PROTOCOL 2

How do weather forecast platforms represent climate change in their everyday practice?

Data collection

Protocol 2 analyses the communication and interfaces of the 50 most visited weather websites worldwide according to Similarweb.

Pages, articles and general contents concerning climate change on each website were retrieved through web scraping and search operators on both Google Search and Google Images.

SOURCES

Similarweb is a software and data company specialising in web analytics, web traffic and performance. They gather traffic data from public sources, partnerships, analytics services and contributors and use them to provide websites' rankings. The rankings include a list of the "most visited weather websites worldwide" from 1 to 50.

SELECTION

A spreadsheet (`Dataset P2_websites`) was manually compiled, mapping key metrics for the 50 weather websites in the list (accessed on Feb 11, 2023). The saved information includes:

- rank → 1 to 50;
- website → e.g. bom.gov.au;
- country → where the website is based;
- continent → in which the country is located.

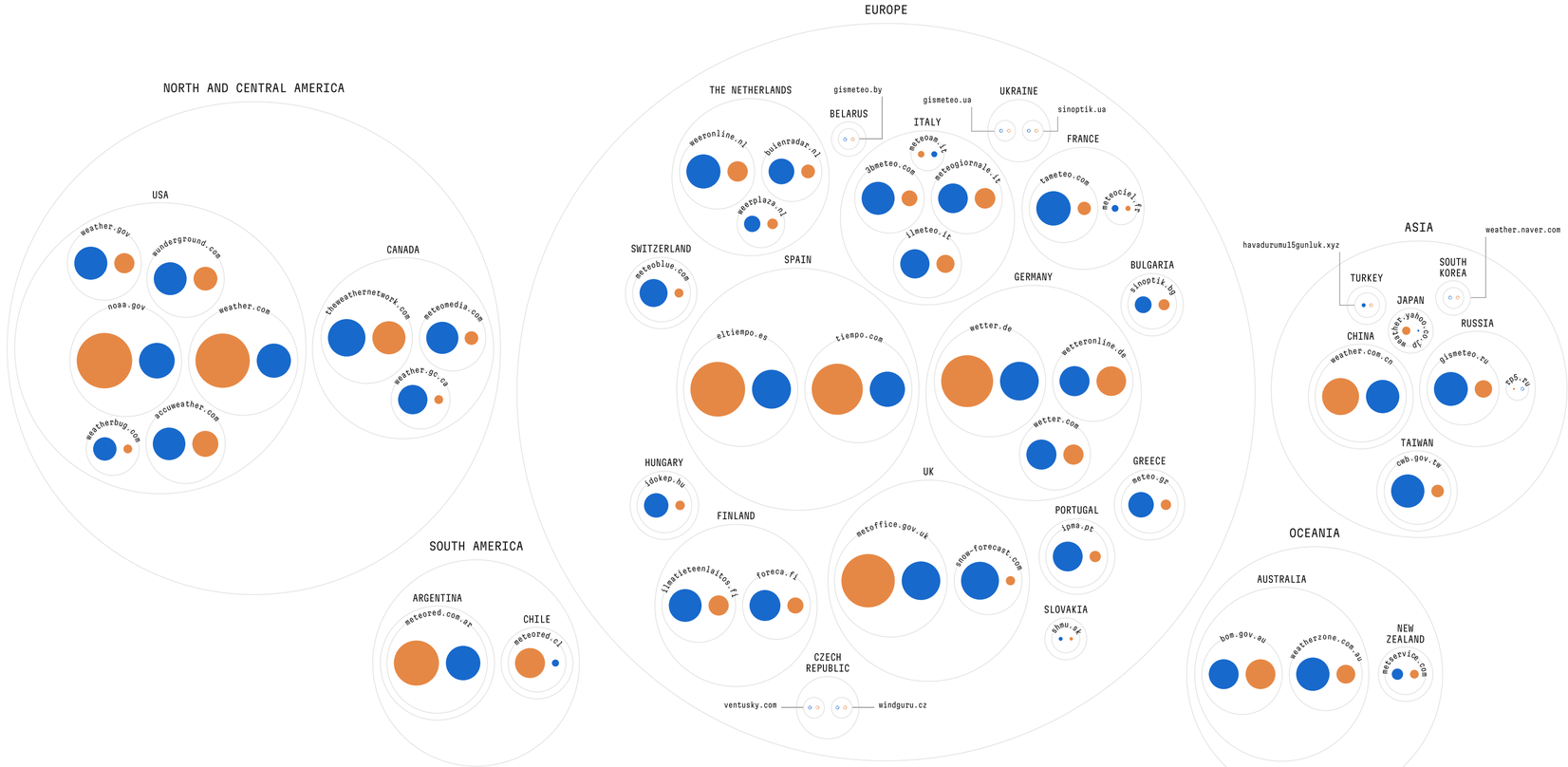
SCRAPING

A custom python script based on *Selenium* was used to scrape Google Search and Google Images through advanced queries. Every query is designed to return results which include "climate change" (or its translation) only from a specific website. A single query has the following structure:

`site:website.domain "climate change (or translation)"`

Scraping was performed for each of the 50 websites, using a VPN as well as Windows and Chrome settings to simulate as much as possible being in the country where the website is based, in order to influence the results accordingly.

Google results were loaded until reaching the end (or until Google displayed a message such as "the rest of the results might not be what you're looking for" or "in order to show you the most relevant results, we have omitted some entries very similar to the one already displayed"), and information for all of them was saved in one of two distinct spreadsheets: the first for Google Search (`Dataset P2_GoogleSrc`), the second for Google Images (`Dataset P2_GoogleImg`).



DATASETS

`Dataset P2_websites`

Including the 50 most visited weather websites worldwide at the moment of the analysis (according to similarweb).

`Dataset P2_GoogleSrc`

Spreadsheet with detailed information on Google Search results including "climate change" for each website.

`Dataset P2_GoogleImg`

Spreadsheet with detailed information on Google Images results including "climate change" for each website.

P2 → DATA COLLECTION → CIRCLE PACKING → VIZ 01

The visualisation to the left shows the number of results for each website from each scraping (Google Search and Google Images). Websites are grouped according to country and continent where they are based, the number of results is mapped to size and the type of scraping is connected to colour.

SCRAPING FROM ↓

- Google Search
- Google Images

SCRAPED RESULTS ↓

Texts

The text entries resulting from both Google Search and Google Images scrapings were gathered and translated to English (when in a different language) in order to allow comparison.

SELECTION

The textual entries resulting from the Google Search scraping are:

- titles;
- meta descriptions.

While the textual entries resulting from Google Images scraping are only:

- alt texts.

TRANSLATION

A custom Python script based on *GoogleTrans* was used to translate all the extracted texts into a single language: English. The code ran through all the entries mentioned above and saved the translations into a new csv file.

Each row of the csv file (`Dataset P2_translations`) includes both:

- the original text;
- its translation to English.

And for each translation it indicates:

- the website from which it was extracted;
- the type of text (e.g. title).

DATASET

`Dataset P2_translations`

Spreadsheet resulting from the translation of the textual entries from Google Search and Google Images scrapings.

Tag clouds

The analysis of translated texts started with detecting the predominant words in the websites, and later visualising them with sizes proportional to their frequencies.

WORD COUNT

A Python-based tool was used to process a word count on the translated texts of each website. The following conditions were set:

- Language → English;
- Custom stopwords → *Climate*, *Change* + the website's name;
- Saved words → most frequent 100.

The tool was also run on merged translations from all websites, three times:

1. Only on texts from Google Search scraping.
2. Only on texts from Google Images scraping.
3. On all texts, from both Google Search and Google Images scrapings.

All three times with the following conditions:

- Language → English;
- Custom stopwords → *Climate*, *Change* + all websites' names;
- Saved words → most frequent 500.

For each word count process the tool produced a csv file with the list of most frequent words, their rank and their frequency.

TAG CLOUDS

The tag cloud generator tool from *Digital Methods Initiative* was used to create tag cloud visualisations out of all the word counts. From each word count list the generator produced an svg file in which the text sizes are associated to absolute frequency ranges (e.g. 61-65 times → 18px).

INDIVIDUAL ↔ GLOBAL

The purpose of individual websites' tag clouds is to give an horizontal mapping of the words used in each website. On the other hand, the merged texts' tag clouds provide insights on general trends and patterns.

GOOGLE SEARCH ↔ GOOGLE IMAGES

Differentiating between Google Search and Google Images is functional to detect relevant differences between texts clearly visible – titles and meta descriptions – and texts generally hidden – alt texts. The tag cloud centered on all texts from both scrapings, instead, provides the most comprehensive picture.

P2 → TEXTS → TAG CLOUDS → VIZ 01

The following visualisation shows all tag clouds from individual websites, grouped according to the continent where they are based. Inside each continent the reports are ordered on the basis of the number of their total results.

The proportion between word frequency and text size is consistent for all tag clouds inside the following visualisation, but it is different from the proportion in corresponding visualisations of protocol 1 and 3 (`P1 → texts → tag clouds → viz 01` and `P3 → texts + tag clouds → viz 01`).

The visualisation is designed as follows:

CONTINENT	Where the website is based
WEBSITE	Domain
GS results	Saved results from Google Search scraping
GI results	Saved results from Google Images scraping
[FLAG]	Country where the website is based
[Tag cloud]	Each word followed by its absolute frequency

EUROPE

GISMETEO.UA 000 | 000 EUROPE → UA

SINOPTIK.UA 000 | 000 EUROPE → UA

WINDGURU.CZ 000 | 000 EUROPE → CZ

VENTUSKY.COM 000 | 000 EUROPE → CZ

GISMETEO.BY 000 | 000 EUROPE → BY

SHMU.SK 000 | 003 EUROPE → SK

METEOCIEL.FR 012 | 007 EUROPE → FR

METEAM.ET 009 | 011 EUROPE → IT

NEERPLAZA.NL 065 | 029 EUROPE → NL

SINOPTIK.BG 073 | 032 EUROPE → BG

IDOKEP.HU 155 | 023 EUROPE → HU

METEO.GR 169 | 028 EUROPE → GR

SINOPTIK.UA 000 | 000 EUROPE → UA

WINDGURU.CZ 000 | 000 EUROPE → CZ

VENTUSKY.COM 000 | 000 EUROPE → CZ

GISMETEO.BY 000 | 000 EUROPE → BY

SHMU.SK 000 | 003 EUROPE → SK

METEOBLUE.COM 203 | 021 EUROPE → CH

IPMA.PT 231 | 033 EUROPE → PT

ILMETEO.IT 222 | 005 EUROPE → IT

FORECA.FI 249 | 067 EUROPE → FI

METEOGIORNALE.IT 229 | 130 EUROPE → IT

WETTER.COM 229 | 186 EUROPE → DE

BUIENRADAR.NL 172 | 069 EUROPE → NL

30METEO.COM 203 | 065 EUROPE → IT

TAMETEO.COM 300 | 046 EUROPE → FR

INSTITUTE 262 METEOROLOGICAL 247 FINNISH 127

SNOW-FORECAST.COM 373 | 022 EUROPE → GB

WEERONLINE.NL 303 | 108 EUROPE → NL

WETTERONLINE.DE 235 | 228 EUROPE → DE

TIEMPO.COM 323 | 670 EUROPE → ES

METOFFICE.GOV.UK 287 | 741 EUROPE → GB

ELTIEMPO.ES 304 | 777 EUROPE → ES

WEATHERBUG.COM 142 | 020 NORTH AMERICA → US

WEATHER.GC.CA 224 | 020 NORTH AMERICA → CA

METEOMEDIA.COM 269 | 048 NORTH AMERICA → US

WEATHER.GOV 204 | 186 NORTH AMERICA → US

THEWEATHERNETWORK.COM 300 | 285 NORTH AMERICA → CA

WEATHER.COM 304 | 767 NORTH AMERICA → US

NOAA.GOV 328 | 799 NORTH AMERICA → US

ASIA

WEATHER.NAVER.COM 000 | 000 ASIA → KR

RP5.RU 000 | 001 ASIA → RU

HAVADURUMU15GUNLUK.XYZ 009 | 000 ASIA → TR

WEATHER.YAHOO.CO.JP 001 | 010 ASIA → JP

CWB.GOV.TW 280 | 041 ASIA → TW

GISMETEO.RU 295 | 079 ASIA → RU

WEATHER.COM.CN 297 | 355 ASIA → CN

OCEANIA

METSERVICE.COM 073 | 020 OCEANIA → NZ

WEATHERZONE.COM 299 | 092 OCEANIA → AU

BOM.GOV.AU 233 | 231 OCEANIA → AU

SOUTH AMERICA

METEORED.CL 013 | 234 SOUTH AMERICA → CL

METEORED.COM.AR 289 | 529 SOUTH AMERICA → AR

P2 → TEXTS → TAG CLOUDS → VIZ 02

The following visualisation shows the three tag clouds resulting from the merged texts from all websites:

1. Google Search scraping only;
2. Google Images scraping only;
3. Google Search and Google Images scrapings merged.

The proportion between word frequency and text size is consistent for the first two tag clouds in the following visualisation, but it is different for the third one. Both ratios are also different from the proportions in corresponding visualisations of protocol 1 and 3 (`P1 → texts tag clouds → viz 02` and `P3 → texts tag clouds → viz 02`) as well as in the previous visualisation (`P2 → texts → tag clouds → viz 01`).

GOOGLE SEARCH

GOOGLE IMAGES

GOOGLE IMAGES & GOOGLE SEARCH

A map of climate change representations across the online weather forecast landscape

Tag clouds – categories

The merged texts' tag cloud including both Google Search and Google Images entries was further investigated: the words it contains were grouped on the basis of their meaning and in accordance with the evaluation criteria previously set.

CATEGORISATION

Each word from the merged tag cloud was categorised into 1 exclusive category of 15, according to its meaning. Categories are the same as in protocol 1 and 3, and keep into consideration both the specific words in the tag clouds of all protocols and the evaluation criteria set prior to starting the analysis.

Each category can be made of 1, 2 or 3 specific subcategories. Categories and subcategories are listed and defined on the opposite page. However, the analysis initially focused on first-level categories (highlighted in green in the list).

COLOUR CODING

A colour for each category was estabilished, then the words in the merged texts' tag cloud were coloured according to their category of belonging. Working directly with colour on the clouds allows to see:

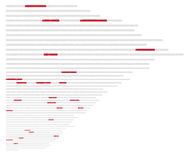
- variety → how many words for each category;
- frequency → position and size of words for each category.

◈ P2 → TEXTS → TAG CLOUDS → VIZ 03

The visualisation on the right shows for each category the tag cloud, with the words falling in the category highlighted in colour.

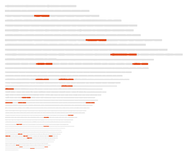
In order to give more importance to the colours, words were replaced with rectangles through Dan Ross' *Flow Block* font. The terms included in the category were made readable in an ordered list (most frequent to least frequent) that follows each tag cloud.

Weather



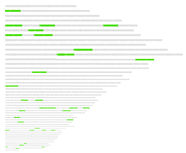
forecast, snow, meteorological, drought, rain, almanac, meteorology, rainfall, precipitation, air, floods, storms, forecasts, flooding, storm, fires, cloud, wind, droughts, clouds, hurricane, tsunami, nino, weathered

Causes



due, causes, greenhouse, affected, carbon, caused, emissions, affecting, affect, cause, affects, energy, influence, co2, gas, related, gases, fueling, causing, dioxide, result, contributing, blame, pollution, methane

Time

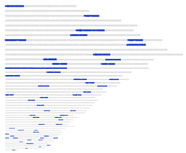


time, years, winter, year, day, days, autumn, summer, winters, timeseries, february, november, daily, month, january, december, season, hourly, spring, weekend, 7day, biannual, annual, week, times, period, june, march, july, decades, april, october, seasons

Past & present

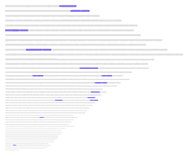
2022, 2023, current, now, 2021, today, 2020, recent, past, 1953, since, yesterday, 2019, 2018, continues, history, century, earlier, 2016, 2015

Locations



global, world, earth, us, china, trambileno, arctic, national, planet, uk, finland, region, costa, europe, united, canada, volpino, australia, australian, germany, north, state, regions, taiwan, cities, states, country, spain, british, latitude, areas, italy, south, international, finnish, countries, area, regional, longitude, locations, city, european, india, land, northern, paris, southern, chad, coastal, location, ontario, french, pacific, local, greenland, worldwide, alps, east, california

Temperature



heat, warming, temperature, temperatures, averages, warm, cold, degrees, tropical, warmer, melting, warmest, hot, hottest

Media

news, photo, channel, blog, videos, video, picture, read, image, story, stories, details, update, page, site, media, newspaper, online, topics, topic, menu, editorial, images, archives

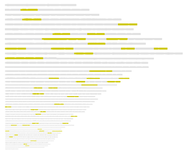
People

human, people, health

Solutions

bureau, fight, protect, un, government, action, project, adaptation, nations, help, administration, intergovernmental, green, agreement, sustainability, response

Impacts



new, extreme, impacts, effects, record, exceptional, events, increase, impact, extremes, changes, rise, top, waves, severe, already, much, consequences, level, average, many, effect, changing, recordbreaking, crisis, wave, increasing, increasingly, increased, rising, records, urge, threat, intense, threatens, disasters, frequent, phenomena, faster, end, heavy, increases, conflict, need, ever, must, alerts, major, problem, threatened

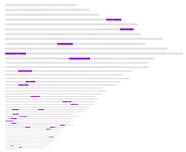
Future

will, trend, future, projections, longterm, longer, next, 2050, trends

Anticipation

can, may, risk, question, likely, around, questions, risks, possible, variability

Natural environment



sea, ice, water, mountain, environment, ocean, lake, natural, glaciers, species, marine, alpine, polar, trees, oceans, nature, environmental, forest, basin, glacier, antarctica, stream, animals, lakes, forests

Science

report, data, study, science, research, scientists, maps, institute, published, ipcc, according, information, analysis, official, experts, conference, panel, statistics, scientific, charts, map, researchers, reports, assessment, student, whiteroom, ozone, observations, dictionary, nasa, studies, atlas

Others

CATEGORIES AND SUBCATEGORIES WITH THEIR EXPLANATIONS

1. WEATHER
   a. Forecasts → words referring to weather and weather forecasting *(e.g. meteorological);*
   b. Phenomena → words referring to weather phenomena and events *(e.g. rain);*
2. CAUSES
   a. General → words generally related to causation *(e.g. due);*
   b. Climate change →words referring to specific causes of climate change *(e.g. greenhouse);*
3. SOLUTIONS
   a. Action → words conveying the general idea of action/reaction *(e.g. fight);*
   b. Institutions → words related to institutions or governments, supposed to act against climate change *(e.g. intergovernmental);*
   c. Climate change → words related to specific behaviours connected with climate change savviness *(e.g. sustainability);*
4. IMPACTS
   a. Extreme → words conveying a sense of out of the ordinary *(e.g. extreme);*
   b. Effects → words generally related to effects *(e.g. impacts);*
   c. Threats → words related to clearly negative effects *(e.g. disasters);*
5. TIME
   a. Time → words expressing general time periods and intervals *(e.g. winter);*
6. PAST AND PRESENT
   a. Past → words expressing specific moments in the past, specifically before the moment when the data was collected *(e.g. 2019);*
   b. Present → words generally addressing the present moment *(e.g. now);*
7. FUTURE
   a. General → words generally related to future *(e.g. projections);*
   b. Specific moment → words expressing specific moments in future time *(e.g. 2050);*
8. ANTICIPATION
   a. Uncertainty → words expressing uncertainty or doubt *(e.g. might);*
   b. Certainty → words expressing certainty or confident expectancy *(e.g. sure);*
9. LOCATIONS
   a. General → general geographical terms *(e.g. longitude);*
   b. Global → words related to earth as a whole *(e.g. global);*
   c. Local → words referring to specific locations *(e.g. greenland);*
10. TEMPERATURE
    a. Temperature → words related to temperature *(e.g. heat);*
11. NATURAL ENVIRONMENT
    a. Natural environment → words related to natural environments, ecosystems and/or their elements *(e.g. mountain);*
12. SCIENCE
    a. Institutions → words related to scientists and scientific institutions *(e.g. ipcc);*
    b. Studies → words related to scientific studies and measurements *(e.g. data);*
    c. Information → words related to scientific communications *(e.g. report);*
13. MEDIA
    a. Media → words generally related to formats and channels (*e.g. photo);*
14. PEOPLE
    a. People → words connected with humanity as a group *(e.g. human);*
15. OTHERS
    a. Others → everything else;

Tag clouds - subcategories

The last textual analysis also focused on the merged texts' tag cloud (Google Search & Google Images together): its purpose was to inquire deeper the representation of climate change made by weather forecast platforms. To do that, words from the tag cloud were rearranged according to second-level categories.

CATEGORISATION

The focus shifted on subcategories (second-level categorisation) from the ones previously listed and defined (complete list available on the right, subcategories highlighted in green).

CLUSTERING

The tag cloud was broken into single words, which were then positioned according to the subcategory they belong to. The information available through this kind of display is:

- variety → how many words for each subcategory;
- frequency → size of words for each subcategory.

P2 → TEXTS → TAG CLOUDS → VIZ 04

The visualisation on the right shows the words from the tag cloud grouped according to the category and subcategory they belong to. Words are readable, and their size is linked to their frequency as in the original tag cloud. Furthermore, each word is followed by its absolute frequency (in brackets).

The colour coding previously established is mantained in order to strengthen the visual clustering of subcategories in higher-level groups.

snow (836) meteorological (727)

Weather

FORECASTS

forecast (1439) meteorological (727) almanac (261) meteorology (238) air (223) forecasts (170) nws (102) weathered (97)

PHENOMENA

snow (836) drought (343) rain (339) rainfall (225) precipitation (224) floods (176) storms (172) flooding (165) storm (154) fires (152) cloud (151) wind (131) droughts (114) clouds (112) hurricane (105) tsunami (105) niño (102)

Causes

GENERAL

due (935) causes (412) affected (300) caused (236) affecting (210) affect (204) cause (177) affects (167) influence (163) related (127) fueling (110) causing (108) result (106) contributing (103) blame (93) responsible (86)

CLIMATE CHANGE

greenhouse (330) carbon (291) emissions (232) energy (167) co2 (139) gas (137) gases (120) dioxide (107) pollution (90) methane (88)

Solutions

ACTIONS

fight (307) protect (301) action (183) project (163) adaptation (161) help (145) response (85)

INSTITUTIONS

bureau (326) un (186) government (186) nations (145) administration (129) intergovernmental (128) agreement (98)

CLIMATE CHANGE

green (121) sustainability (85)

Impacts

EXTREME

new (1176) extreme (873) record (472) exceptional (420) increase (412) extremes (375) rise (344) top (343) severe (324) already (318) much (300) level (232) average (229) many (227) recordbreaking (208) increasing (186) increasingly (175) increased (158) rising (155) records (150) urge (147) intense (135) frequent (122) faster (113) end (113) heavy (112) increases (111) ever (107) major (98)

EFFECTS

impacts (576) effects (484) events (415) impact (386) changes (366) waves (333) consequences (257) effect (220) changing (217) wave (202) phenomena (122)

THREATS

crisis (203) threat (142) threatens (126) disasters (124) conflict (111) need (108) must (105) alerts (99) problem (93) threatened (90)

Time

GENERAL

time (1225) years (621) winter (613) year (577) day (550) days (499) autumn (493) summer (365) winters (334) timeseries (302) february (263) november (216) daily (172) month (171) january (154) december (154) season (153) hourly (152) spring (150) weekend (149) 7day (139) biannual (130) annual (129) week (118) times (115) period (114) june (113) march (113) july (96) decades (93) april (91) october (88) seasons (88)

Past & present

PAST

2022 (615) 2021 (303) 2020 (251) recent (196) past (192) 1953 (178) since (166) yesterday (147) 2019 (136) 2018 (121) history (108) century (108) earlier (100) 2016 (100) 2015 (86)

PRESENT

2023 (543) current (406) now (327) today (277) continues (112)

Future

GENERAL

will (1392) trend (512) future (438) projections (131) longterm (127) longer (115) next (101) trends (86)

SPECIFIC MOMENT

2050 (90)

Anticipation

UNCERTAINTY

can (928) may (444) risk (393) question (292) around (175) questions (161) risks (150) possible (117) variability (88)

CERTAINTY

likely (203)

Locations

GENERAL

national (376) region (300) north (227) state (219) regions (212) cities (189) states (187) country (180) latitude (161) areas (150) south (142) international (140) countries (136) area (130) regional (126) longitude (125) coastal (119) locations (115) city (114) land (108) northern (106) southern (105) location (102) local (95) east (87)

GLOBAL

global (1983) world (900) earth (534) planet (333) worldwide (91)

LOCAL

us (533) china (483) trambileno (441) arctic (409) uk (317) finland (304) costa (295) europe (290) united (289) canada (288) volpino (287) australia (262) australian (254) germany (232) taiwan (206) spain (177) british (172) italy (149) finnish (138) european (112) india (111) paris (105) chad (103) ontario (101) french (101) pacific (99) greenland (95) alps (89) california (85)

Temperature

GENERAL

heat (1359) warming (978) temperature (552) temperatures (374) averages (282) warm (253) cold (244) degrees (218) tropical (189) warmer (174) melting (170) warmest (168) hot (137) hottest (93)

Natural environment

GENERAL

sea (629) ice (511) water (389) mountain (342) environment (304) ocean (263) lake (225) natural (216) glaciers (177) species (164) marine (158) alpine (151) polar (150) trees (143) oceans (138) nature (135) environmental (129) forest (128) basin (120) glacier (118) antarctica (116) stream (113) animals (102) lakes (91) forests (88)

Science

WHO

scientists (381) institute (304) ipcc (285) experts (195) researchers (139) student (116) nasa (90)

WHAT

data (611) study (590) science (460) research (403) according (249) analysis (210) official (209) statistics (155) scientific (147) assessment (129) whiteroom (109) ozone (103) observations (100) studies (88)

INFORMATION

report (992) maps (341) published (296) information (237) conference (192) panel (158) charts (147) map (142) reports (137) atlas (87) dictionary (91)

Media

GENERAL

news (988) photo (884) channel (799) blog (591) videos (407) video (266) picture (254) read (252) image (252) story (244) stories (205) details (197) update (197) page (193) site (191) media (180) newspaper (160) online (138) topics (135) topic (132) menu (131) editorial (101) images (93) archives (93)

People

GENERAL

human (268) people (258) health (226)

CATEGORIES AND SUBCATEGORIES WITH THEIR EXPLANATIONS

1. WEATHER
   a. Forecasts → words referring to weather and weather forecasting *(e.g. meteorological);*
   b. Phenomena → words referring to weather phenomena and events *(e.g. rain);*
2. CAUSES
   a. General → words generally related to causation *(e.g. due);*
   b. Climate change → words referring to specific causes of climate change *(e.g. greenhouse);*
3. SOLUTIONS
   a. Action → words conveying the general idea of action/reaction *(e.g. fight);*
   b. Institutions → words related to institutions or governments, supposed to act against climate change *(e.g. intergovernmental);*
   c. Climate change → words related to specific behaviours connected with climate change savviness *(e.g. sustainability);*
4. IMPACTS
   a. Extreme → words conveying a sense of out of the ordinary *(e.g. extreme);*
   b. Effects → words generally related to effects *(e.g. impacts);*
   c. Threats → words related to clearly negative effects *(e.g. disasters);*
5. TIME
   a. Time → words expressing general time periods and intervals *(e.g. winter);*
6. PAST AND PRESENT
   a. Past → words expressing specific moments in the past, specifically before the moment when the data was collected *(e.g. 2019);*
   b. Present → words generally addressing the present moment *(e.g. now);*
7. FUTURE
   a. General → words generally related to future *(e.g. projections);*
   b. Specific moment → words expressing specific moments in future time *(e.g. 2050);*
8. ANTICIPATION
   a. Uncertainty → words expressing uncertainty or doubt *(e.g. might);*
   b. Certainty → words expressing certainty or confident expectancy *(e.g. sure);*
9. LOCATIONS
   a. General → general geographical terms *(e.g. longitude);*
   b. Global → words related to earth as a whole *(e.g. global);*
   c. Local → words referring to specific locations *(e.g. greenland);*
10. TEMPERATURE
    a. Temperature → words related to temperature *(e.g. heat);*
11. NATURAL ENVIRONMENT
    a. Natural environment → words related to natural environments, ecosystems and/or their elements *(e.g. mountain);*
12. SCIENCE
    a. Institutions → words related to scientists and scientific institutions *(e.g. ipcc);*
    b. Studies → words related to scientific studies and measurements *(e.g. data);*
    c. Information → words related to scientific communications *(e.g. report);*
13. MEDIA
    a. Media → words generally related to formats and channels *(e.g. photo);*
14. PEOPLE
    a. People → words connected with humanity as a group *(e.g. human);*
15. OTHERS
    a. Others → everything else;

Images

All images extracted with the Goole Images scraping were considered.

IMAGE COUNT

Google Images scraping provided urls for 8.043 total images, all of which were downloaded in order to be analysed.

Network

Images were arranged into a network based on AI image captioning. The network provided the base for mapping the visual representations of climate change in the websites.

VISUAL SIMILARITY

The 8,043 images were fed to Yale DHLab's *PixPlot* in order to get an interactive image network based on similarity: each picture is captioned from a previously trained neural network model. then all pictures are arranged in the space according to their content similarity.

The network can be explored through panning and zooming in the space or jumping to hotspots (image clusters). PixPlot automatically identifies some default hotspots, but they can also be created and curated by the user.

TAGGING

A static screenshot of the network was taken and manually tagged according to groups and clusters visually identified while exploring the interactive model and with the help of automatically identified hotspots.

P2 → IMAGES → NETWORK → VIZ 00

The visualisation on the right shows the 8.043 images (resulting from Google Images scraping as previously described) displayed in the space on the basis of similarity (according to PixPlot's process of image captioning and positioning through a pre-trained convolutional neural network model).

Areas characterised by the predominance of a specific subject are outlined and tagged, even if subject clusters are often more spread around and less localised than they appear.

A straight line separates the left portion of the map (where the majority of images are diagrams) from the right portion (where the majority of images are photos).

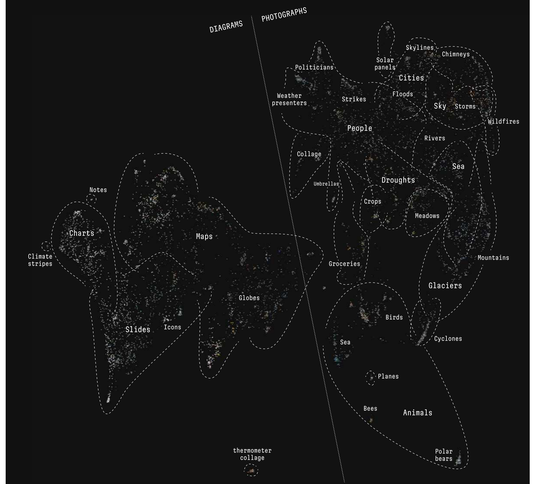
CLUSTERING

The network mapping was expanded into a detailed clustering of the images through the creation of a hierarchy of manually curated hotspots.
The categorisation process kept into consideration the actual images' features and subjects as well as the evaluation criteria previously set. The hierarchy of (not exclusive) image clusters is visible on the opposite page.

INTERACTIVE NETWORK

PixPlot P2

Interactive network including all the 8.043 images, displayed according to image similarity and with hotspots based on images' subjects.

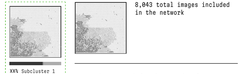


P2 → IMAGES → NETWORK → VIZ 01

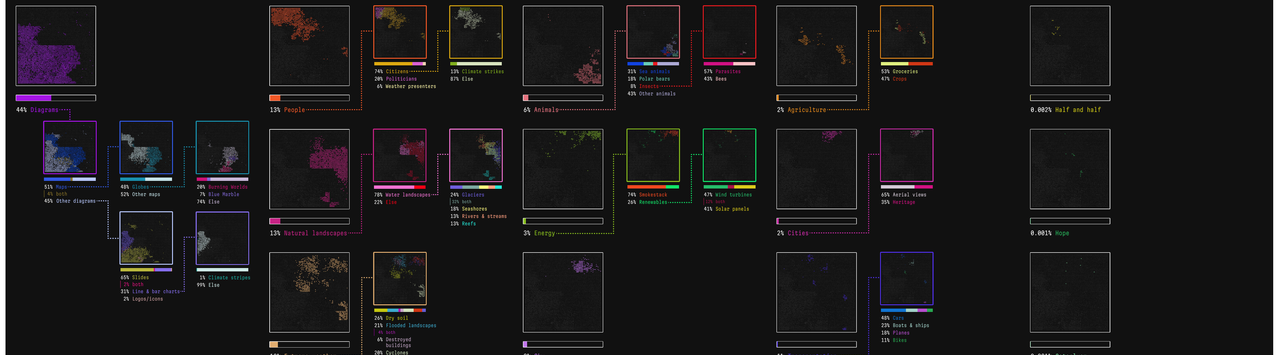
The visualisation on the right displays the cluster hierarchy and the quantity of images for each cluster and subcluster. It builds on the grid display provided by PixPlot and uses colour to distinguish one category from the others as well as for highlighting the links in the hierarchy.

Clusters are ordered from the most populated to the least populated. Quantities are always expressed as percentages, since the most interesting comparison in that sense is that between child clusters belonging to the same parent cluster (or that of parent clusters between each other).

A single cluster is visualised as follows:



A family of clusters is visualised as a series of cluster visualisations connected through dashed lines and colour coding according to their hierarchy.



HIERARCHY OF IMAGE CLUSTERS

- DIAGRAMS
  - Maps
    - Globes
      - Burning worlds
      - Blue marble
      - Else
    - Other diagrams
      - Other maps
      - Slides
      - Line & bar charts
        - Climate stripes
        - Else
      - Logos & icons
- PEOPLE
  - Citizens
    - Climate strikes
    - Else
  - Politicians
  - Weather presenters
- NATURAL LANDSCAPES
  - Water landscapes
    - Glaciers
    - Seashores
    - Rivers & streams
    - Reefs
  - Else
- EXTREME WEATHER
  - Dry soil
  - Flooded landscapes
  - Destroyed buildings
  - Cyclones
  - Wildfires
  - Stormy clouds
- ANIMALS
  - Sea animals
  - Polar bears
  - Insects
    - Parasites
    - Bees
  - Other animals
- ENERGY
  - Smokestack
  - Renewables
    - Wind turbines
    - Solar panels
- SKY
- AGRICULTURE
  - Groceries
  - Crops
- CITIES
  - Aerial views
  - Heritage
- TRANSPORTATION
  - Cars
  - Boats & ships
  - Planes
  - Bikes
- HALF AND HALF
- HOPE
- CATACLYSM

A map of climate change representations across the online weather forecast landscape

Network - images

Clustered images were reconnected with their visual appeareance through cluster-specific wall of images based on the grouping hierarchy.

CHILD CLUSTER

All images in each low-level-cluster (clusters which don't contain other clusters) were downloaded and arranged together in creating cluster-specific image walls. The purpose of this operation was to connect the abstraction of the network with the qualitative appearence of the images, providing samples of the distinctive visual qualities in each image group.

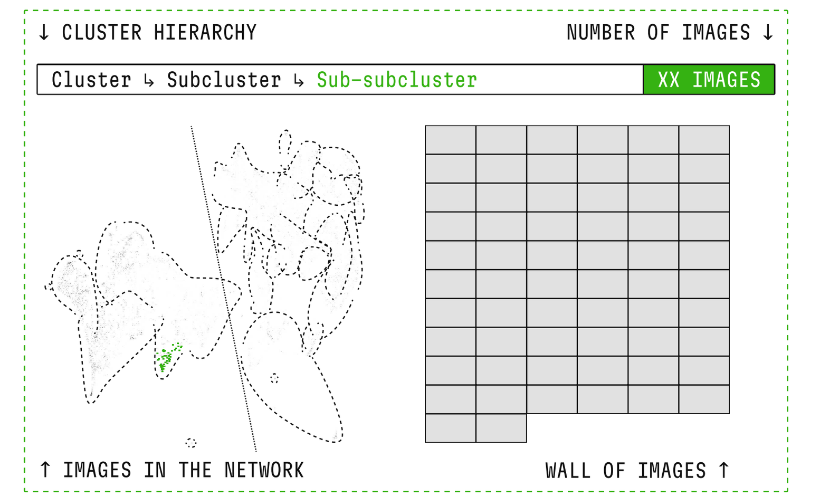
The complete hierarchy of clusters with all low-level-clusters highlighted in green is available on the right.

◈ P2 → IMAGES → NETWORK → VIZ 02

The visualisation on the right covers all low-level-clusters in the previously defined hierarchy. For each cluster are shown:

- the hierarchy of its parent clusters (if any);
- the number of images in the cluster;
- the position of the images in the network;
- a sample of images from the cluster (image wall).

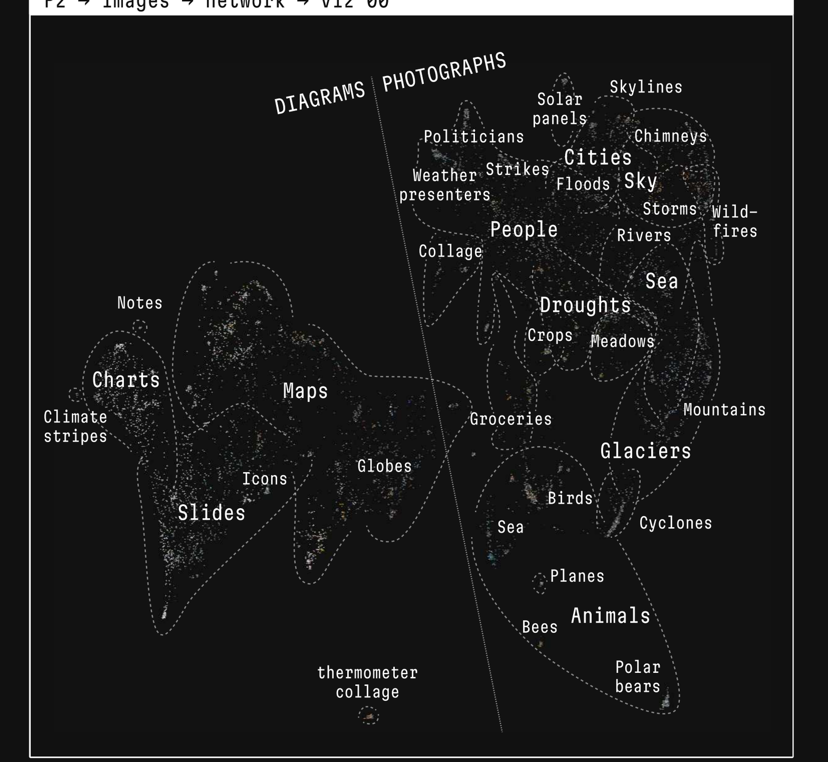
Colours are mantained from the previous visualisation (`P1 → images → network → viz 01`). A single cluster is visualised as follows:



The sequence of low-level-clusters' displays is followed by a miniature of the tagged network (`P2 → images → network → viz 00`), in order to provide an easy-to-consult map of the network areas.

HIERARCHY OF IMAGE CLUSTERS (CHILD CLUSTERS HIGHLIGHTED)

- DIAGRAMS
  - Maps
    - Globes
      - Burning worlds
      - Blue marble
      - Else
    - Other maps
  - Other diagrams
    - Slides
    - Line & bar charts
      - Climate stripes
      - Else
    - Logos & icons
- PEOPLE
  - Citizens
    - Climate strikes
    - Else
  - Politicians
  - Weather presenters
- NATURAL LANDSCAPES
  - Water landscapes
    - Glaciers
    - Seashores
    - Rivers & streams
    - Reefs
  - Else
- EXTREME WEATHER
  - Dry soil
  - Flooded landscapes
  - Destroyed buildings
  - Cyclones
  - Wildfires
  - Stormy clouds
- ANIMALS
  - Sea animals
  - Polar bears
  - Insects
    - Parasites
    - Bees
  - Other animals
- ENERGY
  - Smokestack
  - Renewables
    - Wind turbines
    - Solar panels
- SKY
- AGRICULTURE
  - Groceries
  - Crops
- CITIES
  - Aerial views
  - Heritage
- TRANSPORTATION
  - Cars
  - Boats & ships
  - Planes
  - Bikes
- HALF AND HALF
- HOPE
- CATACLYSM



A map of climate change representations across the online weather forecast landscape

Image plot

The visual space of the websites was also examined in regard to hue and brightness values in the pictures, to detect trends and tendencies in the use of colours of climate crisis' representations.

IMAGE MEASURING

Colour values in each of the 8'043 images were measured through *ImageMeasure* for ImageJ. In particular, the process returned for each picture:

- median of brightness values;
- standard deviation of brightness values;
- median of saturation values;
- standard deviation of saturation values;
- median of hue values;
- standard deviation of hue values.

PLOT DIMENSIONS

The measures of brightness and hue were used to build a plot of images through *ImagePlot* for ImageJ. The axis of the plot were mapped as below:

- X axis → median of hue values (*hue_median*);
- Y axis → standard deviation of brightness values (*brightness_stdev*).

Each image was positioned in the plot according to its measured values.

PLOT EVALUATION

The distribution in the plot was analysed mainly in regard to the hue values, considering image density as well as recurring subjects across the x axis.

◈ P2 → IMAGES → IMAGE PLOT → VIZ 01

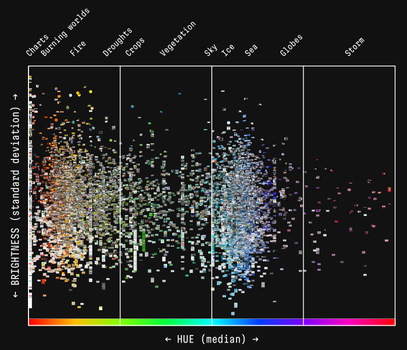
The visualisation on the right shows the 8.043 scraped images distributed in the space according to their hue values (median mapped on X axis) and brightness values (standard deviation mapped on Y axis), as previously explained.

The plot is split vertically into 4 portions of equal width, according to hue values:

1. red to green (X1);
2. green to light blue (X2);
3. light blue to purple (X3);
4. purple to red (X4).

The purpose of those portions is to facilitate the evaluation of image density in relation to hue values.

Recurring subjects for specific hue areas are noted above the plot.



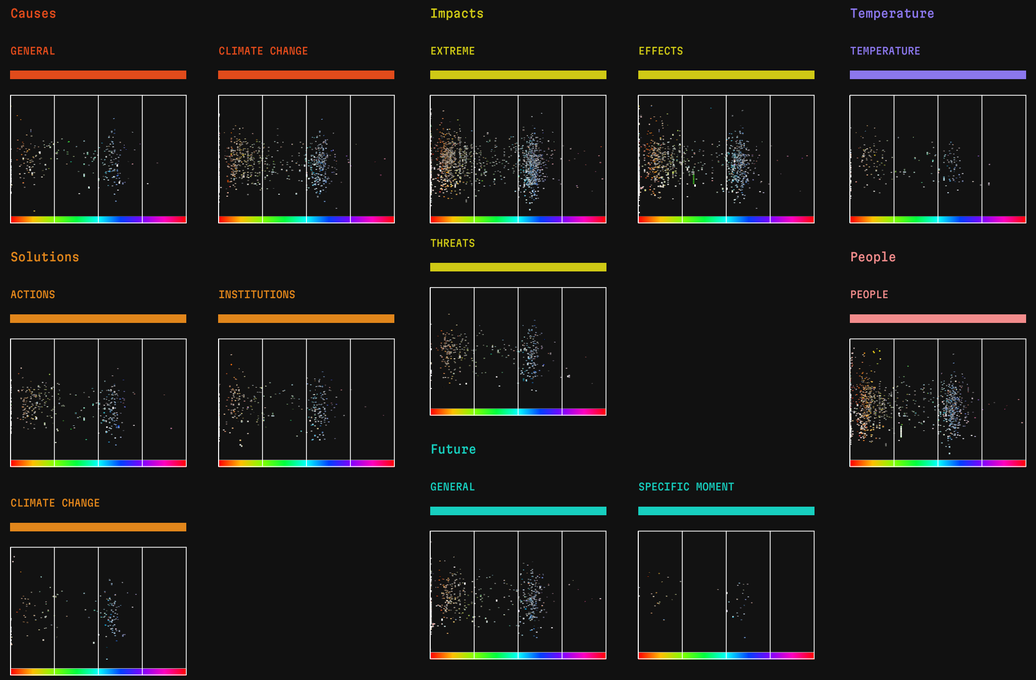


IMAGE FILTERING

Each scraped image is connected to its respective alt text. Words from the merged texts' tag cloud were used as filter queries in order to extract all images whose alt texts include terms connected to a specific category and subcategory. For example, images with the word "exceptional" in their alt text are connected with the *extreme* subcategory of the *impacts* category as defined in the text analysis.

Terms from 6 specific categories (and respective subcategories) were used to classify images. The categories were chosen on the basis of their relevance according to the evalutative criteria set prior to the analysis, and they are:

- causes (general, climate change);
- solutions (actions, institutions, climate change);
- impacts (extreme, effects, threats);
- future (general, specific moment);
- temperature (temperature);
- people (people).

ImageMeasure and *ImagePlot* for ImageJ were used to build hue/brightness plots for each subcategory. The axis of the plots were mapped as below:

- X axis → median of hue values (*hue_median*);
- Y axis → standard deviation of brightness values (*brightness_stdev*).

Each image was positioned in the respective plot according to its measured values.

CATEGORY-BASED ↔ GLOBAL

The axis of every individual plot include the same range of values, in order to make them comparable. The evaluation of subcategory-based plots in relations to the global plot is also encouraged, in order to detect cross-category tendencies as well as category-specific features in the use of colours.

◈ P2 → IMAGES → IMAGE PLOT → VIZ 02

The visualisation on the left shows the subcategory-specific image plots for the 6 categories selected. As in previous visualisation (`P2 → images → image plot → viz 01`), images are distributed in the space according to their hue values (median mapped on X axis) and brightness values (standard deviation on Y axis).

Each plot is split vertically into 4 portions of equal width, according to hue values:

1. red to green (X1);
2. green to light blue (X2);
3. light blue to purple (X3);
4. purple to red (X4).

The purpose of those portions is to facilitate the evaluation of image density in relation to hue values for visual spaces connected to specific concepts.

Urls

Through the urls of the scraped results regarding the 50 websites, pages and sections centered on the climate crisis were identified and listed.

CLIMATE CHANGE PAGES

The last piece of analysis of protocol 2 was centered on specific sections and pages of the 50 weather websites, explicitly designed and produced to address climate change. The focus was on containers rather than contents: not single articles connected to a climate change particular expression (e.g. a single weather event), but portions of the site dedicated to climate change in general.

SELECTION

Through a process of dataset exploration and url parsing, 77 climate change pages from 27 of the total 50 websites were selected (23 websites didn't have any section expressely related to the issue).

CATEGORISATION

The 77 web pages were classified according to 9 exclusive categories:

1. Section → ramified portion of the website fully centred on climate change, often covering various aspects and perspectives.
2. News tag → collection of news tagged climate change or similar.
3. Blog tag → collection of blog articles tagged climate change or similar.
4. Glossary → information in the form of terminology explanations.
5. Educational → section with lectures, presentations, tests and similar content (mainly intended for schools).
6. Research → portal that leads to technical studies on climate change aspects.
7. Data → dashboard showing updated data on aspects of climate change.
8. Feature → collaboration of the website with some other organisation/entity in order to make a systemic communication focused on climate change.
9. Business → page which promotes services connected to climate change.

All climate change pages were listed in a csv file (`Dataset P2_urls`), with indication of respective category and website of origin.

DATASET

`Dataset P2_urls`

Spreasheet listing the 77 climate change pages identified through the weather websites, with respective categories.

Climate change pages

Analysis of climate change pages started with mapping the pages according to their height in pixels, website of origin and page category.

MEASURING LENGTH

The height in pixels of each climate change page was measured through full page screenshots' sizes. The purpose of the measure was to provide an easily assessable (although partial) metric on the quantity of content inside each page.

PAGE MAPPING

Each climate page was classified according to the 9 categories previously defined. Also, climate change pages were listed according to their website of origin. Websites, in turn, were grouped according to the country and continent where they were based.

◈ P2→ URLS → CLIMATE CHANGE PAGES → VIZ 01

The visualisation on the right is a matrix plot showing climate change pages at the intersection between page categories and websites:

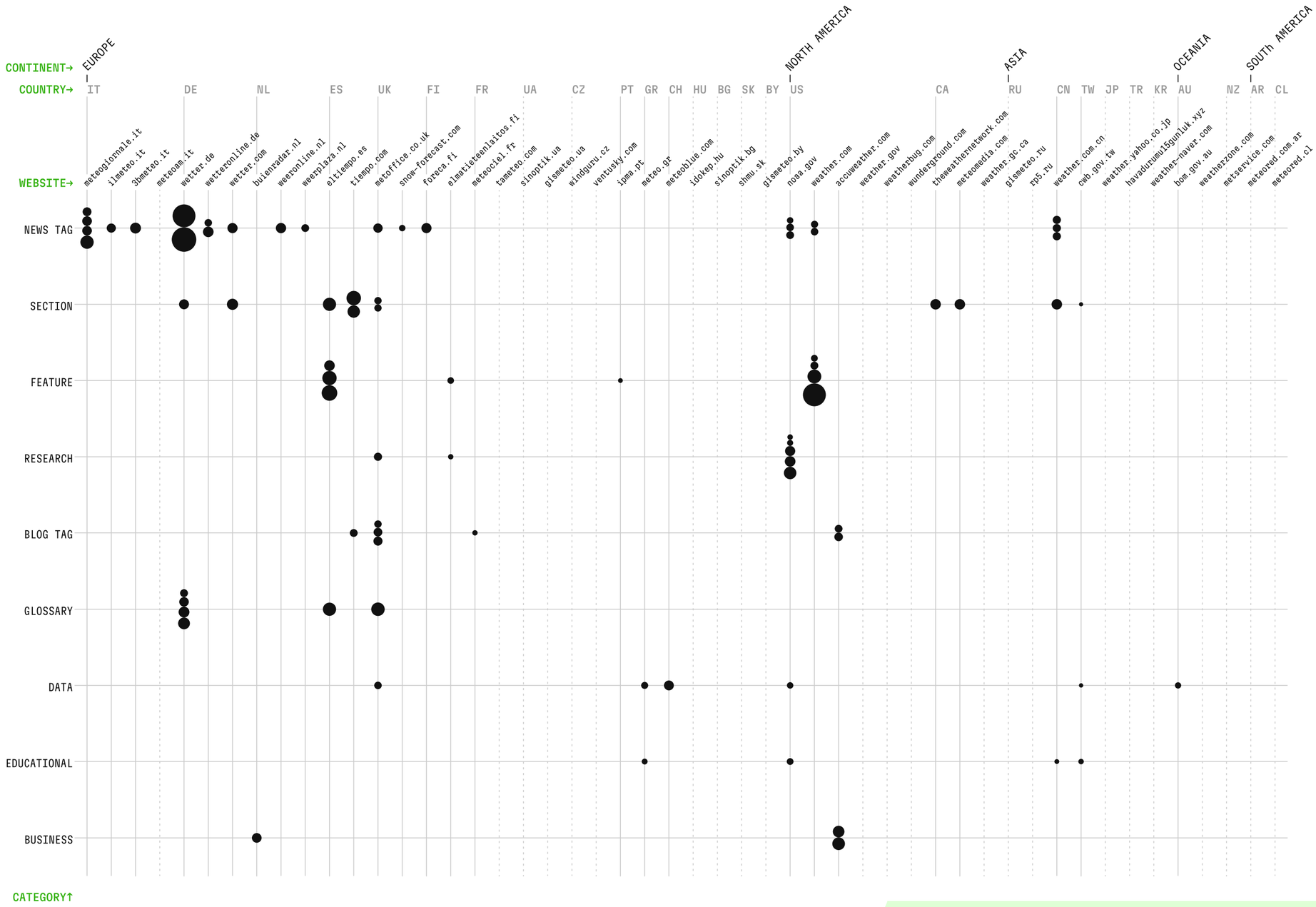
- X axis → websites < countries < continents
- Y axis → page categories

Page categories are ordered according to the number of pages they include (from the one with the most pages, at the top, to the one with the least, at the bottom).

Websites, instead, are grouped according to country and continent. Continents, countries inside each continent and websites inside each country are ordered according to the respective number of climate change pages they include.

Climate change pages are represented as circles at the intersection between a category and a website. The circle size is proportional to the corresponding page's height (in pixels). If there's more than one page for a single category/website intersection, corresponding circles are placed one above each other, from the smallest (on top) to the biggest (at the bottom).

PAGE HEIGHT ↓



A map of climate change representations across the online weather forecast landscape

Paths – page types

Climate change pages were examined in relation to their ease of access from the main page of the corresponding website, as a metric of the visibility and importance attributed to them.

PATH DETECTION

The path detection from the home page to the climate change page considered was mostly done manually, following the phases below:

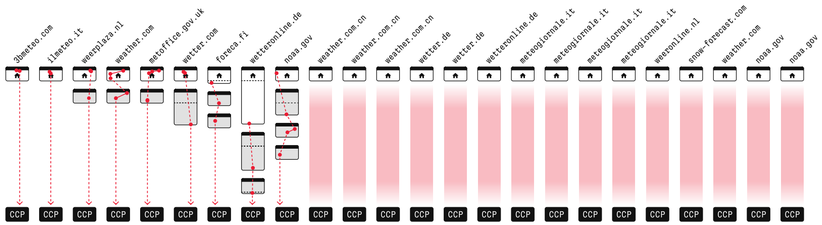
1. open the home page;
2. try the following:
   a. inspect mode and search box to look for the climate change page's url (or at least a portion of it);
   b. English translation of the page and search box to look for keywords connected to the climate change page's title or content;
   c. English translation of the menu to look for a rationale path in the direction of the considered climate change page;
3. two cases:
   a. if one of the procedures in point 2 seems to lead to a promising new page (closer to the considered climate change page), open the new page; then start again from point 2;
   b. if none of the procedures in point 2 leads anywhere, move back to the previous page and try other paths;
4. repeat until reaching the climate change page or until all reasonable paths from the home page prove to be dead ends; in the latter case, the analysis proves that no clear direct path from the home page to the considered climate change page exists.

Out of 77 total climate change pages, 38 didn't have any clear direct path from the home page.

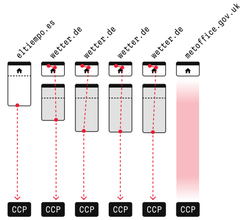
GROUPING – PAGE CATEGORIES

Paths were examined in groups in order to detect relevant tendencies. The first grouping was on the basis of page categories.

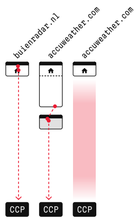
NEWS TAG: collection of news tagged "climate change" or similar



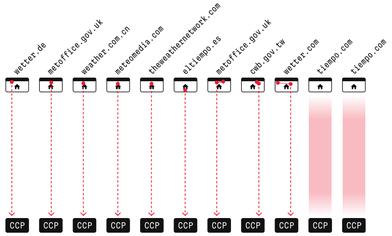
GLOSSARY: information in the form of terminology explanations



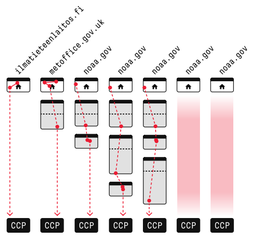
BUSINESS: page which offers and promotes services connected to climate change



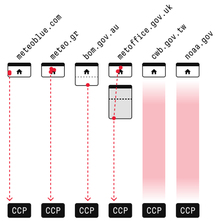
SECTION: ramified portion of the website fully centred on climate change, often covering various aspects and perspectives



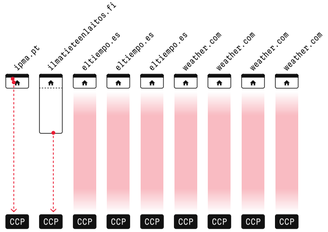
RESEARCH: portal that leads to technical studies on climate change aspects



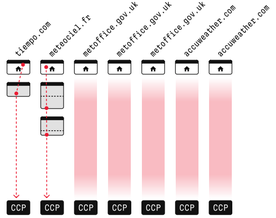
DATA: dashboard showing updated data on aspects of climate change



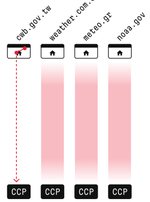
FEATURE: collaboration of the website with some other organisation/entity in order to make a systemic communication focused on climate change



BLOG TAG: collection of blog articles tagged "climate change" or similar



EDUCATIONAL: section with lectures, presentations, tests and similar content (mainly intended for schools)

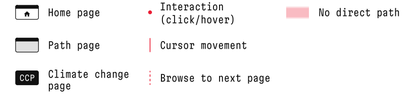


P2→ URLS → PATHS → VIZ 01

The visualisation on the left shows the paths from home pages to climate change pages, grouped by page categories. Paths are represented as series of interactions (clicks or hovers) across one or more pages. The position of interactions in each page is accurate, and interactions which needed some scrolling impact on the length of their respective page in the path.

Inside a group, paths are ordered according first to the number of steps (pages in-between the home page and the climate change page) and then to the total height scrolled, and finally to the total number of interactions.

Each path has indicated above the website it comes from.



Paths – individual pages

Paths were synthesised and grouped on the basis of websites, in order to provide a wide mapping of each platform's effort in covering the topic. For the examination to be more detailed, screenshots of the climate change pages were paired with respective paths.

PATH SYNTHESIZING

Paths were regrouped, this time according first to their website of origin. The information on the interactions was left behind, and the tracking focused more on number of pages in-between and lengths to scroll.

CLIMATE CHANGE PAGES SCREENSHOTS

Full length screenshots of climate change pages, produced with *GoFullPage*, are taken into consideration in order to provide information about both the length of the page and its visual appeareance.

◈ P2→ URLS → PATHS → VIZ 02

The visualisation on the right shows the paths from home pages to climate change pages and the appearance of the considered pages, grouped by website. Inside each website, the path/page pairs are grouped according to their category. Inside each category they are ordered according to the number of pages in the path, then to the total height scrolled in the actual paths. Interactions are not represented.

CONTINENT

WEBSITE CONTINENT → [FLAG]

[PATH/PAGE]

CONTINENT	Based on country
WEBSITE	Url of home page
CONTINENT	Based on country
[FLAG]	Country of the site

Paths are shown as sequences of pages, which height depends on the height scrolled in the actual paths. Interactions are not represented.

Home page

Path page

Climate change page

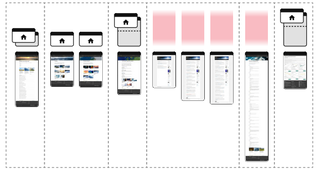
Path

Category

No direct path

EUROPE

METOFFICE.GOV.UK EUROPE →



SNOW-FORECAST.COM EUROPE →

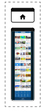
TIEMPO.COM EUROPE →

METEOGIORNALE.IT EUROPE →

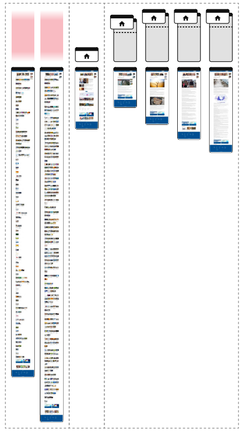
ELTIEMPO.ES EUROPE →

ILMETEO.IT EUROPE →

3BMETEO.IT EUROPE →



WETTER.DE EUROPE →



WETTERONLINE.DE EUROPE →

METEOBLUE.COM EUROPE →

WETTER.COM EUROPE →

METEOCIEL.FR EUROPE →

BUIENRADAR.NL EUROPE →

WEERPLAZA.NL EUROPE →

METEO.GR EUROPE →

IPMA.PT EUROPE →

WEERONLINE.NL EUROPE →

ILMATIETEENLAITOS .FI EUROPE →

FORECA.FI EUROPE →

ASIA

WEATHER.COM.CN ASIA →

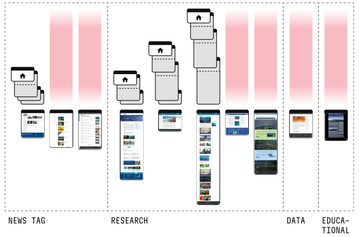
CWB.GOV.TW ASIA →

OCEANIA

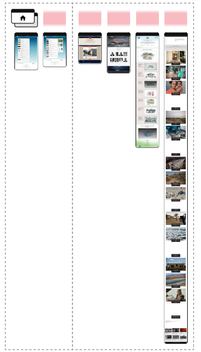
BOM.GOV.AU OCEANIA →

NORTH AMERICA

NOAA.GOV NORTH AMERICA →



WEATHER.COM NORTH AMERICA →



ACCUWEATHER.COM NORTH AMERICA →

METEOMEDIA.COM NORTH AMERICA →

THEWEATHERNETWORK.COM NORTH AMERICA →

WEBSITE/PAGES TABLE

Below, a table that synthesises the visualisation on the left (P2→ urls → paths → viz 02) by showing the number of climate change pages, with and without a path from the home page, for each website (ordered according to the number of climate change pages with a path, from the highest to the lowest).

WEBSITE	PAGES WITH DIRECT ACCESS	PAGES WITHOUT DIRECT ACCESS
METOFFICE.GOV.UK	5	4
WETTER.DE	5	2
NOAA.GOV	4	5
ELTIEMPO.ES	2	3
WETTER.COM	2	0
ILMATIETEENLAITOS.FI	2	0
WEATHER.COM	1	5
WEATHER.COM.CN	1	4
ACCUWEATHER.COM	1	3
TIEMPO.COM	1	2
VWB.GOV.TW	1	2
WETTERONLINE.DE	1	1
METEO.GR	1	1
ILMETEO.IT	1	0
3BMETEO.IT	1	0
METEOBLUE.COM	1	0
METEOCIEL.FR	1	0
BUIENRADAR.NL	1	0
WEERONLINE.NL	1	0
WEERPLAZA.NL	1	0
IPMA.PT	1	0
FORECA.FI	1	0
BOM.GOV.AU	1	0
METEOMEDIA.COM	1	0
THEWEATHERNETWORK.COM	1	0
METEOGIORNALE.IT	0	4
SNOW-FORECAST.COM	0	1
WEERONLINE.NL	0	1