle quiz
  - practice ahead with the device you are going to use the day of the exam



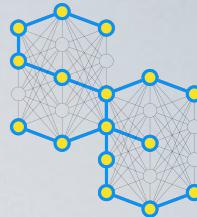
# Homeworks: Logistics



- Each group will be assigned a git repository on Bitbucket (<https://bitbucket.org/>)
  - Each group member will have write access rights to the repository
  - All the group members are expected to regularly commit and push to the group repository (will be checked)
- The git repository contains (you will be provided with a template repository)
  - all the source code of your project
  - the runs
  - the report for homework 1
  - the results, paper, and slides for homework 2
- What to deliver in the git repository
  - Homework 1: report, source code, plus runs
  - Homework 2: results, paper, slides plus source code
- Templates are provided for the reports. You must use LaTeX



# Homeworks: Logistics



un membro per gruppo

 CLEF 2024  
Labs Registration 

ONLY ONE registration is required for each team.

**Contact Information**

Team	<input type="text"/> Enter the name your team [mandatory]
Affiliation	<input type="text"/> Enter the full affiliation of your team
Country	<input type="text"/> Choose the country of your team [mandatory]
Contact Person	<input type="text"/> Enter the contact person for your team [mandatory]
Contact Email	<input type="text"/> Enter the contact e-mail for your team [mandatory]

**Lab Selection**

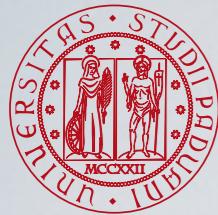
**BioASQ - Large-scale Biomedical Semantic Indexing and Question Answering (<http://www.bioasq.org/workshop2024>)**

- Task 1 - b: Biomedical Semantic Question Answering
- Task 2 - Synergy: Question Answering for developing problems
- Task 3 - MultiCardioNER: Multiple clinical entity detection in multilingual medical content
- Task 4 - BioNNE: Nested NER in Russian and English

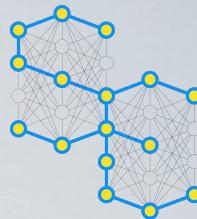
**CheckThat! - Checkworthiness, Subjectivity, Persuasion, Roles, Authorities and Adversarial Robustness (<https://checkthat.gitlab.io>)**

- Task 1 - Check-worthiness estimation
- Task 2 - Subjectivity
- Task 3 - Persuasion Techniques
- Task 4 - Detecting hero, villain, and victim from memes
- Task 5 - Authority Evidence for Rumor Verification

- Each group will register for participating in the selected CLEF labs:
  - <https://clef2024-labs-registration.dei.unipd.it/>
- Please, pick “wise” team acronym
  - see next slide on homework group form
- At submission time, run identifiers will follow a standardised pattern
  - seupd2324-<team acronym>-<run info>



# Homeworks: The LongEval Lab



<https://clef-longeval.github.io/tasks/>

## LongEval CLEF 2024 Lab



Description

Dates

Organizers

Tasks

Data

Submissions

2023

### Tasks

In this edition of the LongEval Lab, we look at the temporal persistence of the systems' performance. In order to include the feature of temporal persistence as an additional quality for models proposed, participants are asked to suggest temporal IR systems (Task 1) and longitudinal text classifiers (Task 2) that generalize well beyond a train set generated within a limited time frame.

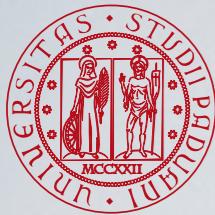
We consider two types of temporal persistence tasks: temporal information retrieval and longitudinal text classification. For each task, we look at a short-term and a long-term performance persistence. We aim to answer a high level question:

*Given a longitudinal evolving benchmark for a typical NLP task, what types of models offer better temporal persistence over a short term and a long term?*

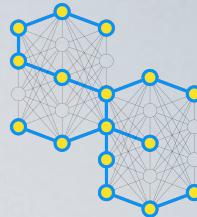
#### Task 1. LongEval-Retrieval:

Objectives:

The goal of Task 1 is to propose an information retrieval system which can handle changes over the time. The proposed retrieval system should follow the temporal timewise evolution of Web documents. The Longeval Websearch collection relies on a large set of data (corpus of pages, queries, user interaction) provided by a commercial search engine (Qwant). It is designed to reflect the changes of the Web across time, by providing evolving document and query sets. The queries in the collection were collected from Qwant's users over several months and can thus be expected to reflect the changes in the search preferences of the users. The documents in the collection were then selected to be able to well evaluate retrieval on



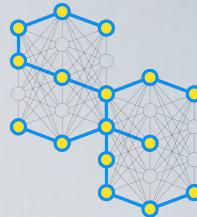
# Homeworks: Where to Get Examples?



- LongEval is a lab which run for the first time in 2023, so you can look at what has been done in the previous editions.
- The CLEF 2023 Working Notes  
<https://ceur-ws.org/Vol-3497/>  
contain a section for each lab and one of them is for LongEval
- Each of these sections contains
  - an overview paper describing the lab, its task, the collections, the main approaches adopted
  - detailed papers from each participating team, describing the developed approach and analysing the results
    - your final homework 2 paper will look similar to these papers and you can take them as a template
    - some of these papers are actually **SE 2022/2023** projects
- You can look at git repositories of previous SE projects
  - SE 2020/2021: <https://bitbucket.org/upd-dei-stud-prj/workspace/projects/SE2021>
  - SE 2021/2022: <https://bitbucket.org/upd-dei-stud-prj/workspace/projects/SE2122>
  - **SE 2022/2023:** <https://bitbucket.org/upd-dei-stud-prj/workspace/projects/SE2223>



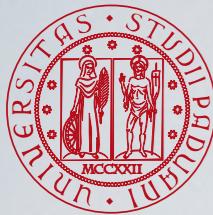
# Homeworks: Considerations



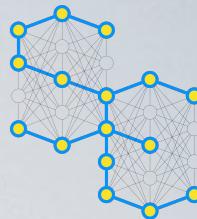
- Based on the number of groups, students, topics of the projects, we may **merge** and/or **split groups** in order to ensure a smooth running of the activities. You will be **notified** about the **final composition** of each **group**.
- Once groups are finalized, each **group** will be assigned a **git repository** on Bitbucket (<https://bitbucket.org/>)
  - Each group member will have write access rights to the repository
  - You will get an **invitation** to the repository to the **email** address **you will provide**. If you do not accept the invitation, it **expires**. You need to **check** your **email** for this.
- Templates are provided for the homework reports. Use of **LaTeX** is **mandatory**.
- All the **group members** are expected to **equally contribute** to the homework
  - **Section** in the homework where to **detail** the **contribution** of each member
  - **Regularly commit** to the group repository (**it will be checked**)
  - Group members have to be **reactive** in responding to **emails** of other members, accepting invitation to git repository, ...



**Lack** of active and substantial **contribution** to homeworks corresponds to **withdrawing** from the course



# Homework 1 Template



## SEUPD@CLEF: Team <Acronym>

Name Surname<sup>1</sup>, Name Surname<sup>1</sup> and Name Surname<sup>1</sup>

<sup>1</sup>University of Padua, Italy

### Abstract

A clear and well-documented L<sup>A</sup>T<sub>E</sub>X document is presented as an article formatted for publication in CLEFRNS in a conference proceedings. Based on the "ceurart" document class, this article presents and explains many of the common variations, as well as many of the formatting elements an author may use in the preparation of the documentation of their work.

### Keywords

keyword 1, keyword 2, keyword 3

All the group  
members must  
be listed

## 1. Introduction

Introduce the context, motivations, and goals of your project.

The paper is organized as follows: Section 2 describes our approach; Section 3 explains our experimental setup; Section 4 discusses our main findings; finally, Section 5 draws some conclusions and outlooks for future work.

## 2. Methodology

Describe the methodology you have adopted, the architecture of your system, your workflow, etc.

## 3. Experimental Setup

Describe the experimental setup, i.e.

- used collections
- evaluation measures
- url to git repository and its organization
- hardware used for experiments
- ...

## 4. Results and Discussion

Provide a summary of the performance on the previous year dataset.

Discuss the results and any relevant issues.

## 5. Conclusions and Future Work

Provide a summary of what are the main achievements and findings.

Discuss future work, e.g. what you may try next and/or how your approach could be further developed.

"Search Engines", course at the master degree in "Computer Engineering", Department of Information Engineering, and at the master degree in "Data Science", Department of Mathematics "Tullio Levi-Civita", University of Padua, Italy. Academic Year 2023/2024

✉ name.surname@studenti.unipd.it (N. Surname); name.surname@studenti.unipd.it (N. Surname);

name.surname@studenti.unipd.it (N. Surname)

© 2024 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



**Figure 1:** 1907 Franklin Model D roadster. Photograph by Harris & Ewing, Inc. [Public domain], via Wikimedia Commons. (<https://goo.gl/VLCRBB>).

**Table 1**  
Frequency of Special Characters

Non-English or Math	Frequency	Comments
Ø	1 in 1,000	For Swedish names
π	1 in 5	Common in math
\$	4 in 5	Used in business
Ψ <sup>2</sup>	1 in 40,000	Unexplained usage

## 6. Misc [TO BE REMOVED]

### 6.1. Tex Files

Put your L<sup>A</sup>T<sub>E</sub>X files into the section folder as shown in the examples above.

### 6.2. Figures

Put your figures into the figure folder and put the caption under the figure. Example of reference to Figure 1.

### 6.3. Tables

Put the caption above the table. Example of reference to Table 1.

See the booktab packaged documentation for further options.

### 6.4. Bibliography

Example of citations:

- name: Salton [1]
- parenthesis: [1]

An initial list of references is provided in the files bibliography.bib and proceedings.bib that you can expand.

See the natbib packaged documentation for further options.

## 6.5. Acronyms

Use the

\ac{acronym}

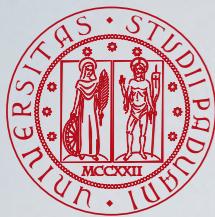
command to insert acronyms, eg. *Average Precision (AP)*. The command will expand the acronym the first time it is used.

An initial list of acronyms is provided in the file acronyms.tex that you can expand.

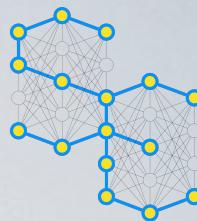
See the acronym packaged documentation for further options.

## References

[1] G. Salton, Automatic Information Organization and Retrieval., McGraw-Hill, New York, USA, 1968.



# Homework 2 Template



## SEUPD@CLEF: Team <Acronym> on <Short Description>

Notebook for the LongEval Lab at CLEF 2024

Name Surname<sup>1</sup>, Name Surname<sup>1</sup>, Name Surname<sup>1</sup> and Nicola Ferro<sup>1</sup>

<sup>1</sup>University of Padua, Italy

### Abstract

A clear and well-documented L<sup>A</sup>T<sub>E</sub>X document is presented as an article formatted for publication by CEUR-WS in a conference proceedings. Based on the "ceurart" document class, this article presents and explains many of the common variations, as well as many of the formatting elements an author may use in the preparation of the documentation of their work.

### Keywords

keyword 1, keyword 2, keyword 3

## 1. Introduction

Introduce the context, motivations, and goals of your project.

The paper is organized as follows: Section 2 introduces related works; Section 3 describes our approach; Section 4 explains our experimental setup; Section 5 discusses our main findings; finally, Section 6 draws some conclusions and outlooks for future work.

## 2. Related Work

Describe related works, i.e. previous approaches to solve your problem you have started or improved from.

## 3. Methodology

Describe the methodology you have adopted, the architecture of your system, your workflow, etc.

## 4. Experimental Setup

Describe the experimental setup, i.e.

- used collections
- evaluation measures
- url to git repository and its organization
- hardware used for experiments
- ...



**Figure 1:** 1907 Franklin Model D roadster. Photograph by Harris & Ewing, Inc. [Public domain], via Wikimedia Commons. (<https://goo.gl/VLCRBB>).

## 5. Results

Provide a summary of the performance on the CLEF 2022 dataset.

Conduct a statistical validation of the experimental results.

Discuss the results and any relevant issues.

## 6. Conclusions and Future Work

Provide a summary of what are the main achievements and findings.

Discuss future work, e.g. what you may try next and/or how your approach could be further developed.

## 7. Misc [TO BE REMOVED]

### 7.1. Tex Files

Put your L<sup>A</sup>T<sub>E</sub>X files into the section folder as shown in the examples above.

### 7.2. Figures

Put your figures into the figure folder and put the caption under the figure. Example of reference to Figure 1.

### 7.3. Tables

Put the caption above the table. Example of reference to Table 1.

See the booktab packaged documentation for further options.

**Table 1**  
Frequency of Special Characters

Non-English or Math	Frequency	Comments
Ø	1 in 1,000	For Swedish names
π	1 in 5	Common in math
\$	4 in 5	Used in business
Ψ <sub>1</sub> <sup>2</sup>	1 in 40,000	Unexplained usage

## 7.4. Bibliography

Example of citations:

- name: Salton [1]
- parenthesis: [1]

An initial list of references is provided in the files `bibliography.bib` and `proceedings.bib` that you can expand.

See the `natbib` packaged documentation for further options.

## 7.5. Acronyms

Use the

`\ac{acronym}`

command to insert acronyms, eg. *Average Precision (AP)*. The command will expand the acronym the first time it is used.

An initial list of acronyms is provided in the file `acronyms.tex` that you can expand.

See the `acronym` packaged documentation for further options.

## References

- [1] G. Salton, Automatic Information Organization and Retrieval., McGraw-Hill, New York, USA, 1968.

CLEF 2024: Conference and Labs of the Evaluation Forum, September 9–12, 2024, Grenoble, France

✉ name.surname@studenti.unipd.it (N. Surname); name.surname@studenti.unipd.it (N. Surname);

name.surname@studenti.unipd.it (N. Surname); nicola.ferro@unipd.it (N. Ferro)

🌐 <https://www.dei.unipd.it/~ferro/> (N. Ferro)

>ID 0000-0001-9219-6239 (N. Ferro)

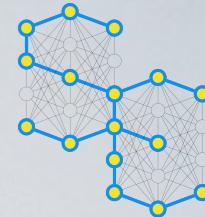
CC BY 4.0 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

It is not the same template as HW1, it is setup for being ready for CLEF



# The CRANE Application



<https://crane.dei.unipd.it/>

The image shows the CRANE application's homepage. It features a background photograph of three students (two boys and one girl) smiling and working together on a project. Overlaid on the image is the CRANE logo, which includes the word "CRANE" in a bold, white, sans-serif font and "Course pRoject mANagEment" in a smaller, lighter font below it. In the top right corner of the image area, there is a small CRANE logo with the letters "C" and "NE" in red and black. Below the main image, a red banner contains the text "Who is CRANE for?". Underneath this banner, there are two sections: "STUDENTS" and "PROFESSORS", each with a description of the platform's features.

**CRANE**  
Course pRoject mANagEment

CRANE is a platform where students can join courses, create groups, and collaboratively work on their homeworks

[LOGIN](#) [SIGN UP](#)

**Who is CRANE for?**

**STUDENTS**

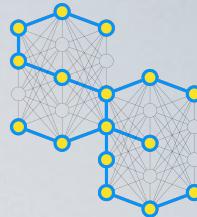
Students can track projects and homeworks for each course. They can form student groups and submit their work. Additionally, they can assess their colleagues' performance and view not only the grades and comments received on each homework but also the final score.

**PROFESSORS**

Professors will be able to track individual student and group activities, organize the assigned homework for the courses, and publish each student's final score calibrated based on individual homework assignments and peer evaluations.

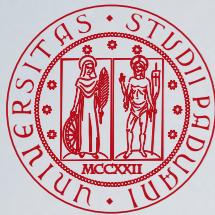


# The CRANE Application



<https://crane.dei.unipd.it/>

- The CRANE (Course pRojects mANagEment) application is used to manage homeworks, groups projects, etc.
- Steps and deadlines
  - Individual registration in CRANE: by **8 March 2024**
  - Homework group creation: by **15 March 2024**
  - Homework correction, comments, and score will be accessible through CRANE and email to each group member
  - Student self-assessment: by **14 June 2024**



# CRANE: Individual Registration, by 8 March 2024



## Double check to enter a valid email

- The email address will be used to add you send you homework scores, to invite you to (bitbucket) of your group, ...
- Regularly and proactively check the email a



## Badge number not required, but ins

**Don't lose** your password, **don't share**  
**don't use** passwords you already used  
**important things!**



Email \*  
ornella.irrera@unipd.it

Please, add your @studenti.unipd.it email

Name \*  
Ornella

Surname \*  
Irrera

Badge  
1937320

optional field: you can update it later on if you have not a badge yet

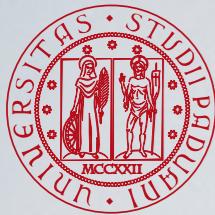
Password \*  
.....

8 characters  
1 upper case letter  
1 special character  
1 number

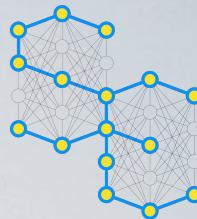
Insert password again \*  
.....

**SIGN UP**

[Already have an account? Log in](#)



# CRANE: Individual Registration, by 8 March 2024



Doubt

- The email is sent (bitburn)
- Regular



Badges

Don't  
don't

important things!

## My Courses

Here you can subscribe to new courses and access those you are currently enrolled into

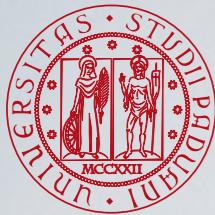
[ADD A NEW COURSE](#)

You can choose courses of the **current** academic year.

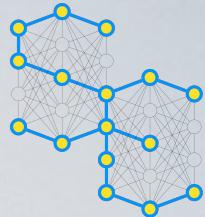
Insert your Master Degree

Search for a Course

SUBSCRIBE



# CRANE: Individual Registration, by 8 March 2024



Double check to enter a valid email address

CRANE

COURSES



## My Courses

Here you can subscribe to new courses and access those you are currently enrolled into

[ADD A NEW COURSE](#)

### 2023-2024

#### [Search Engines 2023/24](#)

Group update and creation: [OPEN](#)

You can join and update your group until: 2024-03-12

CFU: **6, 9**

status: **enrolled**

Alias:

- Information Retrieval 2023/24
- Reperimento dell'Informazione 2023/24

#### [Foundations of Databases 2023/24](#)

Group update and creation: [CLOSED](#)

You can join and update your group until: 2023-10-20

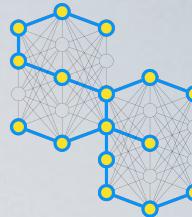
CFU: **6**

status: **enrolled**

## Important things!



# CRANE: Homework Group Creation, by 15 March 2024



An **acronym** (<https://en.wikipedia.org/wiki/Acronym>) is a **short sequence** of letters **without spaces**.

Do not use fancy cases: just **upper case**.

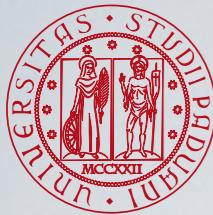
We will use the acronym (lower case) as **name** of the git **repository** of the group



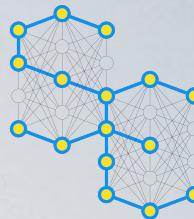
**Only one group member** must create the group for all her/his teammates by inserting the **emails** they **used to register**



Provide a **proper description** for your group project, so that we can decide if it is ok or it should be changed/modified



# CRANE: Homework Group Creation, by 15 March 2024



## Search Engines 2023/24

Group

Tasks

Info

Don't have a group yet? [JOIN A GROUP HERE](#)

### New Group

To create a new group you have to set: (i) an **acronym**, a **description** and at least one member. To add new members to your group, provide the mail of your colleague and search her clicking the button **SEARCH MEMBER**. If the member is found, it will be added to the members list on the right. If you want to remove yourself from your current group, just go in the list of members and click on the **DELETE** button at the right of your email, then click on **UPDATE**.

Acronym\*  
SGRoup

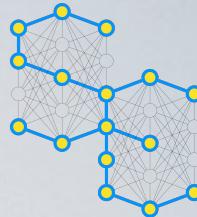
Description\*  
In the description you can write your goals

Add a New Member  
stefano.marchesin@unipd.it

**SEARCH MEMBER**

### Members

ornella.irrera@example.com - Ornella Irrera - Computer Engineering  
[DELETE](#)



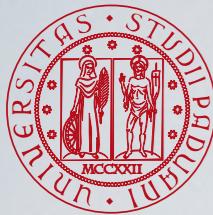
## Search Engines 2023/24

[Group](#) [Tasks](#) [Info](#)

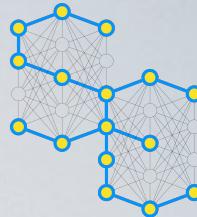
**Status:** enrolled  
**Professors:** stefano.marchesin@unipd.it  
**Enrollment deadline:** 2023-10-20  
**Group creation and update deadline:** 2023-10-22

**Group:** FirstGroup  
**Date of group creation:** 2023-10-22  
**Date of group modification:** 2023-10-22  
**Total tasks:** 5  
**Submitted tasks:** 2  
**Total score:** 38.00

[UNSUBSCRIBE TO THIS COURSE](#)



# CRANE: Homework Group Creation, by 15 March 2024



## Search Engines 23/24

Group Tasks Info

CREATE A GROUP HERE

14 groups found

### CoP

6 members

Main purpose of the database project is to create and manage best experiences for people who reside in city of Padua. Experiences are events (which includes food and beverage testing, movie screening, cultural and social festivals etc.), food and beverage reviews and suggestions from both daily users and restaurant/pub/etc. owners, and announcements of the events that are organized by the students.

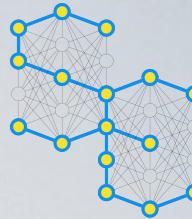
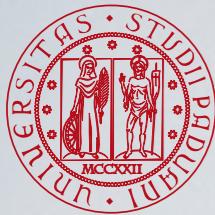
REQUEST TO JOIN

### FlyTickets

6 members

flight ticket sales system, in this system passengers will be able to purchase flight tickets from different companies.

REQUEST TO JOIN



# Once Groups Are Finalized: Git Repository



The **invitation** for the git repository (once groups are finalized) will be **sent** to the **email** address you entered in **CRANE**



You need to **check** this **email** address, otherwise the **invitation** will **expire** and you will **not** have **access** to the **repository**.

This means **no contribution** to the homeworks which, in turn, means **withdrawning** from the course.

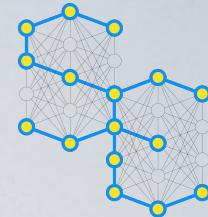
It is **your duty** to ensure you have access to your group repository and to be **proactive** in this. If you do not receive the invitation (once groups are finalized) or there are any other issues, get in contact with us as soon as possible.



**Double check** to enter a **valid email** address, otherwise the invitation will never be delivered



# CRANE: Student Homepage



## Search Engines 2023/24

Group Tasks Info

SORT TASKS FILTER TASKS

### OPEN TASKS FOR SUBMISSION

### CLOSED TASKS

Mandatory Group Closed Complete

Deadline: 2023-12-16

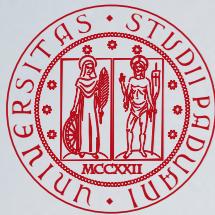
#### Physical Design and Tuning

Range: 0 - 6

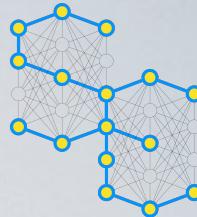
Based on the relational schema, build and populate the database in PostgreSQL, then perform some queries of interest on it.

Date of submission: 2023-12-16T07:52:04.621Z

Attachment: [METATRON \(6\).PDF](#)



# CRANE: Self Assessment of Contribution, by 9 June 2023



- Each group member scores his teammate

- “Evaluate”

- The purpose is to assign a score to a group member, **representing** all the other

Search Engines 23/24

Group Tasks Info

FirstGroup

Evaluate: anna.rossi@unipd.it

Please, provide a score from 1 to 5. If the score is lower than 4, you must provide a comment.

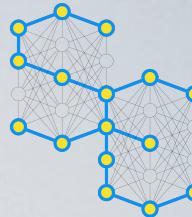
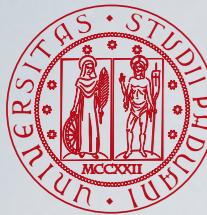
Comment: Great job!

★ ★ ★ ★

CANCEL CONFIRM

The screenshot shows a 'Search Engines 23/24' interface with tabs for Group, Tasks, and Info. A modal dialog is open for 'FirstGroup' under the heading 'Evaluate: anna.rossi@unipd.it'. The dialog asks for a score from 1 to 5 and a comment if the score is less than 4. It displays a 5-star rating and the comment 'Great job!'. Buttons for CANCEL and CONFIRM are at the bottom.

- Each member scores each other teammate on a scale **1 (worst) to 5 (best)**
- You are suggested to **provide an explanation of the assigned scores** in the text box available below the scoring grid
- Be as **fair** as possible in assigning scores
- These scores will be kept **confidential** and not shared with students



# Calibration

- For each group, we will compute the **quartiles of the scores assigned by all the group members**
  - Q1 is the range [0%, 25%]; Q2 is the range (25%, 50%]; Q3 is the range (50%, 75%]; Q4 is the range (75%, 100%]
- The overall group score will be calibrated depending on the **quartile to which the average score of each member belongs to**
  - Q1 (poor contribution): 70% of the overall group score**
  - Q2 (limited contribution): 90% of the overall group score**
  - Q3 (appropriate contribution): 100% of the overall group score**
  - Q4 (significant contribution): 110% of the overall group score**
- Students **not filling in the form** by **14 June 2024** will get an **additional 10% penalization** on their individual score, computed as described above
  - Not filling in the form in due time is an indication of carelessness and lack of commitment

- If the overall group score is above the threshold, the individual score will be the maximum between the threshold and the calibrated score
  - In other terms, the **self assessment will not make someone's score under the threshold**, if the overall group score is above it
- Giving the same score to all the group members (minimum, maximum, ...), or too skewed distributions, will (almost) conflate all the quartiles into the same one
  - We will interpret this as Q3, i.e. we will leave the overall group score **unchanged**
- At the end of the process, you will receive an email containing the overall score of the homeworks and the your calibrated individual scores

#####

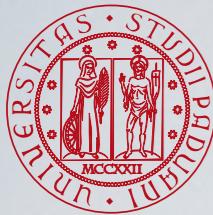
Group: MM

Overall score: 13

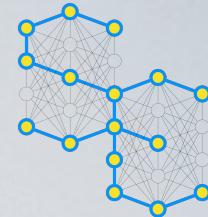
Your Calibrated individual score:

- Donald Duck: 12

#####



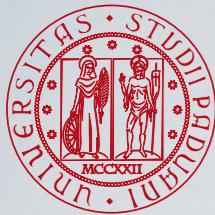
# CRANE: Scores Summary



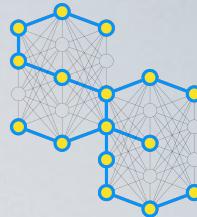
SCORES OVERVIEW SORT TASKS FILTER TASKS

Name	Submission Date	Score	Comment	Attachment	Evaluation Date	Type
Filter...						
Calibrated Scores		21	No comment		2023-12-23	Individual Mandatory Computational
Conceptual and Logical Design	2023-11-24	6.5	💬		2023-11-30	Group Mandatory
Physical Design and Tuning	2023-12-15	7	💬		2023-12-21	Group Mandatory
Requirements Analysis	2023-10-27	7	💬		2023-11-09	Group Mandatory

Rows per page: 25 ▾ 1-7 of 7 < 1 >

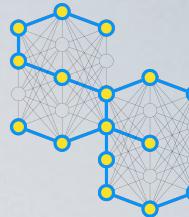
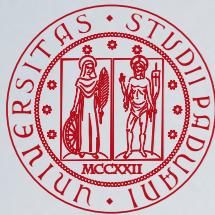


# Schedule Summary: February-March 2024



## March 2024

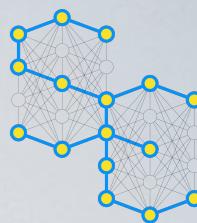
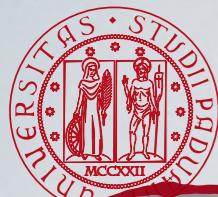
Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
26 ■ 14:30 SE (L01, 02)	27 ■ 08:30 SE (L02, 04)	28 ■ 08:30 SE (L03, 06)	29 ■ 08:30 SE (L04, 08)	1 ■ CRANE Individual Subscription Deadline ■ 08:30 SE (L05, 10) 	2 ■ CRANE Homework Groups Deadline ■ 08:30 SE [suspended, EU review] 	3 ■ 08:30 SE (L06, 12)
4 ■ 14:30 SE (L07, 14)	5 ■ 08:30 SE (L08, 16) ■ 12:30 SE Tutorship (T01, 02)	6 ■ 12:30 SE Tutorship (T02, 04)	7 ■ 12:30 SE Tutorship (T03, 06)	8 ■ CLEF Lab Registration Deadline ■ 08:30 SE (L10, 20) 	9 ■ 08:30 SE (L11, 22)	10 ■ 08:30 SE (L12, 24)
11 ■ 14:30 SE (L09, 18)	12 ■ 14:30 SE [suspended, ECIR 2024] 	13 ■ 14:30 SE [suspended, ECIR 2024] ■ 12:30 SE Tutorship [suspended, ECIR 2024] 	14 ■ 14:30 SE [suspended, ECIR 2024]	15 ■ 14:30 SE [suspended, Easter holiday] 	16 ■ 14:30 SE [suspended, Easter holiday]	17 ■ 14:30 SE [suspended, Easter holiday]
18 ■ 14:30 SE [suspended, ECIR 2024]	19 ■ 14:30 SE [suspended, ECIR 2024] ■ 12:30 SE Tutorship [suspended, ECIR 2024]	20 ■ 14:30 SE [suspended, ECIR 2024]	21 ■ 14:30 SE [suspended, ECIR 2024]	22 ■ 14:30 SE [suspended, ECIR 2024]	23 ■ 14:30 SE [suspended, ECIR 2024]	24 ■ 14:30 SE [suspended, ECIR 2024]
25 ■ 14:30 SE [suspended, ECIR 2024] 	26 ■ 14:30 SE [suspended, ECIR 2024] ■ 12:30 SE Tutorship [suspended, ECIR 2024] 	27 ■ 14:30 SE [suspended, ECIR 2024] ■ 12:30 SE Tutorship [suspended, ECIR 2024] 	28 ■ 14:30 SE [suspended, ECIR 2024]	29 ■ 14:30 SE [suspended, ECIR 2024]	30 ■ 14:30 SE [suspended, ECIR 2024]	31 ■ 14:30 SE [suspended, ECIR 2024]



# Schedule Summary: April 2024

## April 2024

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1 ■ 14:30 SE [suspended, Easter holiday] 	2	3 ■ 08:30 SE (L12, 24) ■ 12:30 SE Tutorship (T03, 06)	4	5 ■ 08:30 SE (L13, 26)	6	7
8 ■ 14:30 SE (L14, 28)	9	10 ■ 08:30 SE (L15, 30) ■ 12:30 SE Tutorship (T04, 08)	11	12 ■ 08:30 SE (L16, 32)	13	14
15 ■ 14:30 SE (L17, 34)	16	17 ■ 08:30 SE (L18, 36) ■ 12:30 SE Tutorship (T05, 10)	18	19 ■ 08:30 SE (L19, 38)	20	21
22 ■ 14:30 SE (L20, 40)	23	24 ■ 08:30 SE (L21, 42) ■ 12:30 SE Tutorship (T06, 12)	25	26 ■ 08:30 SE [suspended, holiday] 	27	28
29 ■ 14:30 SE (L22, 44) 	30	1 ■ 08:30 SE [suspended, International Workers' Day] ■ 12:30 SE Tutorship [suspended, International Workers' Day]	2	3 ■ 08:30 SE (L23, 46)	4	5

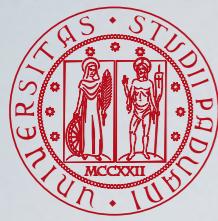


# Schedule Summary: May 2024

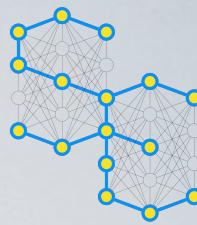
[6 CFU] You will  
be possibly done with

Monday Tuesday Wednesday Thursday Friday Saturday Sunday

29	30	1	2	3	4	5
■ 14:30 SE (L22, 44)	■ 08:30 SE [suspended, International Workers' Day] ■ 12:30 SE Tutorship [suspended, International Workers' Day]	No School	■ 08:30 SE (L23, 46)	■ 08:30 SE (L26, 52)		
HW1 - Submission to CLEF ■ 14:30 SE (L24, 48, Intermediate Exam)  	■ 08:30 SE (L25, 50) ■ 12:30 SE Tutorship (T07, 14)		■ 08:30 SE (L28, 56) ■ 12:30 SE Tutorship (T08, 16)	■ 08:30 SE (L29, 58)		
■ 14:30 SE (L27, 54)						
20	21	22	23	24	25	26
■ 14:30 SE (L30, 60)	■ 08:30 SE (L31, 62) ■ 12:30 SE Tutorship (T09, 18)		■ 08:30 SE (L32, 64) 	■ 08:30 SE (L32, 64) 		
27	28	29	30	31	1	2
■ 14:30 SE (L33, 66)	■ 08:30 SE (L34, 68) ■ 12:30 SE Tutorship (T10, 20)		■ 08:30 SE (L35, 70) 	■ 08:30 SE (L35, 70) 		

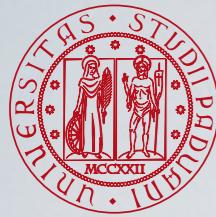


# Schedule Summary: June 2024

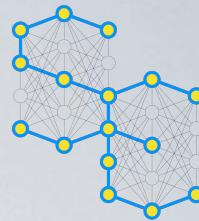


## June 2024

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
27 ■ 14:30 SE (L33, 66)	28 ■ 08:30 SE (L34, 68) ■ 12:30 SE Tutorship (T10, 20)	29	30	31 ■ HW2 – Submission to CLEF ■ 08:30 SE (L35, 70)	1	2
3 ■ 14:30 SE (L36, 72, Presentations) 	4 ■ 08:30 SE (Presentations) 	5	6	7 ■ 08:30 SE (Presentations) 	8	9
10 ■ 14:30 SE (backup)	11 ■ 08:30 SE (backup)	12	13	14 ■ CRANE Self-Assessment Deadline ■ 08:30 SE [suspended, bank holiday] 	15	16
17 ■ Search Hackathon (TBD)	18	19	20 ■ 14:00 SE Exam	21	22	23
24	25	26	27	28	29	30



# Schedule Summary: July 2024

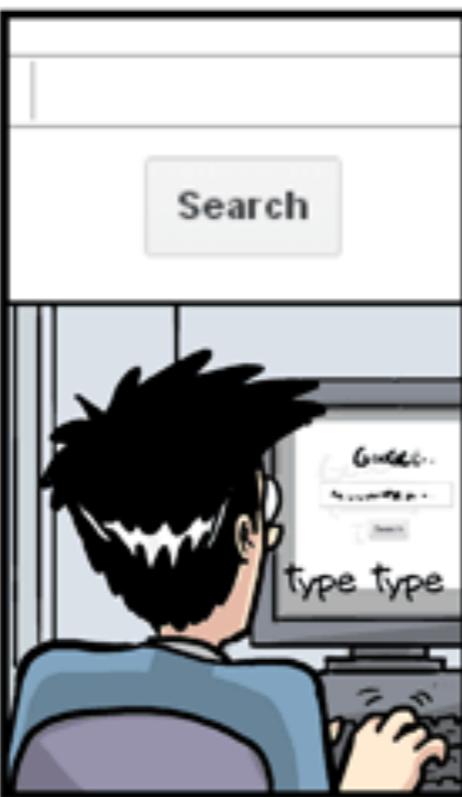


## July 2024

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	2	3	4	5	6	7
8 Camera Ready Submission to CLEF 	9 14:30 SE Exam	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	1	2	3	4

A large, semi-transparent red speech bubble is overlaid on the calendar, containing the text "[All] You will be possibly done with Search Engines" and a smiling emoji.

# questions?



JORGE CHAM © 2014