

Parcours Data

4 mai

Données & Métadonnées
Open Data
Étude de cas

7 mai

Plateformes techniques
TD : Introduction à Neo4j

14 mai

Veille & OSINT
TD : Construire une veille
avec Inoreader

TBD

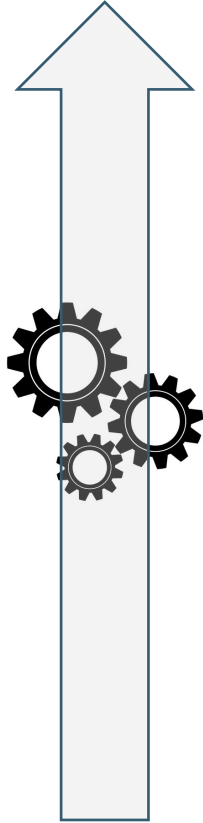
Projet cartographie



The world's largest taxi firm, Uber, owns no cars. The world's most popular media company, Facebook, creates no content. The world's most valuable retailer, Alibaba, carries no stock. And the world's largest accommodation provider, Airbnb, owns no property. Something big is going on.

Tom Goodwin, SVP Strategy and Innovation, Havas Media, 2015

DATA



sagesse
compréhension
connaissance
information
donnée

NDATA

METADATA

DATA

METADATA

METADATA

METADATA

METADATA

*MET***ADATA**





structurelles

techniques

droits

METADONNÉES DE GESTION

traçabilité

pérennisation



File Name : beyonce1.jpg
Directory : Downloads
File Size : 237 kB
File Modification Date/Time : 2019:04:08 22:58:36+02:00
File Access Date/Time : 2019:04:08 22:58:38+02:00
File Inode Change Date/Time : 2019:04:08 22:58:36+02:00
File Permissions : rw-r--r--
File Type : JPEG
File Type Extension : .jpg
MIME Type : image/jpeg
Exif Byte Order : Little-endian (Intel, II)
Copyright : 2017 Getty Images
Quality : 100%
XMP Toolkit : Adobe XMP Core 5.0-c061 64.140949, 2010/12/07-10:57:01
Creator Tool : Adobe Photoshop CS6 (Windows)
Authors Position : Stringer
Instance ID : xmp.iid:60100F7A520A11E99738CE282231B326
Document ID : xmp.did:60100F7B520A11E99738CE282231B326
Rights : 2017 Getty Images
Creator : Kevork Djansezian
Title : 695236579AB00210_The_59th_G
Derived From Instance ID : xmp.iid:3B02E05752011E99738CE282231B326
Derived From Document ID : xmp.did:3B02E05852011E99738CE282231B326
Current IPTC Digest : f8f9d1c70b7ed2cce4f21239f928fed4
Coded Character Set : UTF8
Application Record Version : 2
Copyright Notice : 2017 Getty Images
IPTC Digest : f8f9d1c70b7ed2cce4f21239f928fed4
DCT Encode Version : 100
APP14 Flags 0 : [14], Encoded with Blend=1 downsampling
APP14 Flags 1 : (none)
Color Transform : YCbCr
Image Width : 480
Image Height : 545
Encoding Process : Baseline DCT, Huffman coding
Bits Per Sample : 8
Color Components : 3
Y Cb Cr Sub Sampling : YCbCr4:4:4 (1 1)
Image Size : 480x545
Megapixels : 0.262



ExifTool Version Number : 11.30
File Name : architecture-buildings-china-1366957.jpg
Directory : Downloads
File Size : 2.0 MB
File Modification Date/Time : 2019:04:09 10:46:09+02:00
File Access Date/Time : 2019:04:09 10:46:14+02:00
File Inode Change Date/Time : 2019:04:09 10:46:12+02:00
File Permissions : rw-r--r--
File Type : JPEG
File Type Extension : jpg
MIME Type : image/jpeg
JFIF Version : 1.01
Resolution Unit : inches
X Resolution : 72
Y Resolution : 72
Profile CMM Type : Linotronic
Profile Version : 2.1.0
Profile Class : Display Device Profile
Color Space Data : RGB
Profile Connection Space : XYZ
Profile Date Time : 1998:02:09 06:49:00
Profile File Signature : acsp
Primary Platform : Microsoft Corporation
CMM Flags : Not Embedded, Independent
Device Manufacturer : Hewlett-Packard

Device Model : sRGB
Device Attributes : Reflective, Glossy, Positive, Color
Rendering Intent : Perceptual
Connection Space Illuminant : 0.9642 1 0.82491
Profile Creator : Hewlett-Packard
Profile ID : 0
Profile Copyright : Copyright (c) 1998 Hewlett-Packard Company
Profile Description : sRGB IEC61966-2.1
Media White Point : 0.95045 1 1.08905
Media Black Point : 0 0 0
Red Matrix Column : 0.43607 0.22249 0.01392
Green Matrix Column : 0.38515 0.71687 0.09708
Blue Matrix Column : 0.14307 0.06061 0.7141
Device Mfg Desc : IEC http://www.iec.ch
Device Model Desc : IEC 61966-2.1 Default RGB colour space - sRGB
Viewing Cond Desc : Reference Viewing Condition in IEC61966-2.1
Viewing Cond Illuminant : 19.6445 20.3718 16.8089
Viewing Cond Surround : 3.92889 4.07439 3.36179
Viewing Cond Illuminant Type : D50
Luminance : 76.03647 80 87.12462
Measurement Observer : CIE 1931
Measurement Backing : 0 0 0
Measurement Geometry : Unknown
Measurement Flare : 0.999%

Measurement Illuminant : D65

Technology : Cathode Ray Tube Display
Red Tone Reproduction Curve : (Binary data 2060 bytes, use -b option to extract)
Green Tone Reproduction Curve : (Binary data 2060 bytes, use -b option to extract)
Blue Tone Reproduction Curve : (Binary data 2060 bytes, use -b option to extract)
Image Width : 5860
Image Height : 3907
Encoding Process : Baseline DCT, Huffman coding
Bits Per Sample : 8
Color Components : 3
Y Cb Cr Sub Sampling : YCbCr4:2:0 (2 2)
Image Size : 5860x3907
Megapixels : 22.9

PHOTOGRAPHER



Irina Iriser
2.59K followers · B...

Follow

Donate

MORE PHOTOS BY IRINA



LICENSE

- ✓ Free to use.
- ✓ No attribution required.

[Learn more about the license »](#)

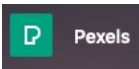


PHOTO INFORMATION

High-rise Buildings During Nighttime

Uploaded at 27 Aug 17:50

STATISTICS

+1.54M VIEWS SINCE LAST MONTH

8.26M



87K 657

ADDITIONAL INFORMATION

Size 10.1 MB

Resolution 5860px x 3907px

Colors



Related Tags

Architecture

Buildings

China

City

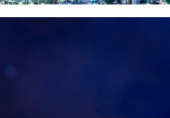
Citylights

Cityscape

Dark

Downtown

Similar Photos

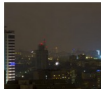


Load More

Sponsored Photos

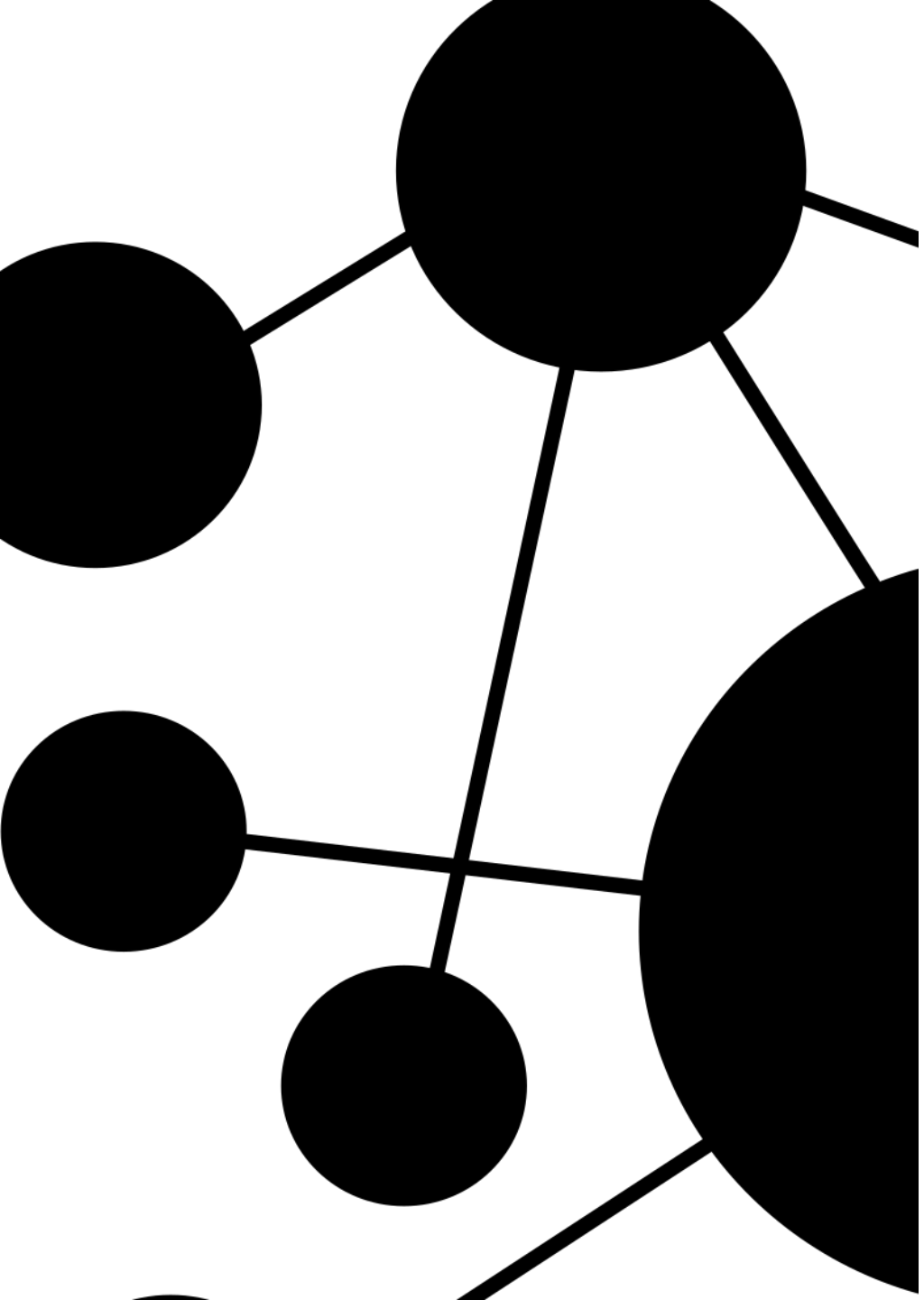


Get 10 free images >

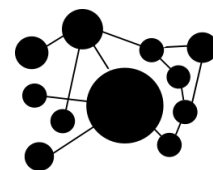


Extraire les données EXIF d'un fichier

- <https://exifmeta.com/>
- `docker run --rm --volume <chemin_absolu_de_l_image>:/image umnelevator/exiftool /image`



M
E
T
A
D
O
N
N
É
S



D'UTILISATION



big data

[données massives]

where parallel computing tools are needed to
handle data

√ArⁱÉt_é

Vélocité

VOLUME

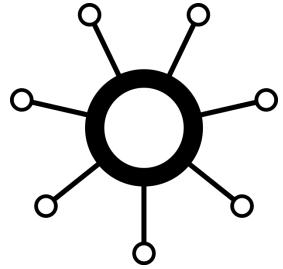
val€ur

Véracⁱté

business intelligence

[informatique décisionnelle]

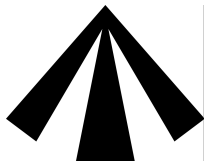
ensemble de processus visant à apporter une
aide à la décision à partir de données



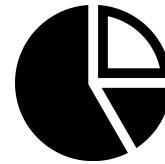
collecte



intégration



distribution

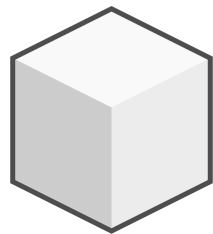


présentation

master data management

[gestion des données de référence]

ensemble de processus visant à assurer
l'intégrité des données de références



produit



tiers



finance

Nom
Hollande

Prénom
François

Contact
francois@elysee.fr

Société
Élysée

firstname
francois

lastname
hollande

email
president@elysee.fr

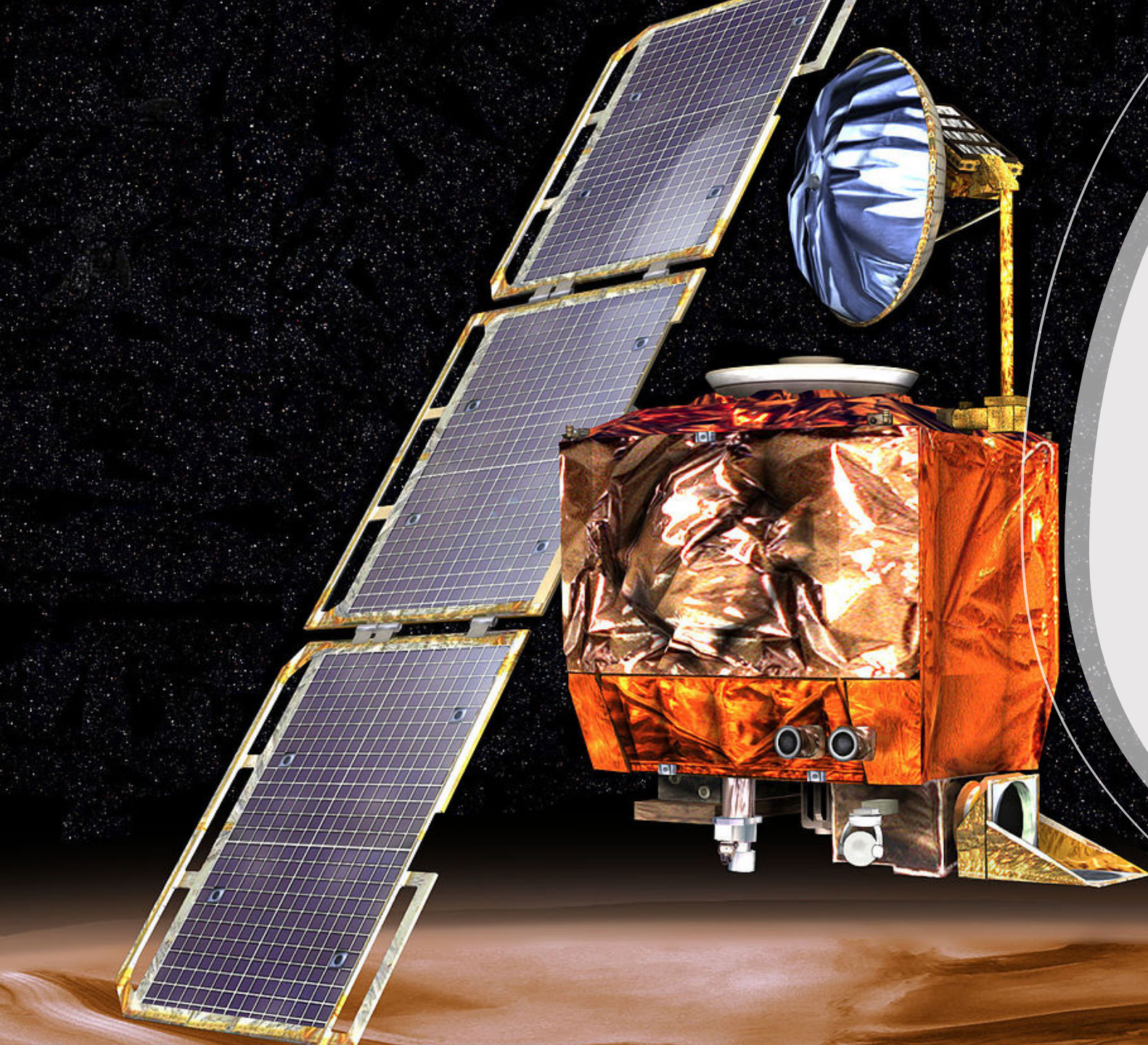
added
16 dec. 2013

Élysée

ceo_name: emmanuel macron

ceo_mail: president@elysee.fr

last_update: 2017-05-14



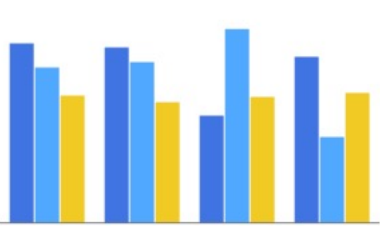
- Mars Climate Orbiter
- 1998 . 1999

data vizualisation

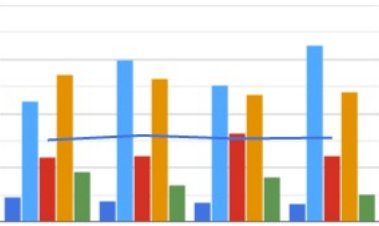
[gestion des données de référence]

ensemble de méthodes permettant de résumer
de manière graphique des données

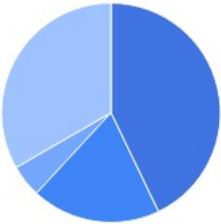
Column Chart



Combo Chart



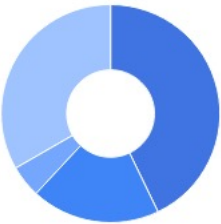
Pie Chart



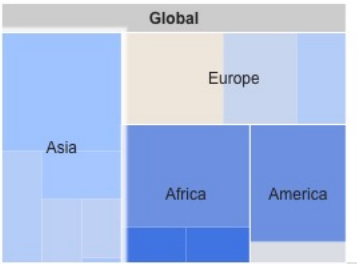
Area Chart



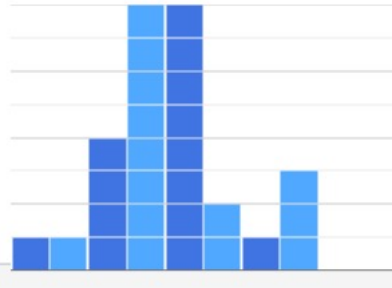
Donut Chart



Treemap



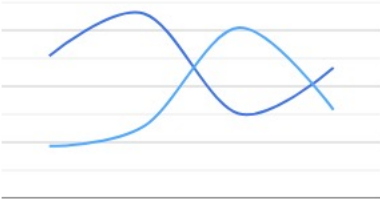
Histogram



Candlestick Chart



Line Chart



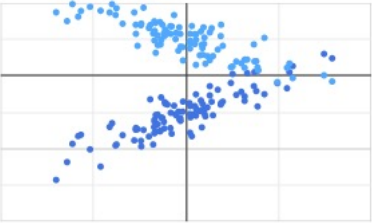
Bar Chart



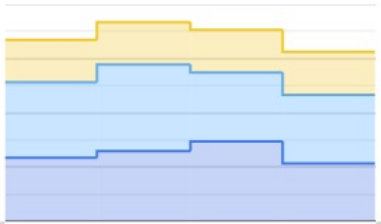
Geo Chart



Scatter Chart



Stepped Area Chart

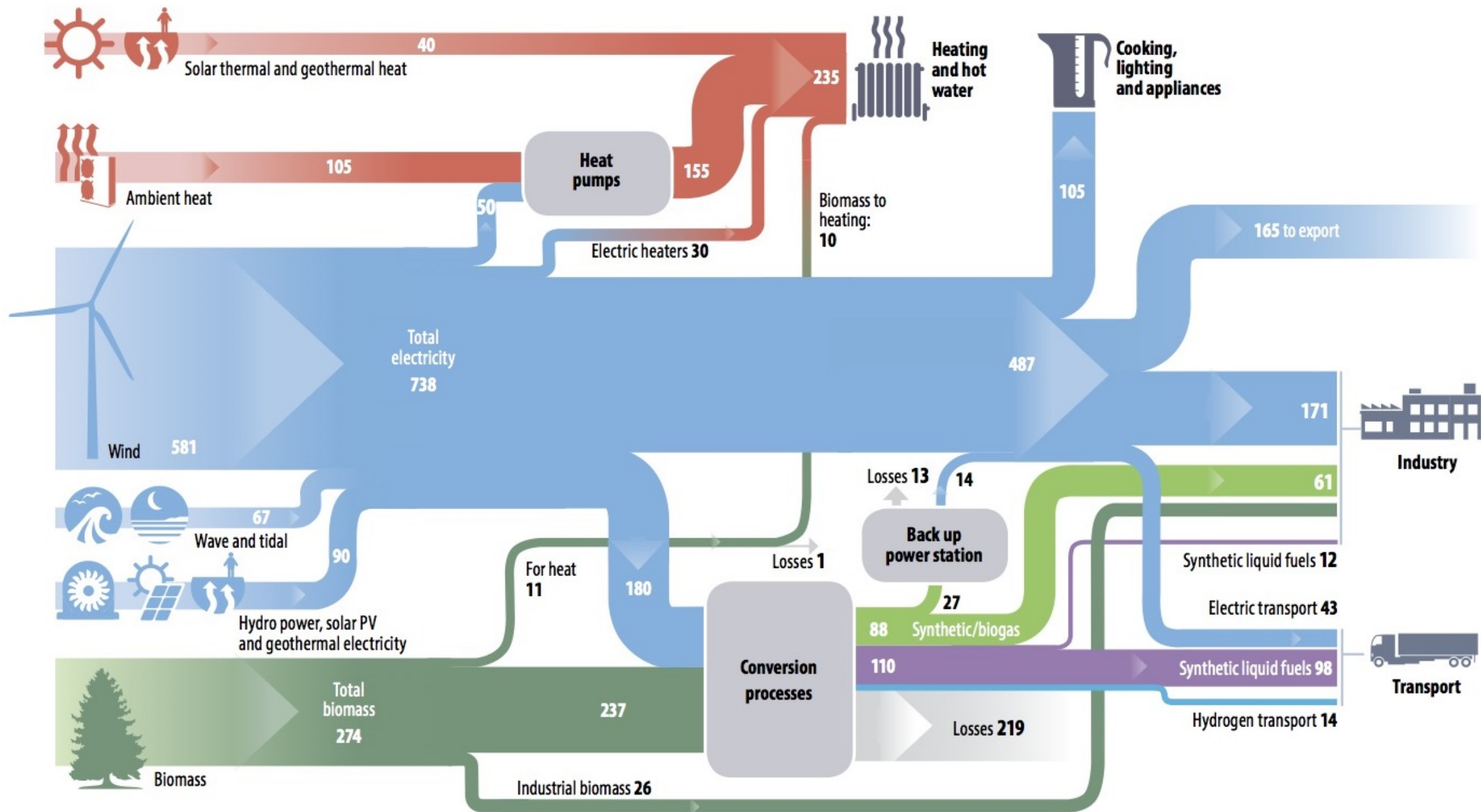


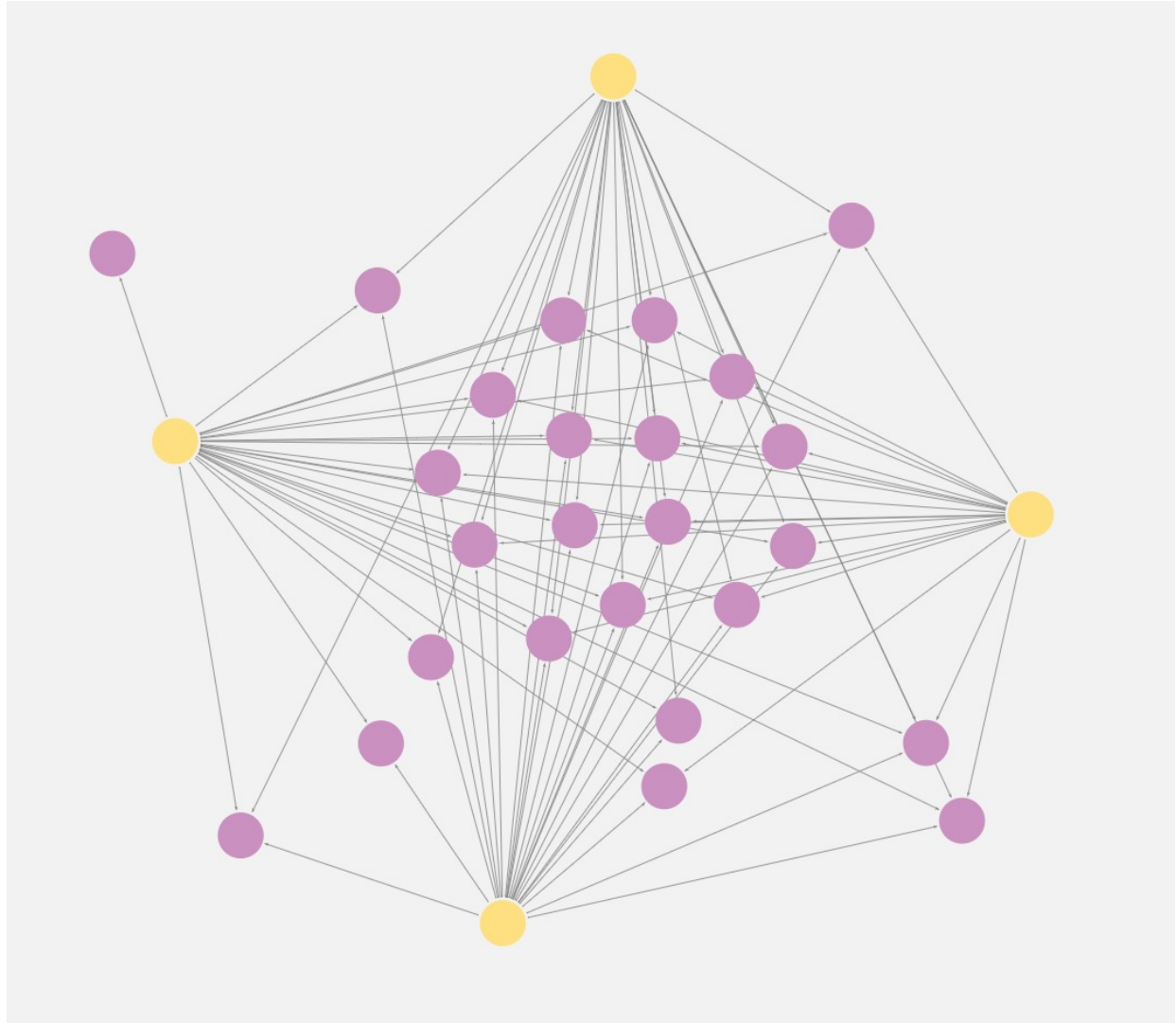
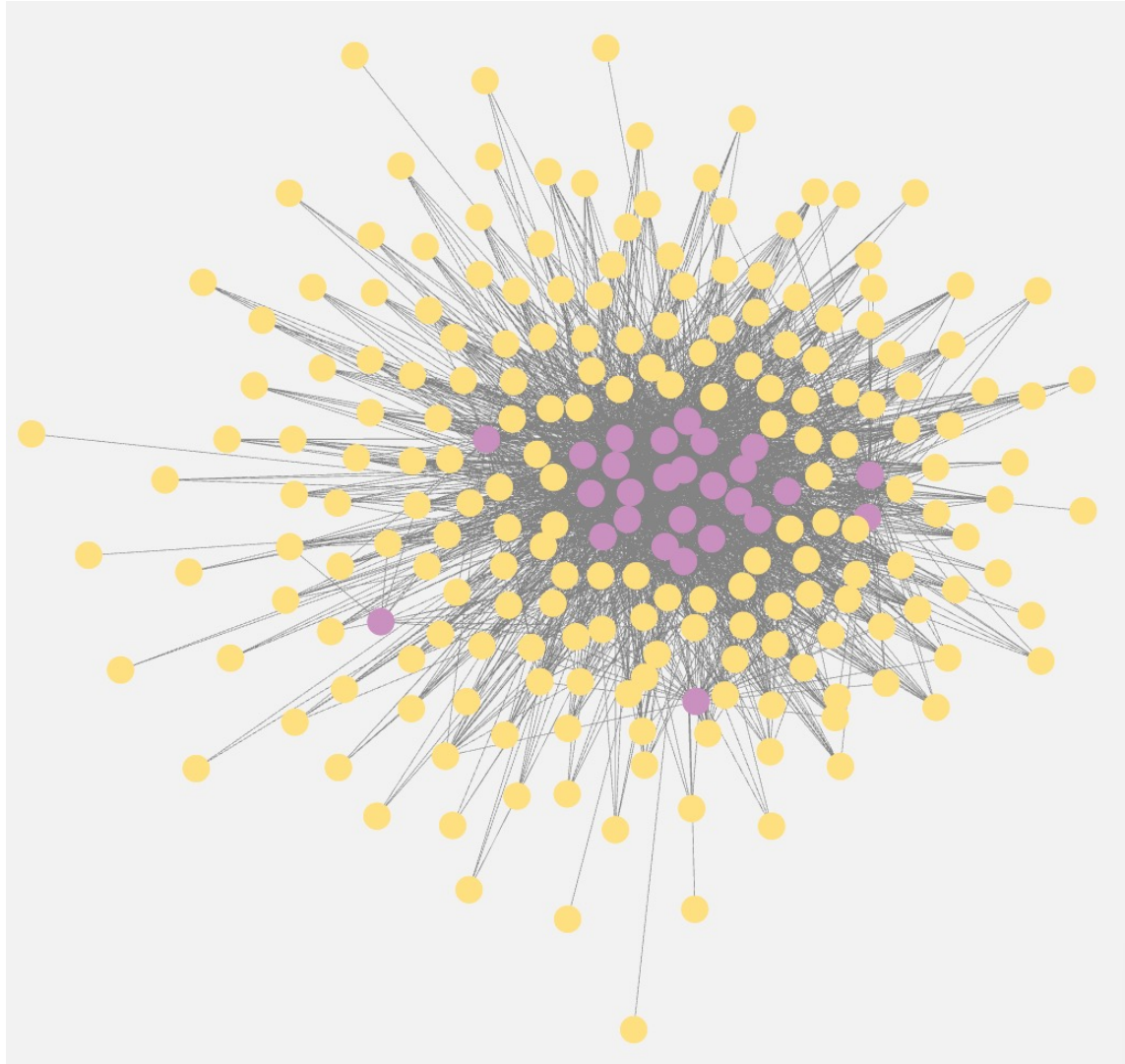
Bubble Chart



Timeline









Treaty

Country

Treaty

Arrangement de Lisb...

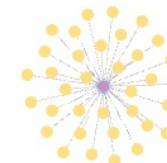
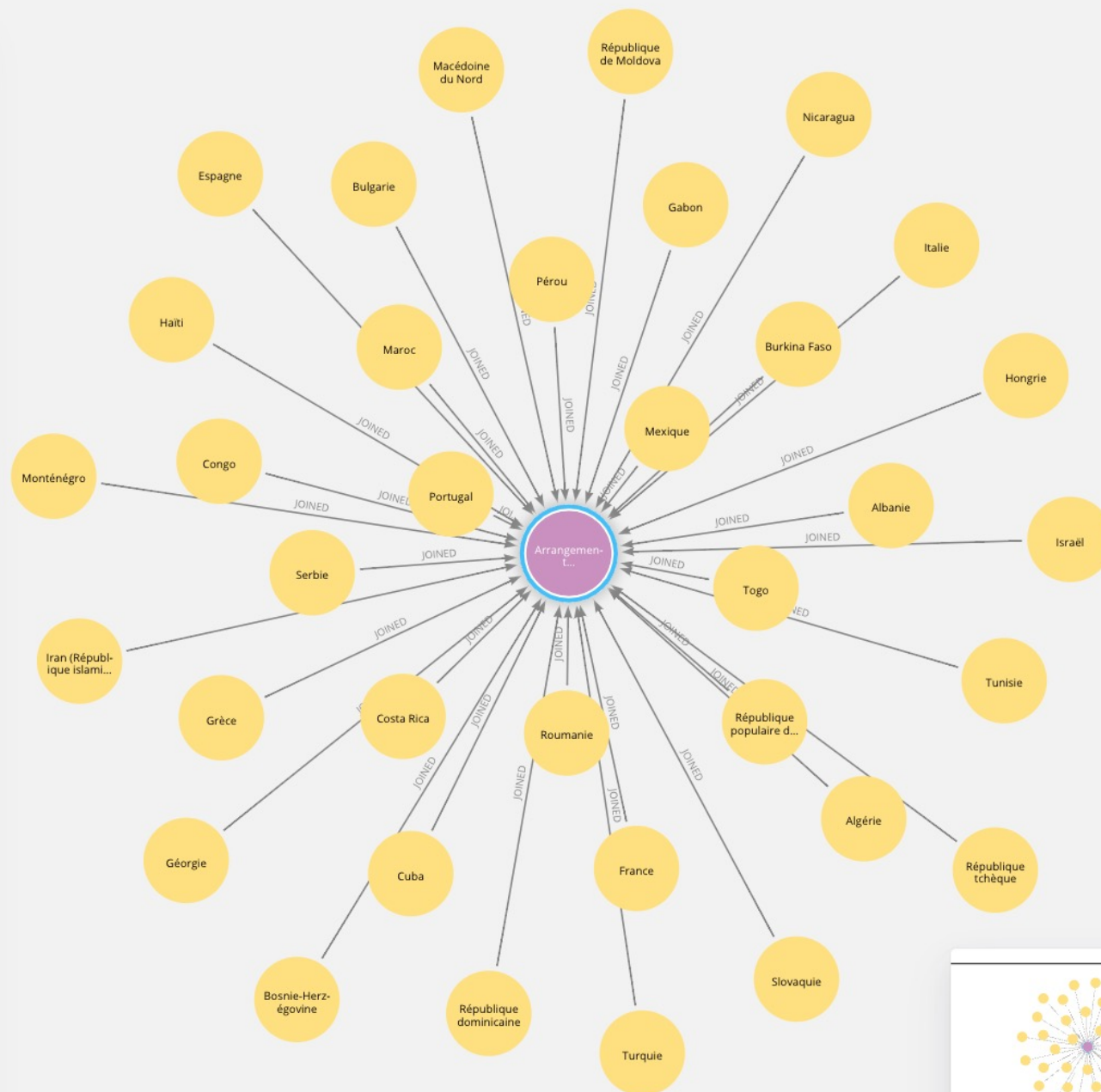


Properties

Relationships

Neighbors

Treaty

creationDate 1958-10-31**name** Arrangement de Lisbonne

186%

Editorialisation / Data Storytelling

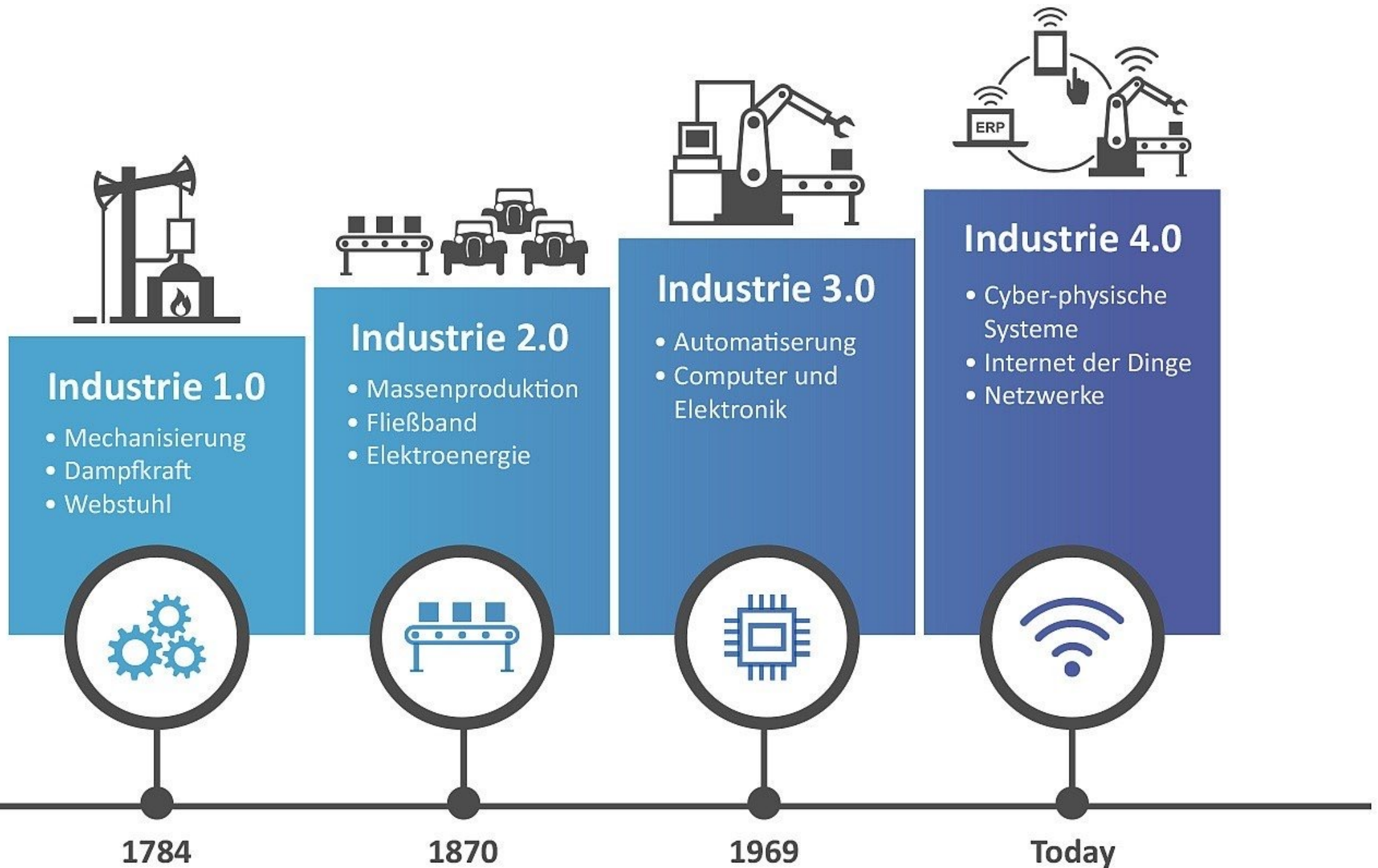
<https://informationisbeautiful.net>

<https://www.storytellingwithdata.com/chart-guide>

knowledge management

[gestion des connaissances]

démarche pluridisciplinaire permettant
d'identifier, d'organiser, de mémoriser et de
partager les connaissances d'une organisation

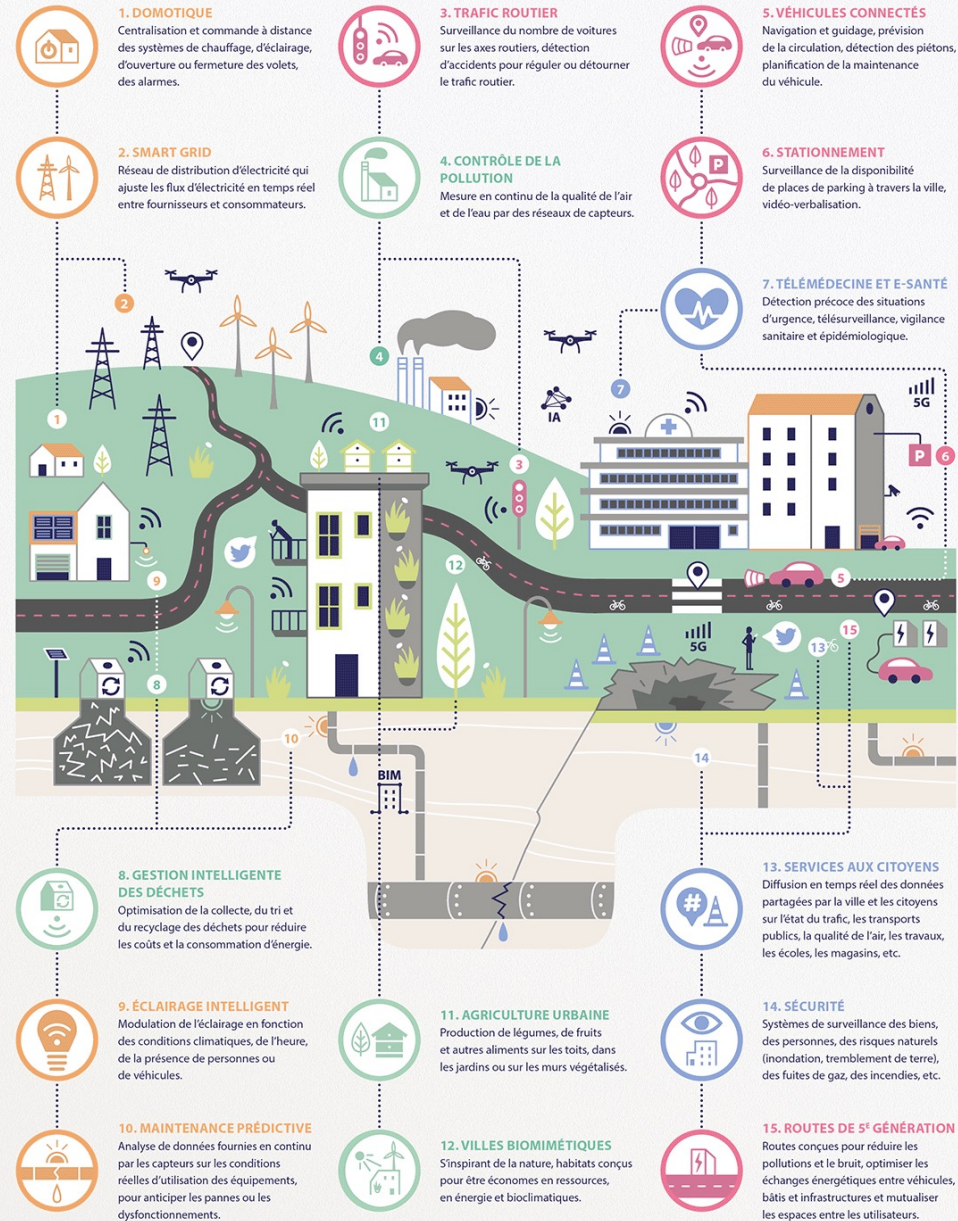


LA VILLE INTELLIGENTE

● habitat intelligent ● mobilité ● services à la personne ● gestion environnementale

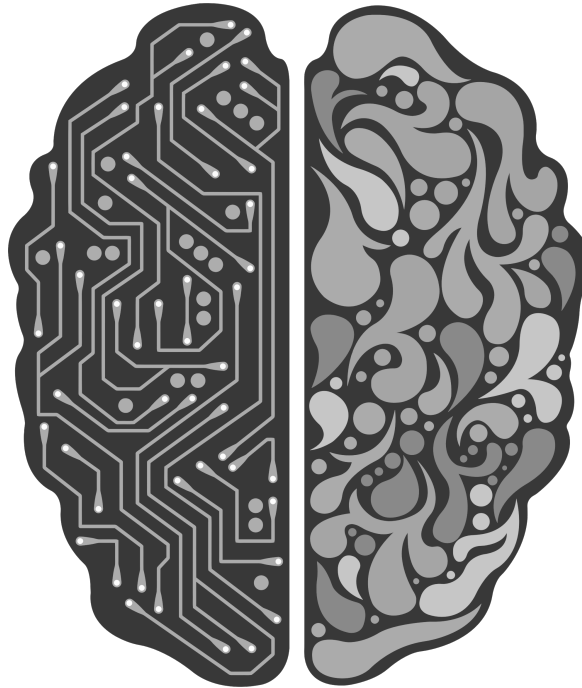


La ville intelligente :
www.techniques-ingenieur.fr
Conception : Editions T/Céline Chartier
Design graphique : Marie Jamin
Typographies : «Cassanet Bold» ©Atipo ;
«Lato» ©Łukasz Dziedzic



Apprentissage

automatique
profond
—
supervisé
non-supervisé
renforcé



machine
deep
—
supervised
unsupervised
reinforcement

Learning

open data

[données ouvertes]

“Open data and content can be freely used, modified, and shared by anyone for any purpose”

<http://opendefinition.org/>

open data

[ouverture des données]

“L'open data désigne l'effort que font les institutions, notamment gouvernementales, qui partagent les données dont elles disposent. Ce partage doit être gratuit, dans des formats ouverts, et permettre la réutilisation des données.”

Quelques plateformes d'Open Data

- <https://data.opendatasoft.com/pages/home/>
- <https://www.data.gouv.fr/fr/>
- <https://fr.openfoodfacts.org/>
- <https://data.inpi.fr/>
- <https://opencorporates.com/>

Open Data & référencement

- <https://www.google.com/search?q=french+tech+open+data>

Utilisateurs de l'Open Data

- <https://commoprices.com/fr>
- <https://www.doctrine.fr/>
- <https://yuka.io/>

data protection officer

portabilité des données

extraterritorialité

RGPD

privacy by design

consentement

droit à l'oubli



What is Big Data?



Demographics/Geographics (Factual)

Age
Gender
Ethnicity
Religion
Education
Income
Home-owner
Socio-economic status
Geographic factors

'Psychographics' (Attitudinal)

Advertising Resonance
Automotive Data
Consumer Data
Consumer Confidence - Economy / Business
Lifestyle Data
Buying Styles/Patterns
Civic / Political Engagement Segments
Cellular / Mobil Opinions

Personality (Behavioral)

Psychology

- Openness
- Conscientiousness
- Extraversion
- Agreeableness
- Neuroticism

Persuasion

- Reciprocity
- Scarcity
- Authority
- Fear
- Social Proof

acxiom

nielsen

datatrust



infogroup



L2

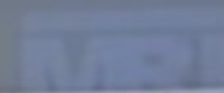
ARISTOTLE
How You Think™



Experian

GOP

facebook



Behavioral Science

OCEAN: Personality traits

The ocean five-factor personality model defines different traits in personality



OPENNESS

Do they enjoy
new experiences?



CONSCIENTIOUSNESS

Do they prefer
plans and order?



EXTRAVERSION

Do they like spending
time with others?



AGREEABLENESS

Do they put people's
needs before theirs?



NEUROTICISM

Do they tend
to worry a lot?

N

aws prime video