

Information Service Engineering

Lecture 1: Information, Natural Language, and the Web



Karlsruher Institut für Technologie



FIZ Karlsruhe

Leibniz Institute for Information Infrastructure

Prof. Dr. Harald Sack

FIZ Karlsruhe - Leibniz Institute for Information Infrastructure

AIFB - Karlsruhe Institute of Technology

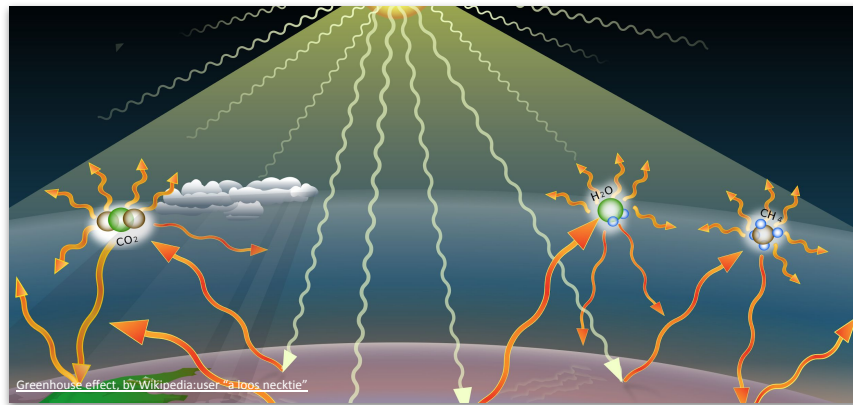
Summer Semester 2021

- 1.1 How to get Information (from the Web)?**
- 1.2 Communication, Language, and Understanding
- 1.3 How to measure Information?
- 1.4 The ever-growing Web of Information
- 1.5 Search Engines on the Web
- 1.6 The Meaning of Information

How to get Information (from the Web)?

Find a solution for the following question:

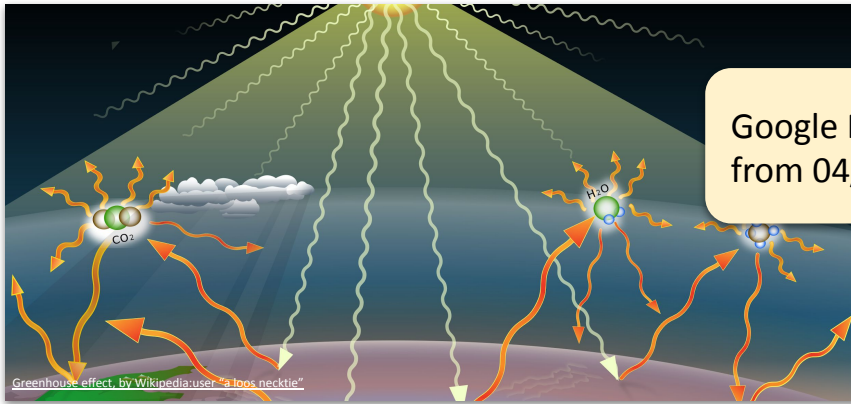
- **When was the Greenhouse Effect discovered?**



Ask an
Information
Service



Find a solution for the following question:

- **When was the Greenhouse Effect discovered?**



Google Result
from 04/2020



When was the Greenhouse effect discovered?  

[All](#) [Images](#) [News](#) [Videos](#) [Maps](#) [More](#) [Settings](#) [Tools](#)

About 14.000.000 results (0,50 seconds)

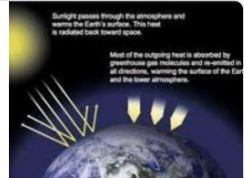
1896

The greenhouse effect was discovered more than 100 years ago

In **1896**, the world renowned Swedish scientist and Nobel Prize Winner Svante Arrhenius (**1859-1927**), described how CO₂ influences the climate. Jun 5, 2018

phys.org › Earth › Environment

[80 years since the first calculations showed that the Earth was ...](#)




[About Featured Snippets](#) [Feedback](#)

People also ask

Who first discovered greenhouse effect? 

What is greenhouse effect and global warming? 

What is the cause of greenhouse effect? 

[Feedback](#)

en.wikipedia.org › wiki › Greenhouse_effect ▾

[Greenhouse effect - Wikipedia](#)

The **greenhouse effect** is the process by which radiation from a planet's atmosphere warms the planet's surface to a temperature above what it would be without ...

[Runaway greenhouse effect](#) · [Anti-greenhouse effect](#) · [History of climate change](#)

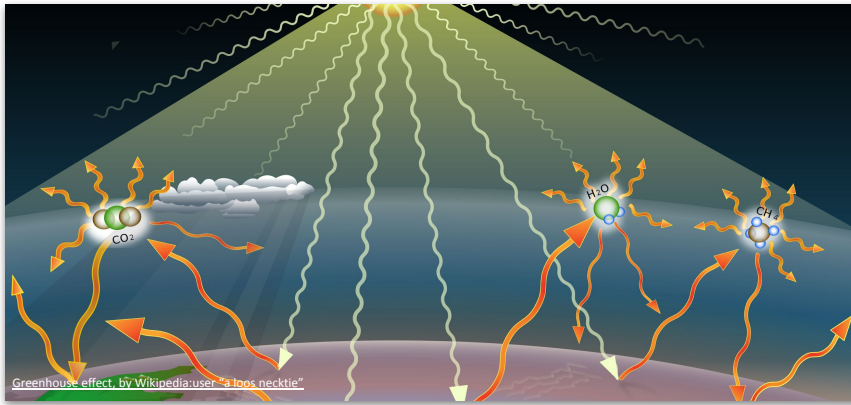
www.lenntech.com › greenhouse-effect › global-warming-history ▾

[History of the greenhouse effect and global warming - Lenntech](#)

History of the **greenhouse effect** and global warming. ... It was also discovered that water vapor absorbed totally different types of radiation than carbon dioxide.

Find a solution for the following question:

- **When was the Greenhouse Effect discovered?**

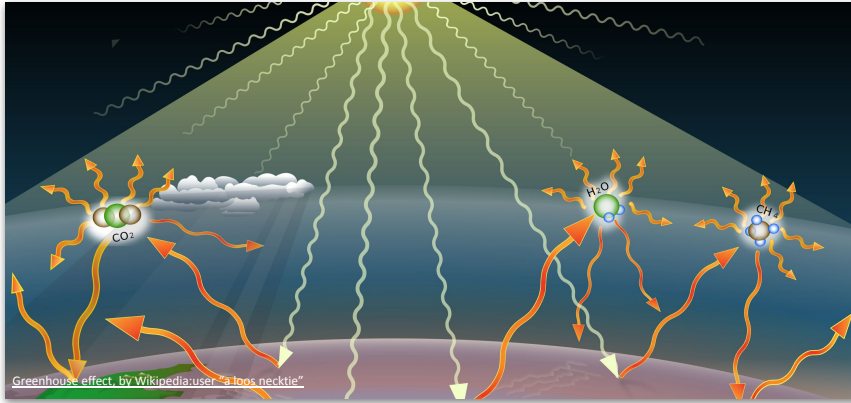


Google Result
from 04/2021

The screenshot shows a Google search interface. The search bar contains the text 'When was the Greenhouse Effect discovered?'. Below the search bar, there are navigation options for 'All', 'News', 'Images', 'Videos', 'Maps', and 'More'. The search results show 'About 16.300.000 results (0,82 seconds)'. The top result is for the year '1859', with a snippet stating: 'Irish physicist John Tyndall is commonly credited with discovering the greenhouse effect, which underpins the science of climate change. Starting in 1859, he published a series of studies on the way greenhouse gases including carbon dioxide trapped heat in the Earth's atmosphere. 2 Sep 2016'. Below this is a link to a news article: 'www.climatechangenews.com › 2016/09/02 › the-woman... Meet the woman who first identified the greenhouse effect'. There is also a 'People also ask' section with several related questions: 'When was greenhouse effect first discovered?', 'Who discovered green house effect?', 'How long has the greenhouse effect been around?', and 'When did we start noticing global warming?'. At the bottom, there are more search results, including one from 'www.rigb.org' titled 'Who discovered the greenhouse effect? | The Royal Institution ...' and one from 'en.wikipedia.org' titled 'Greenhouse effect - Wikipedia'.

Find a solution for the following question:

- **When was the Greenhouse Effect discovered?**



Bing Result
from 04/2020

When was the Greenhouse effect discovered?

ALL IMAGES VIDEOS MAPS NEWS SHOPPING | MY SAVES

7,540,000 Results Any time ▾

When was the greenhouse effect discovered?

The greenhouse effect is a natural process that is millions of years old. It plays a critical role in regulating the overall temperature of the Earth. The greenhouse effect was first discovered by **Joseph Fourier** in **1827**, experimentally verified by **John Tyndall** in **1861**, and quantified by **Svante Arrhenius** in **1896**.

* The greenhouse effect was discovered by Joseph Fourier in 1827, first reliably experimentally on by John Tyndall in 1861, and first reported quantitatively by Svante Arrhenius in 1896.
Image: Slideshare.net

What is the greenhouse effect? | What's Your Impact

whatsyourimpact.org/greenhouse-effect

Was this helpful? 👍 👎

PEOPLE ALSO ASK

- What are the major causes of global warming? ▾
- What are facts about the greenhouse effect? ▾
- What is climate skeptics taught me about global warming? ▾
- When did scientist discovered the ozone hole? ▾

History of the greenhouse effect and global warming

<https://www.lenntech.com/greenhouse-effect/global-warming-history.htm> ▾

History of the greenhouse effect and global warming Svante Arrhenius (1859-1927) was a Swedish scientist that was the first to claim in **1896** that fossil fuel combustion may eventually result in ...

Feedback

Who discovered the greenhouse effect? | The Royal ...

<https://www.rigb.org/blog/2019/may/who-discovered-the-greenhouse-effect>
 Who discovered the greenhouse effect? | The Royal Institution: Science Lives Here 160 years ago, on 18 May 1859, the Irish physicist John Tyndall wrote in his journal 'the subject is completely in my hands'...

Greenhouse effect - Wikipedia

https://en.wikipedia.org/wiki/Greenhouse_effect

Overview **History** Description Details Greenhouse gases

The existence of the greenhouse effect, while not named as such, was proposed by Joseph Fourier in 1824. The argument and the evidence were further strengthened by Claude Pouillet in 1827 and 1838. John Tyndall was the first to measure the infrared absorption and emission of various gases and vapours. From 1859 onwards, he showed that the effect was due to a very small proportion of the atmosphere, with the main gases having no effect, and was largely due to water vapour, though small percentages of hydroca...

Wikipedia · Text under CC-BY-SA license

Discovery Of The Greenhouse Effect - Greenhouse Gases

<https://www.climate-policy-watcher.org/.../discovery-of-the-greenhouse-effect.html>
 23.03.2020 · It was during the 19th century that scientists realized that gases – such as CO2 – found within the atmosphere cause a 'greenhouse effect' that regulates the atmosphere's temperature.

History of the greenhouse effect and global warming

<https://www.lenntech.com/greenhouse-effect/global>
 History of the greenhouse effect and global warming. Svante Arrhenius (1859-1927) was a Swedish scientist that was the first to claim in 1896 that fossil fuel combustion may eventually result in enhance...

History of climate change science - Wikipedia

https://en.wikipedia.org/wiki/History_of_climate_change_science
 The history of the scientific discovery of climate change began in the early 19th century when ice ages and other natural changes in paleoclimate were first suspected and the natural greenhouse effect was...

How Joseph Fourier discovered the greenhouse effect

<https://www.irishtimes.com/news/science/how-joseph-fourier-discovered...>
 21.03.2019 · How Joseph Fourier discovered the greenhouse effect That's Maths: French physicist's study of heat conduction led him to analyse why Earth was so warm Thu, Mar 21, 2019, 05:00

greenhouse effect | Definition, Diagram, Causes, & Facts ...

<https://www.britannica.com/science/greenhouse-effect>
 French mathematician Joseph Fourier is sometimes given credit as the first person to coin the term greenhouse effect based on his conclusion in 1824 that Earth's atmosphere functioned similarly to a...

Greenhouse Effect | National Geographic Society

<https://www.nationalgeographic.org/encyclopedia/greenhouse-effect>
 Since the Industrial Revolution in the late 1700s and early 1800s, people have been releasing large quantities of greenhouse gases into the atmosphere. That amount has skyrocketed in the past century...

Greenhouse Effect



The greenhouse effect is the process by which radiation from a planet's atmosphere warms the planet's surface to a temperature above what it would be without this atmosphere. Radiatively active gases (i.e., greenhouse gases) in a planet's atmospher...

Wikipedia

Related people

See all (5+)

Joseph Fourier	Svante Arrhenius	John Tyndall	Veerabhadran Ramanathan	Stephen Schneider

People also search for

See all (15+)

Greenhouse Gas	Carbon Dioxide in Earth's At...	Water Vapor	Earth's Energy Budget	Atmosphere

Data from: Wikipedia · Apextribune
 Wikipedia text under CC-BY-SA license
 Suggest an edit

Search Engine Indexes

- Web pages
- News articles
- Tweets
- etc.

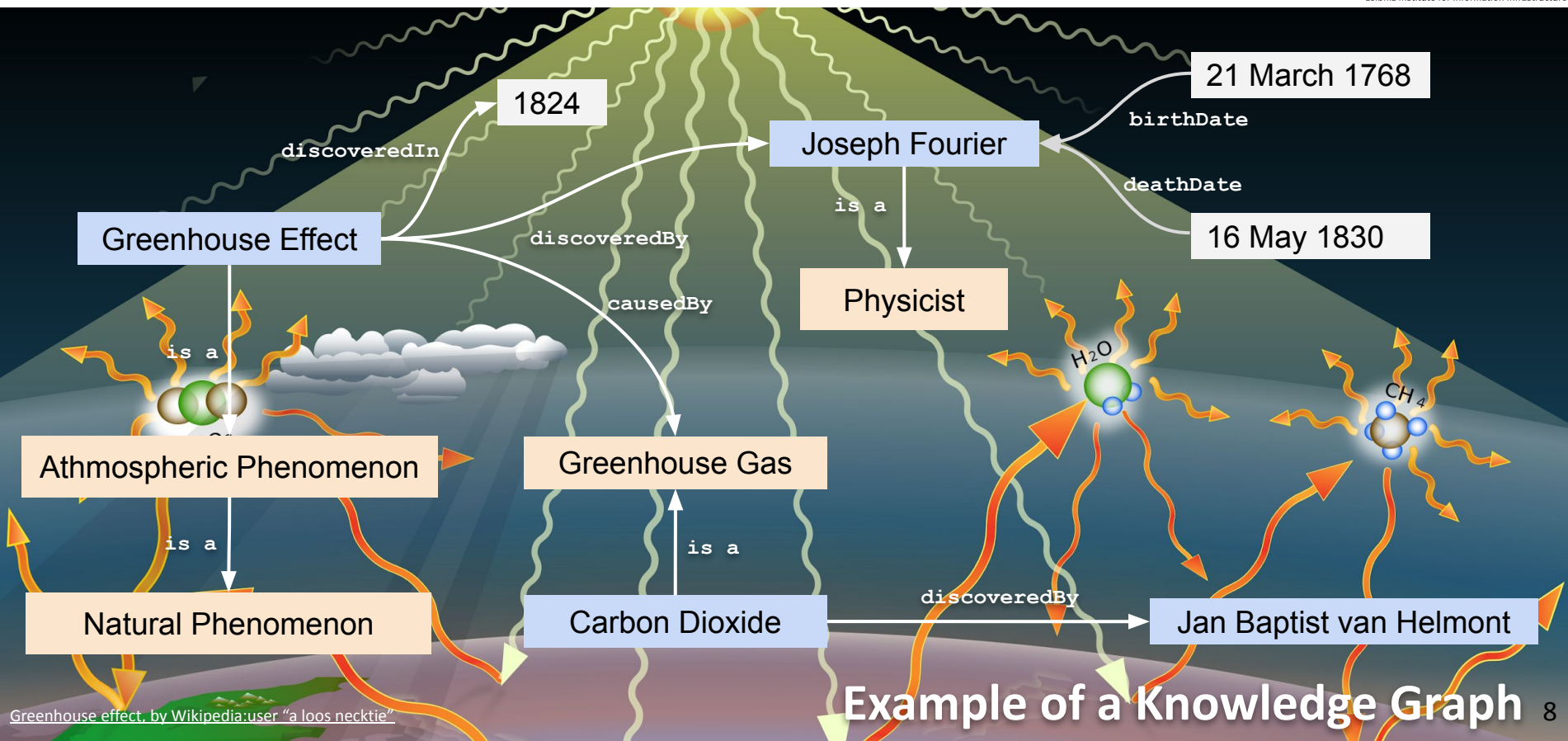
Knowledge Graph

- Facts (structured data)

1. Information, Natural Language and the Web / 1.1 How to get Information (from the Web)?

How to get Information (from the Web)?

What does it mean to know “the Greenhouse Effect”?



Example of a Knowledge Graph 8

Greenhouse effect, by Wikipedia:user "a loos necktie"

1.1 How to get Information (from the Web)?

1.2 Communication, Language, and Understanding

1.3 How to measure Information?

1.4 The ever-growing Web of Information

1.5 Search Engines on the Web

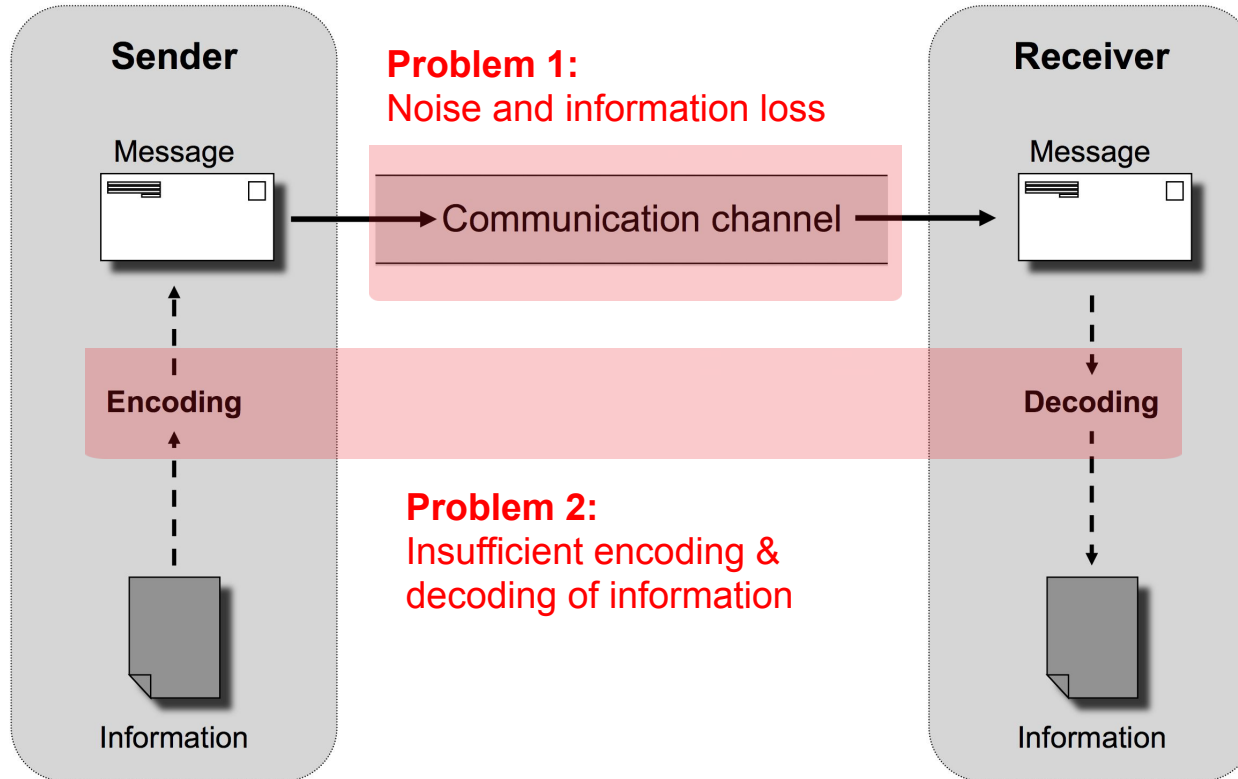
1.6 The Meaning of Information

What is Communication?

- **Communication** is a process by which **information** is **exchanged** between individuals through a **common system of symbols, signs, or behavior.**

[Merriam-Webster]

Communication Models - How Information is transferred



Ch. Meinel, H. Sack:
[Digital Communication - Communication Multimedia](#), Security, Springer, 2014.

What is Language?

- **Language**, a system of conventional **spoken, manual, or written symbols** by means of which human beings, as members of a social group and participants in its culture, **express** themselves.
- The functions of language include **communication**, the expression of identity, play, imaginative expression, and emotional release.

[Encyclopaedia Britannica]



Natural Language

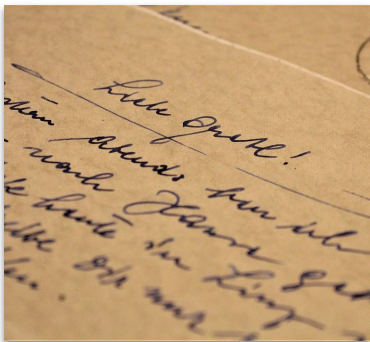
- Natural language is a discrete categorical system of symbols that combine to **convey meaning**.
- It has **evolved naturally and historically in humans** through use and repetition **without planning**.
- A natural language is **different from a constructed or formal language (artificial language)**, such as an auxiliary language (Esperanto), a programming language (C, Python, Java, etc.), arithmetic language, or a language used to study logic.

How do we communicate Natural Language?



Speech

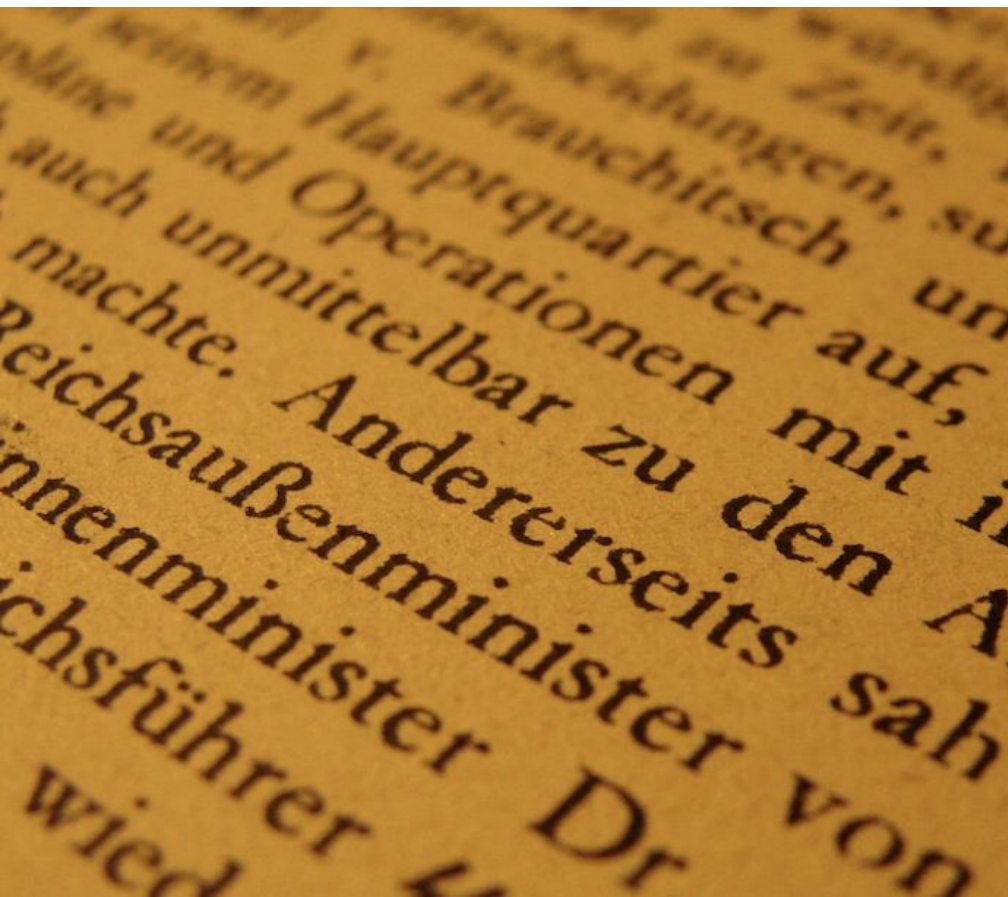
- Listening
- Speaking



Text

- Reading
- Writing

How do we encode Textual Information?



Alphabetic Writing System

- A single character refers to a single sound (*phonemic alphabet*).
- In **consonant alphabets** (*abjads*), vowels can be deduced from context.

How do we encode Textual Information?

Syllabic and Logographic Writing System

- Syllabic systems also involve a mapping between characters and sound, (*but refer on larger units*).
- A logograph is a symbol that represents a unit of meaning.

種白蒺藜今生同州沙苑牧馬草地最多而近道亦有之綠葉細蔓
 披布沙上七月開花黃紫色如碗豆花而小九月結實蒺藜子便可
 採其實味甘而微腥稠綠色與蠶種子相類而差大又與馬蒺子
 相類但馬蒺子微大不堪入藥項細辨之今人多用然古方云蒺藜
 子皆用有刺者治風明目最良神仙方亦有單用蒺藜云下問黑白
 但取堅實者春去刺用煎主痔漏陰汗及婦人發乳帶下葛洪治卒
 中五尸驚蒺藜子蜜丸
 服如胡臣二枚日三愈
 刺盡用酒拌再蒸從
 午至酉出日乾用
 大蓋煮取半蓋仰卧先滿口含飯以汁一合灌鼻
 中下過再灌之噴出一兩個意肉以赤蛹蟲即差
 引腰脊痛搗末蜜和丸酒
 服如胡臣大九日三服
 補肝散治三十年失明蒺藜子七
 寸
又方 治腫蒺藜子一外熬令黃搗篩以麻油和如泥
 熨攻心如刺吐清汁七月七日揮蒺
 藜子陰乾作灰先食服方寸匕日三
 上如破
又方 疔急小兒爛喉齋繞身即死以
 塗之
又方 疔急小兒爛喉齋繞身即死以
 洗三寸截之以水五外煮取二外去滓內銅器中
 又煮取一外內小器中如稠糖下取傳瘡腫上
千金方 疔急小兒爛喉齋繞身即死以
 疔急小兒爛喉齋繞身即死以

以蒺藜子苗黃湯洗
 末米湯下一匙相去
 四百里不下再服
梅師方 治難產礙胎在腹中如已見兒并胞
 衣不出胎死蒺藜子貝母各四兩為
孫真人食忌 治白癩風以白蒺藜子
 生薑為末作湯服之
仙秘百云 服蒺藜子一碩當七八月熟時收口乾春去刺然後杵
 生服之一年已後冬不寒夏不熱服之三年身輕長生
 復少髮白復黑齒落重上服之三年身輕長生
衍義曰 蒺藜
 等一掌柱蒺藜即今之道傍布地而生或生牆上有小黃花結莢刺
 此正是牆有茨者花收摘陰乾為末每服三二錢飯後以溫酒調服
 治白癩風又一種白蒺藜出同州沙苑牧馬地黃紫花作
 笑結子如羊內腎補腎藥今人多用風家惟用刺蒺藜

Artificial Language

```

public static void main(String[] args) throws FileNotFoundException {

    String filePath = "/Users/hsa/Documents/Workspace/Scientometrics2016/testdata_input";
    File fout = new File("/Users/hsa/Documents/Workspace/Scientometrics2016/SemanticSim_testdata_10000");
    FileOutputStream fos = new FileOutputStream(fout);

    try {
        BufferedReader lineReader = new BufferedReader(new FileReader(filePath));
        BufferedWriter bw = new BufferedWriter(new OutputStreamWriter(fos));
        String lineText = null;
        int nr=0;

        while ((lineText = lineReader.readLine()) != null) {

            String[] splitStr = lineText.split("\\s+");
            nr++;

            // DBpedia settings - without giving prefix list, using property patterns instead
            // it will filter properties which start with "http://dbpedia.org/ontology"
            ResourceSimilarityMeasure rsmForDBpedia = new ResourceSimilarityMeasure(
                PropertyRestriction.SamePropertyPath,
                "http://dbpedia.org/sparql",
                null,
                "http://dbpedia.org/ontology/",
                null, null, null, null);

            System.out.println(
                rsmForDBpedia.getSimilarity("<http://dbpedia.org/resource/Albert_Einstein>", "<http://dbpedia.org/re
  
```

1. Information, Natural Language and the Web

1.2 Communication, Language, and Understanding?

What does it mean to understand?

- **Understanding** is the ability to grasp the meaning of information.
- **Information** is conveyed in a **message** using a specific **language**.
- **Information is understood** by the receiver of a message, if the receiver **interprets** the information **correctly**.

- 1.1 How to get Information (from the Web)?
- 1.2 Communication, Language, and Understanding
- 1.3 How to measure Information?**
- 1.4 The ever-growing Web of Information
- 1.5 Search Engines on the Web
- 1.6 Deficiencies of the traditional Web
- 1.7 The Meaning of Information

1. Information, Natural Language and the Web

1.3 How to measure Information?

What is information?

- **Information** is that which informs.
- **Information** is conveyed as the **content of a message**.
- **Information** can be **encoded** into various forms for **transmission** and **interpretation**.
- **Information** is any **propagation of cause and effect** within a system.
- **Information's existence** is not necessarily coupled to an **observer**.
- **Information** reduces **uncertainty (Information Theory)**.

Information according to Information Theory

- **Information reduces uncertainty.**
- **Uncertainty of an event** is measured by its **probability of occurrence** and is **inversely proportional** to that.
- The **more uncertain** an event, the **more information** is required to resolve uncertainty of that event.

Information Content

Discrete Random Variable

- A **random variable** x takes a value x from the alphabet \mathbf{X} with probability $p_x(x)$.
- The **vector of probabilities** is $\vec{p}(x)$ (*probability mass function*).

Information Content

Discrete Random Variable

- **Examples:**

- Coin Tossing:

- $\mathbf{X}=[\text{head}; \text{tail}] ; \vec{p}(x) = \left[\frac{1}{2}; \frac{1}{2}\right]$

- Dice Tossing:

- $\mathbf{X}=[1; 2; 3; 4; 5; 6] ; \vec{p}(x) = \left[\frac{1}{6}; \frac{1}{6}; \frac{1}{6}; \frac{1}{6}; \frac{1}{6}; \frac{1}{6}\right]$

- German language:

- $\mathbf{X}=[a; b; c; d; \dots; z] ; \vec{p}(x) = [0.0651; 0.0189; 0.0306; 0.0508; \dots ; 0.0103]$

Shannon Information Content

- The **Shannon Information Content** of an outcome with **probability p** is:

$$SIC = -\log_2 p$$

- The unit to measure information is **bit**.
(*binary digit, basic indissoluble information unit*)



[Claude E. Shannon](#)
(1916-2001)

Shannon Information Content

- **Examples:**

- Coin Tossing:

- $\mathbf{X}=[\text{head}; \text{tail}]$; $\vec{p}(x) = [\frac{1}{2}; \frac{1}{2}]$; **SIC=[1; 1] bit**

- My birthday:

- $\mathbf{X}=[\text{birthday}; \text{no birthday}]$;

- $\vec{p}(x) = [\frac{1}{365}; \frac{364}{365}]$; **SIC=[8.512; 0.004] bit**

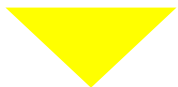
Entropy

- The **Information Content (Entropy)** H of a message M is based on the information content of each symbol $s \in M$ and its relative frequency of occurrence (probability):

$$H(\vec{p}_x) = |M| \cdot \sum_{i=1}^n p_i \cdot (-\log_2 p_i)$$

Entropy Example

$$H(\vec{p}_x) = |M| \cdot \sum_{i=1}^n p_i \cdot (-\log_2 p_i)$$



$$= 31 \times 3.48375 = 107.99625$$

Your string is: **information service engineering**

Alphabet of symbols in the string: **a c e f g i m n o r s t v**

Frequencies of alphabet symbols:

- 0.065 ->
- 0.032 -> a
- 0.032 -> c
- 0.161 -> e
- 0.032 -> f
- 0.065 -> g
- 0.161 -> i
- 0.032 -> m
- 0.161 -> n
- 0.065 -> o
- 0.097 -> r
- 0.032 -> s
- 0.032 -> t
- 0.032 -> v

Shannon entropy can be calculated as follow:

$$\begin{aligned} H(X) &= -[(0.065 \log_2 0.065) + (0.032 \log_2 0.032) + (0.032 \log_2 0.032) + (0.161 \log_2 0.161) + \\ & (0.032 \log_2 0.032) + (0.065 \log_2 0.065) + (0.161 \log_2 0.161) + (0.032 \log_2 0.032) + (0.161 \log_2 0.161) + \\ & (0.065 \log_2 0.065) + (0.097 \log_2 0.097) + (0.032 \log_2 0.032) + (0.032 \log_2 0.032) + (0.032 \log_2 0.032)] \\ H(X) &= -[(-0.255) + (-0.16) + (-0.16) + (-0.425) + (-0.16) + (-0.255) + (-0.425) + (-0.16) + (-0.425) + \\ & (-0.255) + (-0.326) + (-0.16) + (-0.16) + (-0.16)] \\ H(X) &= -[-3.48375] \\ H(X) &= 3.48375 \end{aligned}$$

- 1.1 How to get Information (from the Web)?
- 1.2 Communication, Language, and Understanding
- 1.3 How to measure Information?
- 1.4 The ever-growing Web of Information**
- 1.5 Search Engines on the Web
- 1.6 Deficiencies of the traditional Web
- 1.7 The Meaning of Information

There seem to be no Limits of Growth



[Cisco Annual Internet Report \(2018 - 2023\)](#)

“The Web is big.

Really big.

**You just won't believe how vastly, hugely,
mind-bogglingly big it is.”**

(...according to Douglas Adams)



ZOOM

HOSTS **208,333**

PARTICIPANTS IN MEETINGS

REDDIT

SEES **479,452**

PEOPLE ENGAGE WITH CONTENT

NETFLIX

USERS STREAM **404,444**

HOURS OF VIDEO

DOORDASH

DINERS ORDER **555** MEALS

INSTAGRAM

USERS POST **347,222** STORIES

YOUTUBE

USERS UPLOAD **500 HRS** OF VIDEO

TWITTER

GAINS NEW USERS **319**

WHATSAPP

USERS SHARE **41,666,667** MESSAGES

\$3,805 IS SPENT ON MOBILE APPS

CONSUMERS SPEND **\$1,000,000** ONLINE

2020 every MINUTE



MINUTE of the DAY

PRESENTED BY DOMO



\$3,805 IS SPENT ON MOBILE APPS



CONSUMERS SPEND
\$1,000,000
ONLINE



1,388,889

PEOPLE MAKE VIDEO/
VOICE CALLS



MICROSOFT
TEAMS CONNECT
52,083
USERS



AMAZON

6,659 SHIPS

PACKAGES



FACEBOOK

USERS SHARE

150,000

MESSAGES



INSTAGRAM

BUSINESS PROFILE ADS
SEE CLICKS **138,889**

SPOTIFY

ADDS
28 TRACKS
TO ITS MUSIC LIBRARY



VENMO

USERS SEND
\$239,196
WORTH OF PAYMENTS



TIKTOK

IS INSTALLED
2,704 TIMES

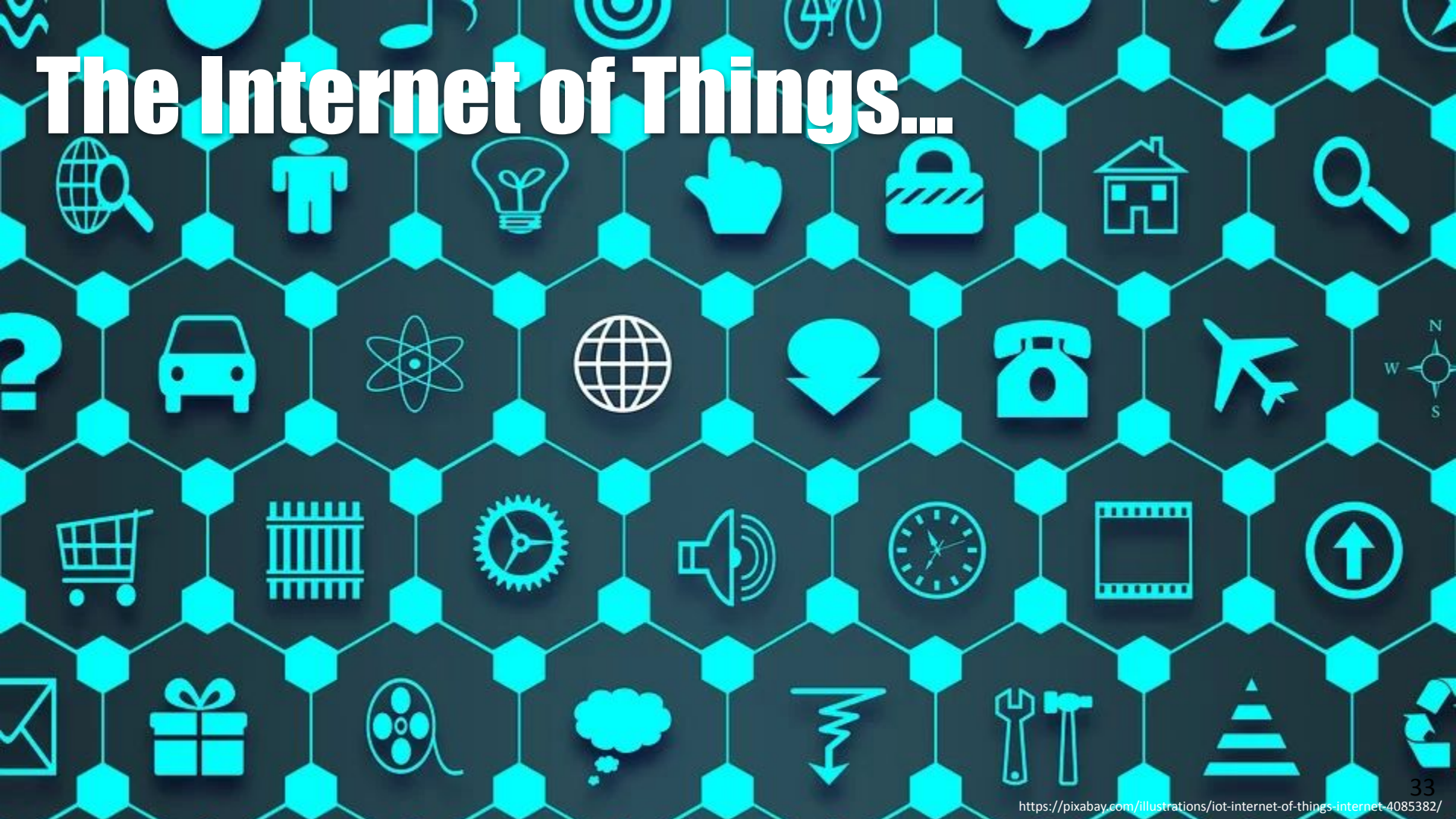


LINKEDIN

USERS APPLY FOR
69,444 JOBS

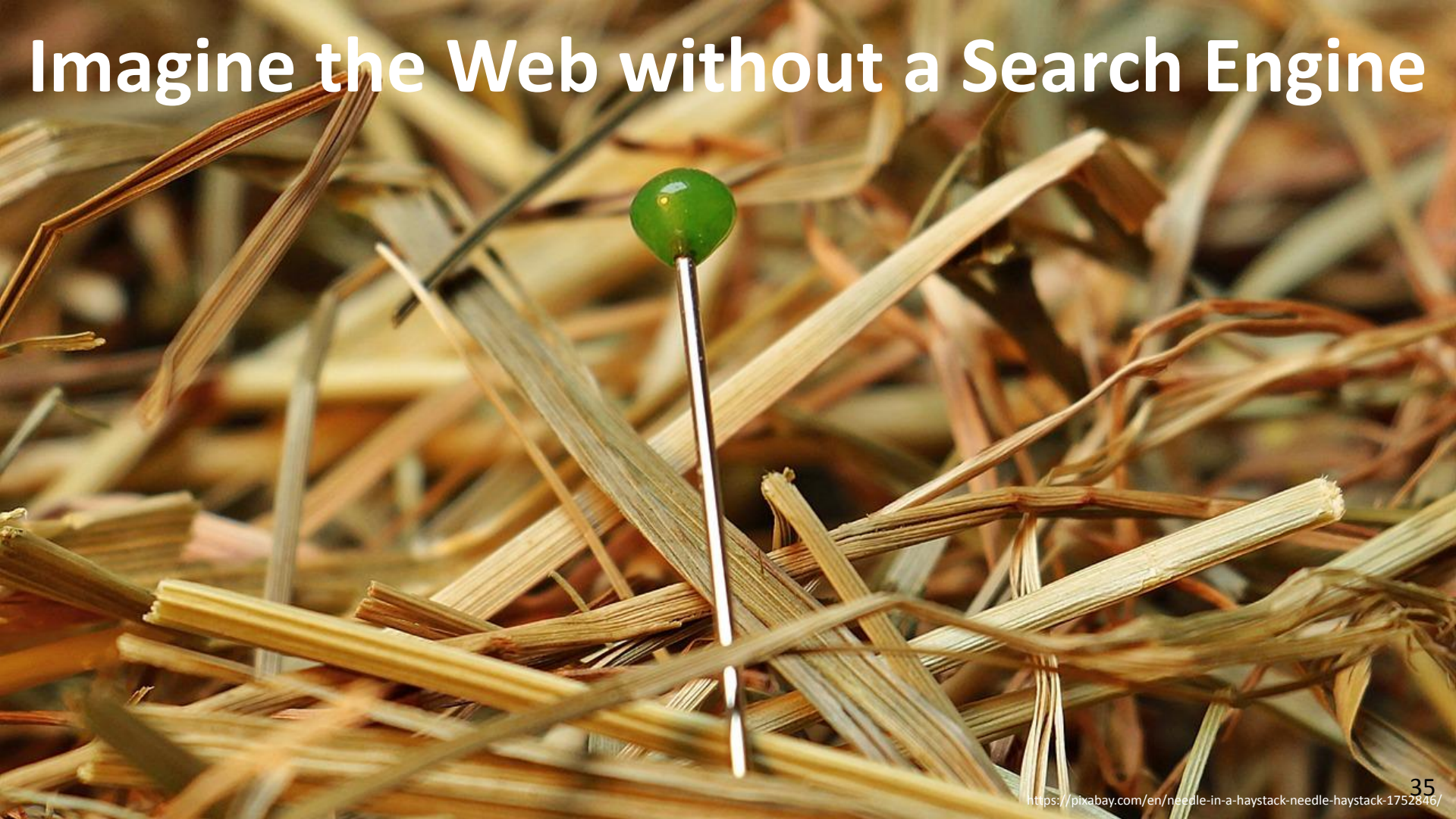


The Internet of Things...



- 1.1 How to get Information (from the Web)?
- 1.2 Communication, Language, and Understanding
- 1.3 How to measure Information?
- 1.4 The ever-growing Web of Information
- 1.5 Search Engines on the Web**
- 1.6 The Meaning of Information

Imagine the Web without a Search Engine



climate chang|



climate change

Remove

climate change definition

climate change hoax

climate change facts

climate change articles

climate change effects

climate change youtube

climate change reddit

[NASA: Climate Change and Global Warming](#)

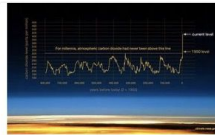
Feedback on these suggestions

<https://climate.nasa.gov>

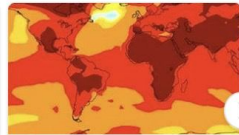
Mar 16, 2020 · NASA's Climate Kids website brings the exciting science of climate change and sustainability to life, providing clear explanations for the big questions in climate science. Targeting upper-elementary-aged children, the site includes interactive games, hands-on activities, and engaging articles that make climate science accessible and fun.



Arctic Ice Melt Is Changing Ocean Currents – Climate Change: Vital Signs of the ...



Graphic: The relentless rise of carbon dioxide – Climate Change: Vital Signs of the ...



Study Confirms Climate Models are Getting Future Warming Projections ...

[climate change | Causes, Effects, & Facts | Britannica](#)

<https://www.britannica.com/science/climate-change>

Climate change, the periodic modification of Earth's climate caused by changes in the atmosphere and interactions between the atmosphere and various other geologic, chemical, biological, and geographic factors. Learn how climate has changed since the last ice ...

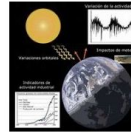
[Evidence for Climate Change](#) · [Climate Change Within a Human Life Span](#) · [Greenhouse Gases](#)

[What Is Climate Change? | NASA](#)

<https://www.nasa.gov/audience/forstudents/k-4/>

To learn about climate change, you first must know what climate is.

Climate Change



Climate change occurs when changes in Earth's climate system result in new weather patterns that remain in place for an extended period of time. This length of time can be as short as a few decades to as long as millions of years. Scientists have identified many episodes of climate change during Earth's geological history; more recently since the industrial revolution the climate has increasingly been affected by human activities driving global warming, and the terms are commonly used interchangeably in that context.

Wikipedia

People also search for

See all (10+)



Global Warming



Greenhouse Effect



Deforestati...



Climate Change and Agriculture



Attribution of Recent Climate Ch...

Data from: Wikipedia

[Suggest an edit](#)

See results for

[Global Warming](#)

Global warming is the long-term rise in the average

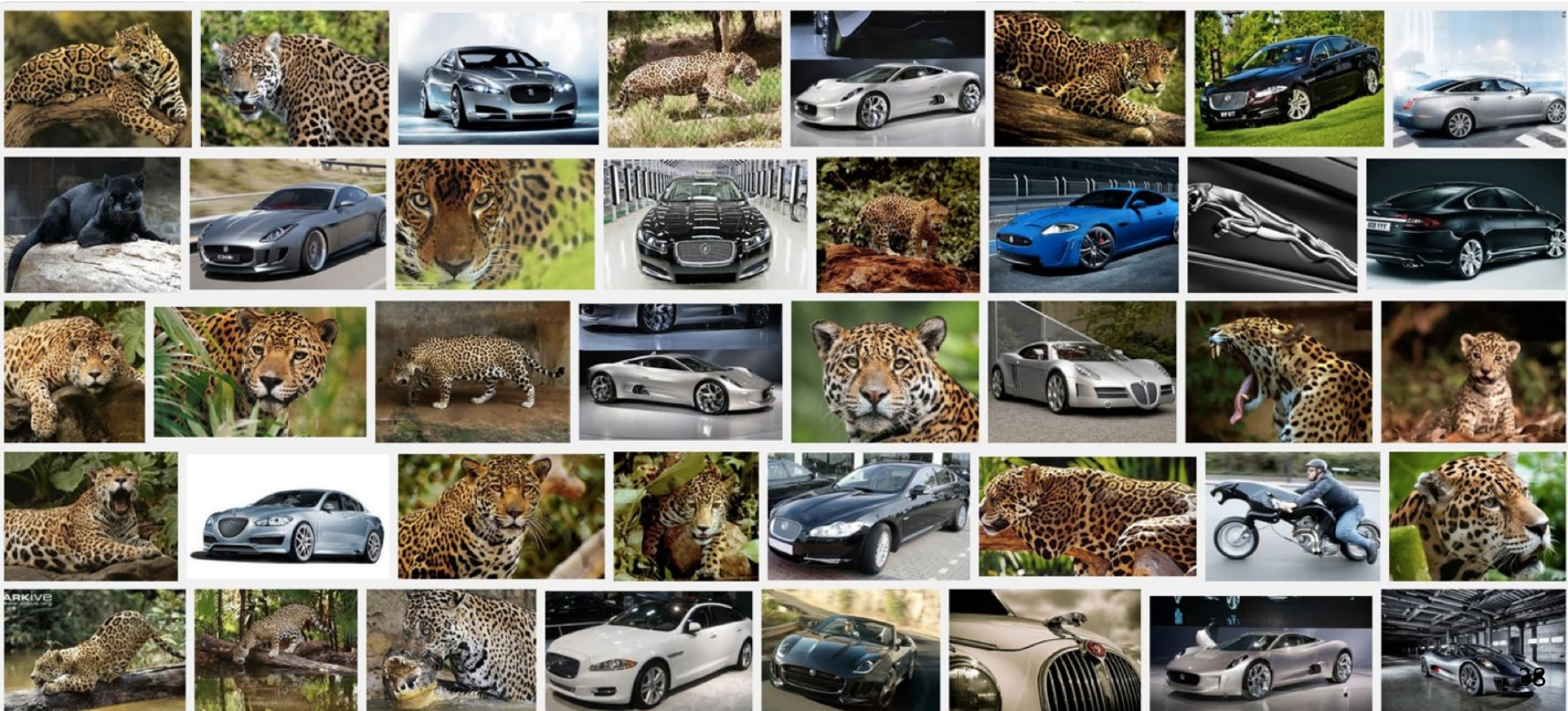


- Instant Search
- Autocompletion
- Knowledge Graph
- Question Answering
- News
- Images & Video
- Search Recommendations
- Quotes
- eCommerce
- etc.

Search Engines and the Web

- The World Wide Web is a **distributed hypermedia system** with
 - multimedia documents and
 - linked via hyperlinks

Do you always find what you are looking for?



おすすめ番組

映画

- ▶ テレビジャパン・シネマシアター「火垂るの墓」＜英語字幕付き＞
- ▶ テレビジャパン・シネマシアター「ルパン三世」
- ▶ テレビジャパン・シネマシアター「謝罪の王様＜英語字幕付き＞」

ドラマ

- ▶ ドラマ10「美女と男子」
- ▶ ドラマ「刑事7人」
- ▶ 戦後70年「一番電車が走った」
- ▶ 木曜時代劇「まんまこと～麻之助裁定帳～」

ドキュメンタリー

- ▶ NHKスペシャル
- ▶ ザ・プレミアム 鈴木亮平 “絶景！ミステリー遺産” に挑む！アドリア海縦断・7日間の大冒険
- ▶ ザ・プレミアム 風雲！大歴史実験
- ▶ NHKスペシャル
- ▶ BS1スペシャル「戦火のマエストロ・近衛秀麿～ユダヤ人の命を救った音楽家～」
- ▶ 世界ふれあい街歩き
- ▶ 奇跡のレッスン～世界の最強コーチと子どもたち～

- Which information is important and how do you know?
- Which information can be trusted?
- Which information is related by content?
- **What does the information mean?**

- ▶ よみがえりマイスター
- ▶ 夏のうまい旅祭り 妄想二ホン キッチンが走る！
- ▶ 新 クイズ面白ゼミナール

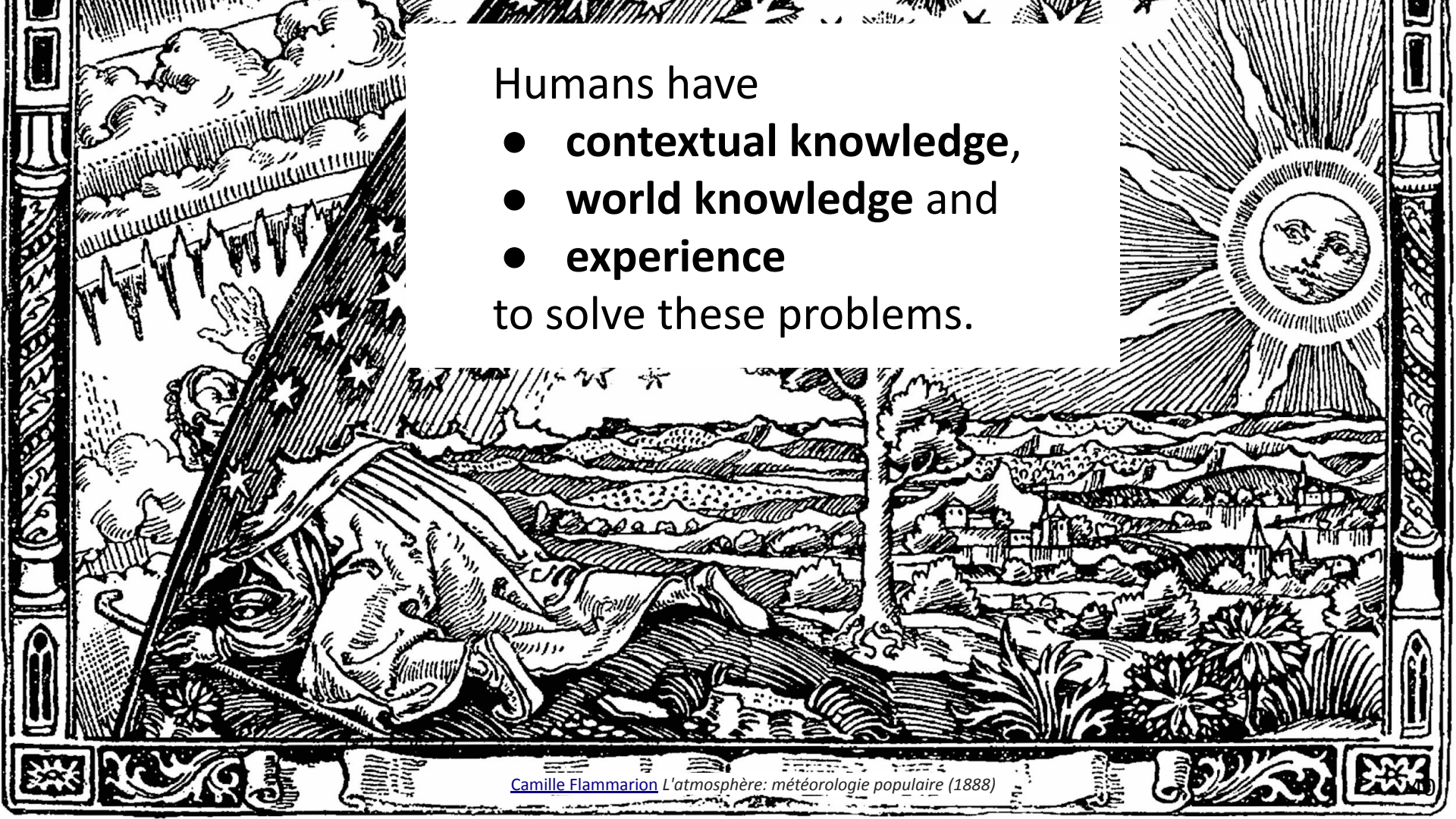
音楽

- ▶ NHKスペシャル

Humans have

- contextual knowledge,
- world knowledge and
- experience

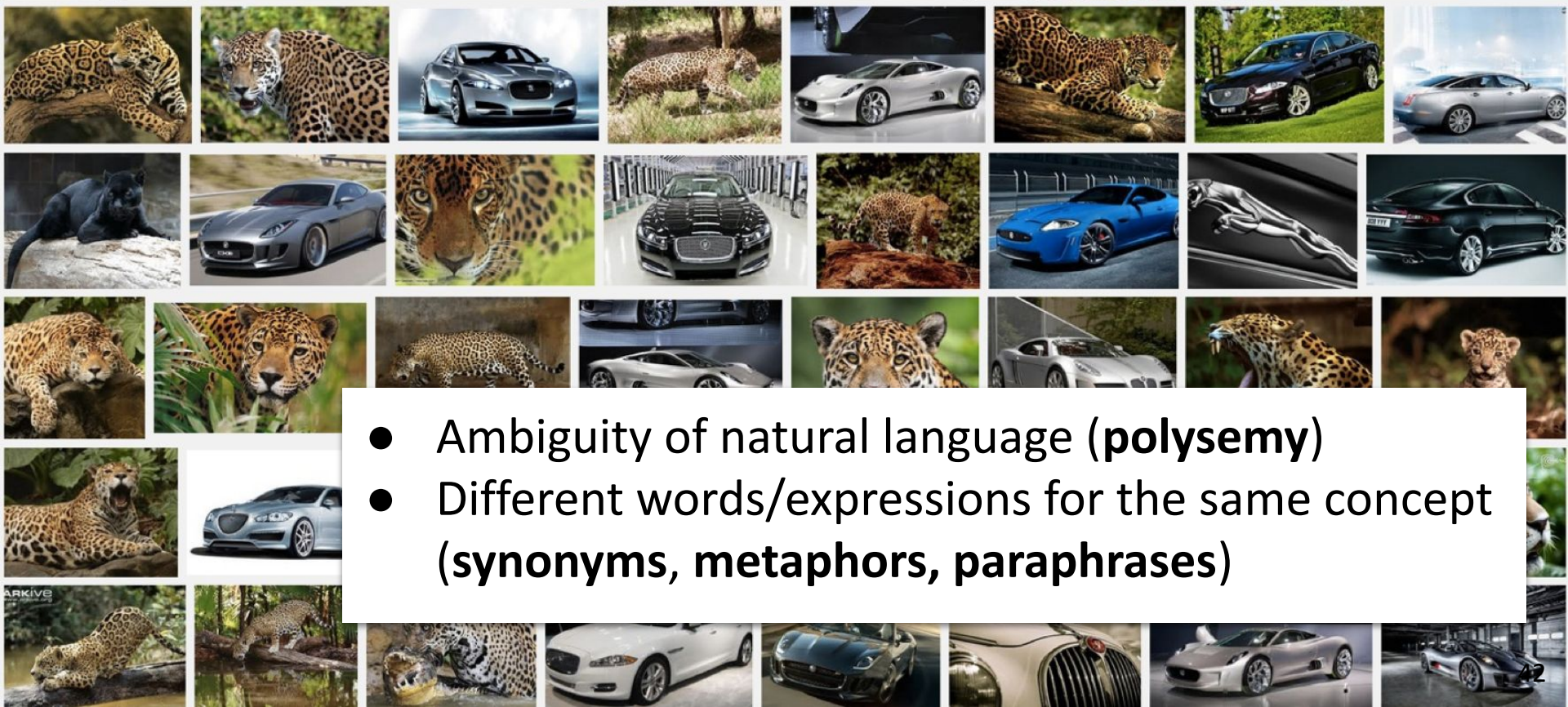
to solve these problems.



The (Document) Web is for Humans

- The Web is based on the **markup language HTML**
- HTML describes
 - how information is **presented**
 - how information is linked
 - **but not, what the information means**

The Information Retrieval Dilemma



- Ambiguity of natural language (**polysemy**)
- Different words/expressions for the same concept (**synonyms, metaphors, paraphrases**)

- 1.1 How to get Information (from the Web)?
- 1.2 Communication, Language, and Understanding
- 1.3 How to measure Information?
- 1.4 The ever-growing Web of Information
- 1.5 Search Engines on the Web
- 1.6 The Meaning of Information**

What is Meaning?

- *(In the philosophy of language, metaphysics, and metasemantics,)*
- **Meaning** is a relationship between two sorts of things:
 - **Signs** and
 - the kinds of things they **intend, express, or signify.**
- **Words** (and nonverbal symbols) are **necessarily meaningful**

What does it mean “to understand”?

- **Understanding** (in general) is the ability to grasp the meaning of information.
- **Information** is conveyed in a **message** using a specific **language** from a sender to a receiver.
- **Information is understood** by the receiver of a message, if the receiver **interprets** the information **correctly**.

Meaning and Comprehension

- **Correct Interpretation (Understanding)** depends on
 - Syntax,
 - Semantics,
 - Context,
 - Pragmatics, and
 - Experience.

Syntax

- =[greek] *Arrangement, Ordering*
- In **grammatics**, syntax denotes the study of the **principles** and processes **by which sentences are constructed** in particular languages.
- In **formal languages**, syntax is just a set of rules, by which **well formed expressions** can be created from a fundamental set of symbols (alphabet).

Semantics

- =[greek] *pertains to the character, the study of meaning*
- Semantics is part of the linguistics focussed on **Sense and Meaning** of language or symbols of language.
- Semantics is the **study of interpretation of signs or symbols** as used by agents or communities within particular circumstances and **contexts**.
- Semantics asks, **how sense and meaning of complex concepts can be derived from simple concepts** based on the **rules of syntax**.
- The semantics of a message depends on **context** and **pragmatics**.

Context

- [lat.] *contextus = interweaved*
- Context denotes the **surrounding of a symbol** (concept) in an expression resp. its **relationship with surrounding expressions** (concepts) and further related elements (**verbal context**).
- Contexts denotes **all elements of any sort of communication that define the interpretation of the communicated content**, as e.g.
 - social context
 - temporal context
 - cultural context

Pragmatics

- =[greek] *action*
- Pragmatics reflects the **intention by which the language is used** to communicate a message.
- In linguistics pragmatics denotes the **study of applying language in different situations.**
- It also **denotes the intended purpose** of the speaker.
- Pragmatics studies **the ways in which context contributes to meaning.**

Semantics vs. Pragmatics

- The boundary between semantics and pragmatics is open for debate.
- In this lecture, we say that **semantics looks at the literal meaning of a sentence**, while **pragmatics investigates the meaning of an utterance**, that is, the use of the sentence.

Example sentences

1. Stand up, Bob! (*imperative; command*)
2. Could you please stand up, Bob? (*interrogative; question*)
3. Bob stands up. (*assertive; statement*)

Do all speech acts (1, 2 and 3) have the same meaning?

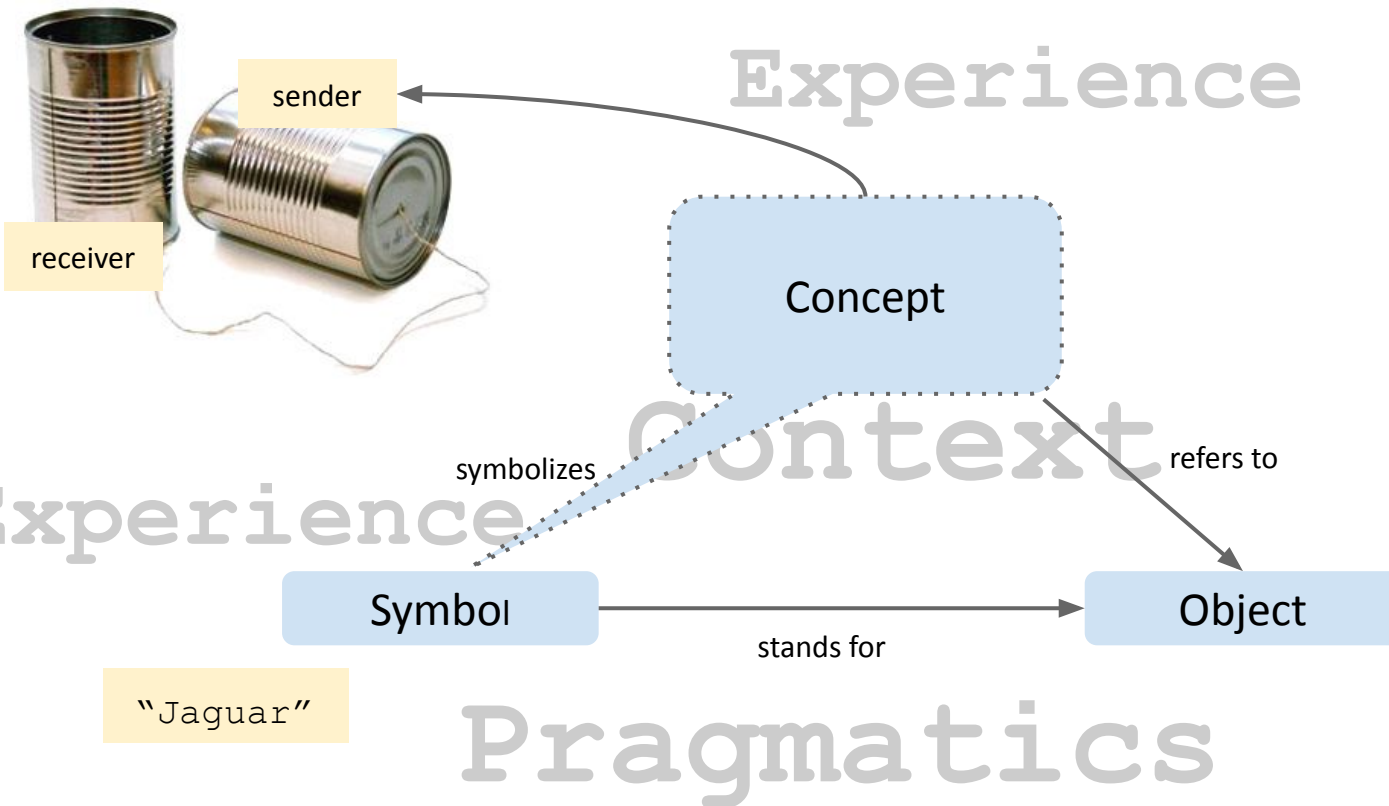
Experience

- **Experience** considers all information that you have learned and put in context with the world you are living in.
- Experience in this sense is often referred to as **common sense knowledge** or **world knowledge**.

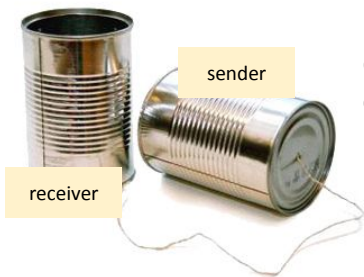
Successful Communication

- For **successful communication**,
 - information has to be correctly transmitted (**Syntax**)
 - the meaning (**Semantics**) of the transmitted information must be interpreted correctly (= **understanding**)
- **Understanding** furthermore depends on
 - the **context** of both sender and receiver and
 - the **pragmatics** of the sender
- **Context** of sender and receiver depend on
 - the **experience** (knowledge of the world) of both sender and receiver

Communication of Meaning



Language and Semiotics



Communication is a result of
consistent social agreement



Dyadic Model by
Ferdinand de Saussure

Signified

abstract concept

reciprocal evocation

Signifier

arbitrary

/tʃɛə(r)/

concrete sound pattern

stands for

refers to

Object

concrete



- 1.1 How to get Information (from the Web)?
- 1.2 Communication, Language, and Understanding
- 1.3 How to measure Information?
- 1.4 The ever-growing Web of Information
- 1.5 Search Engines on the Web
- 1.6 The Meaning of Information

1. Information, Natural Language and the Web

Bibliography

- R. Baeza-Yates, B. Ribeiro-Neto, [Modern Information Retrieval](#), 2nd ed., Addison Wesley, 2010.
- Ch. Meinel, H. Sack: [Digital Communication - Communication Multimedia, Security, Springer](#), Heidelberg, 2014. (available via KIT network)
- C.K. Ogden, I.A. Richards: [The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism](#), 1923.
- The Internet Domain Survey Host Count, <http://www.isc.org/>.
- Douglas Adams: *The Hitchhiker's Guide to the Galaxy*, Pan Books, UK, 1979.

1. Information, Natural Language and the Web

Syllabus Questions

- What is a knowledge graph and how do (web) search engines make use of a knowledge graph?
- What is communication?
- What are the differences between natural and artificial languages?
- What does it mean to understand?
- How does information theory measure information?
- What is (information) entropy and how is it computed?
- What are the main challenges for search engines to „understand“ information of the Web?
- What does the correct interpretation of information depend on?
- What is syntax, semantics, context, pragmatics, and experience?
- What is required for successful communication?