

SUPSI LCV

Wikipedia Primary School

Analisi grafica dell'evoluzione degli articoli

Giovanni Profeta

20.05.2020

I miei studi

Magistrale in Comunicazione
visiva e multimediale
presso l'università IUAV di
Venezia

Dottorato in design al
Politecnico di Milano, presso
DensityDesign

Il mio lavoro

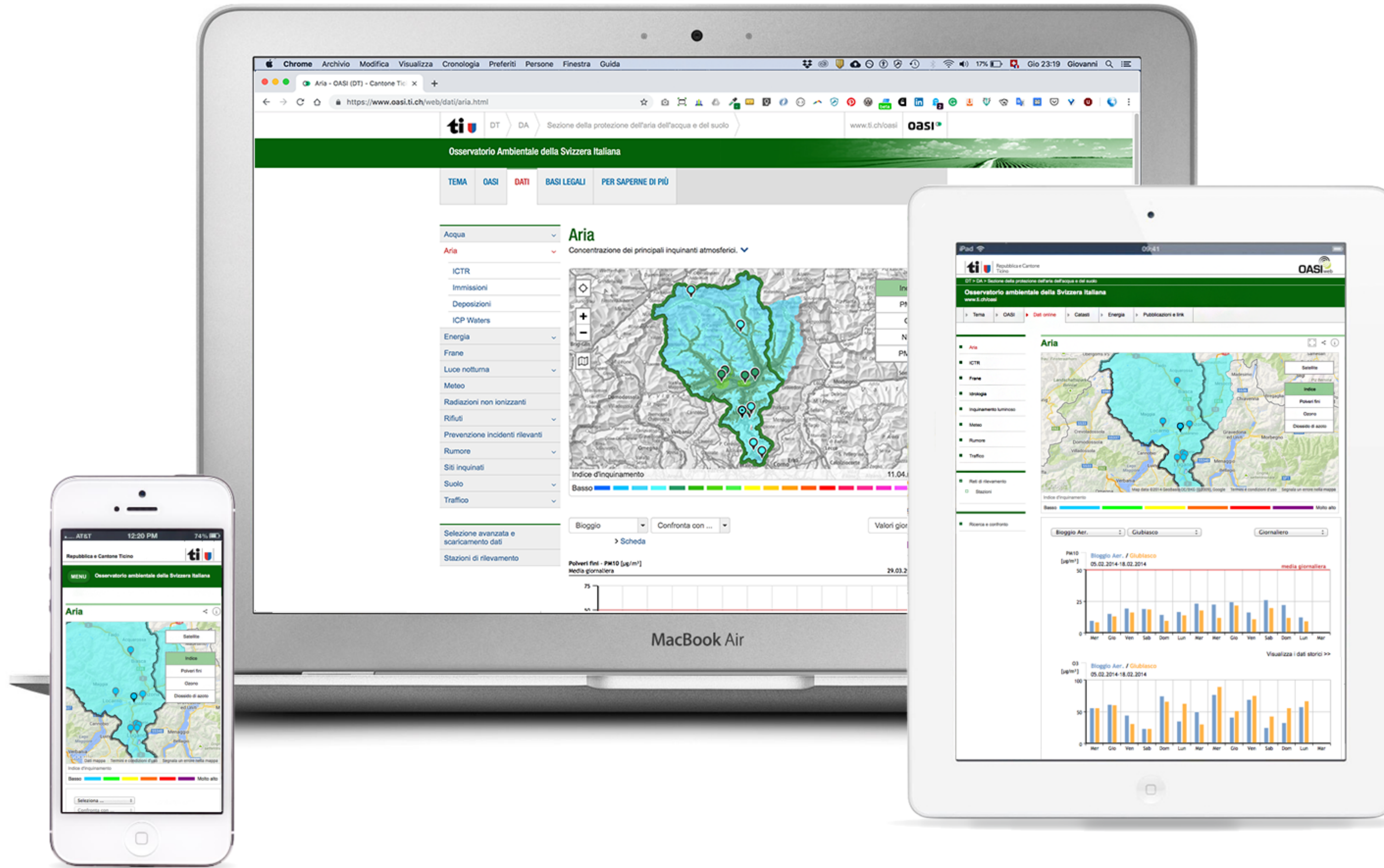
Ricercatore presso il Laboratorio
Cultura Visiva della Scuola
Universitaria Professionale della
Svizzera Italiana (SUPSI)

Docente presso il bachelor in
Comunicazione Visiva SUPSI

Assistente del Master in
Interaction Design SUPSI



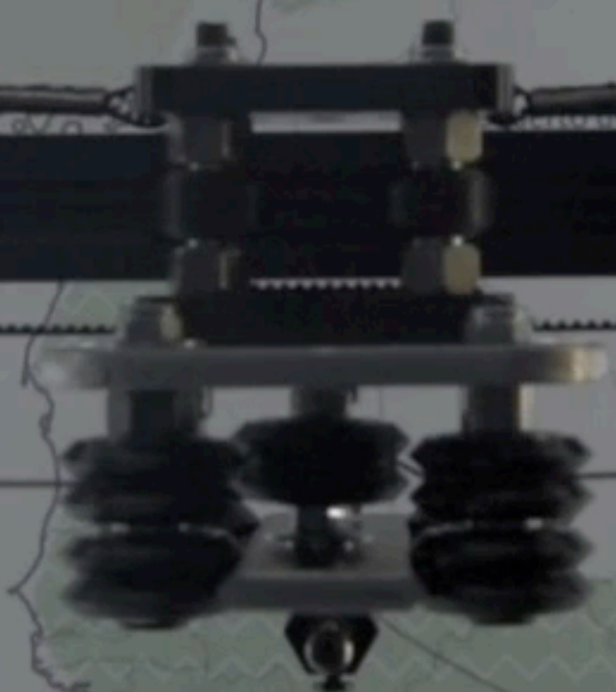
Giovanni Profeta, Massimiliano Cannata
Seminario di Map Design, Bachelor SUPSI in Comunicazione Visiva (2018)



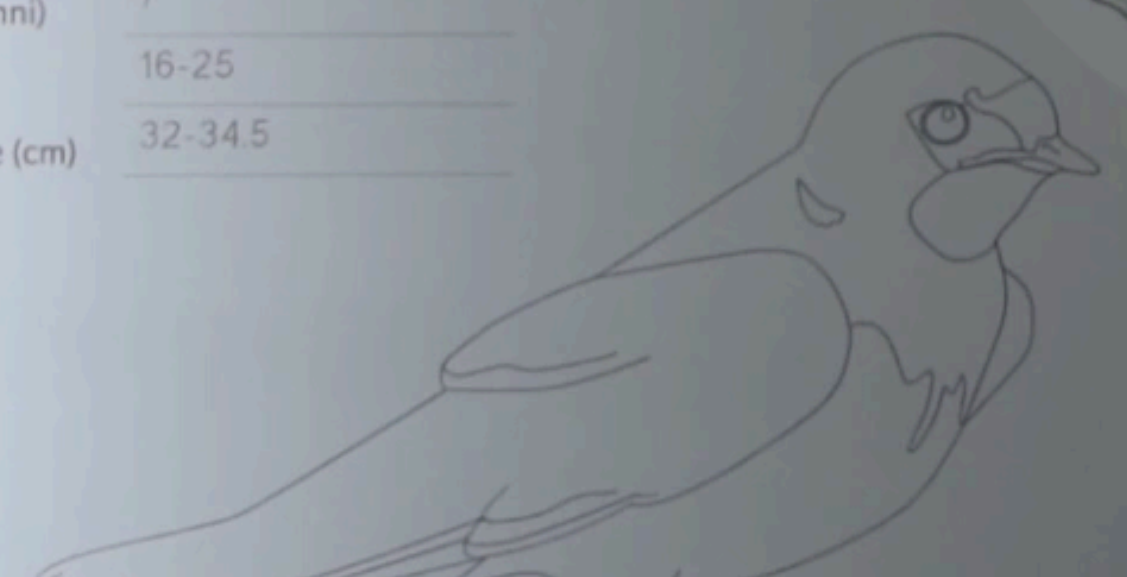
Giovanni Profeta, Serena Cangiano, Massimo Botta
Osservatorio Ambientale della Svizzera Italiana (OASI) (2012-2013). Website: <https://www.oasi.ti.ch>

Dall'Africa e ritorno

Le rondini del Piano di Magadino



Nome scientifico	Hirundo rustica
Vita media (anni)	7
Peso (g)	16-25
Apertura alare (cm)	32-34.5



Wikipedia Primary School

Team di ricerca

Iolanda Pensa
Tobias Schönwetter
Luca Botturi
Florence Devouard
Giancarlo Gianocca
Erica Litrenta
Giovanni Profeta
Marta Pucciarelli
Isla Haddow-Flood
Kelsey Wiens

Enti finanziatori

Fondo Nazionale Svizzero
Fondo Nazionale Sud Africano

Partner

Wikimedia CH
Wikimedia ZA
Africa Centre
Wiki Africa

Wikipedia Primary School

obiettivi dell'analisi grafica

Analizzare lo stato dell'arte degli articoli

Verificare l'impatto del progetto su Wikipedia

Fasi del progetto

Stato dell'arte

Ipotesi

Estate 2015

Stato intermedio

Comparazione

Primavera 2016

Stato finale

Conclusioni

Inverno 2016



200 articoli organizzati in 40 categorie

Acid, Africa Day, Agriculture in South Africa, AIDS orphan, Alcohol abuse, Alcohol dependence, Alcoholism, Aloe, Aloe vera, Animal husbandry, Animal husbandry in South Africa, Apartheid, Atom, Base (chemistry), Bias, Bicycle, Biome, Bird nest, Boiling, Borehole, Bullying, Car, Castle of Good Hope, Cell (biology), Ceramic, Chapter Two of the Constitution of South Africa, Child abuse, Children's Act, 2005, Children's Day, Circuit diagram, Clay, Climate, Coal, Coal in South Africa, Control variable, Convenience food, Cooking, Cradle of Humankind, Day of Reconciliation, Decantation, Dependent and independent variables, Design, Discrimination, Domestic violence, Domestic violence in South Africa, Drakensberg, Duration (music), Economy of South Africa, Ecosystem, Electrical conductor, Electrical network, Electricity generation, Energy in South Africa, Fair trade, Farm, Fiber, Fire safety, Flag of South Africa, Food, Food group, Food pyramid (nutrition), Food security, Food vs. fuel, Frances Baard, Free State (province), Freedom Day (South Africa), Gangster, Gas, Gauteng, Gender role, Glass, Griqua people, Hand washing, Health, Healthcare in South Africa, Herero and Namaqua genocide, Heritage Day (South Africa), History of South Africa, History of writing, HIV, HIV/AIDS, Home safety, Human nutrition, Human Rights Day, Human swimming, Hydraulics, Insulator (electricity), Kaditshwene, Khoikhoi, Khoisan languages, Kingdom of Mapungubwe, Later Stone Age, Latitude, Liquid fuel, List of South Africans, Longitude, Mahatma Gandhi, Makhonjwa Mountains, Malaria, Mandela Day, Map, Mapungubwe Collection, Mapungubwe National Park, Mind map, Mining industry of South Africa, Mpumalanga, Music of Africa, Music of South Africa, National anthem of South Africa, National symbols of South Africa, National Women's Day, Nelson Mandela, Nest, Noise pollution, Oliver Tambo, Outer space, Outline of animal-powered transport, Outline of domestic violence, Paper, People Against Gangsterism and Drugs, Petrol-paraffin engine, Pitch (music), Pneumatics, Politics of South Africa, Pollution, Rail transport, Rail transport in South Africa, Reproduction, Right to food, Road map, Rock (geology), San healing practices, San people, San religion, San rock art, Sand, Sarah Baartman, Scale (map), Scientific method, Settling, Sexism, Sieve, Soil, Sound, South Africa, South African cuisine, South African locomotive history, Standard Model, Stereotype, Steve Biko, Tap (valve), Textile, Trade, Transport, Transport in South Africa, Two-foot-gauge railways in South Africa, Vibration, Walter Sisulu, Water filter, Water pollution, Water privatisation in South Africa, Water supply and sanitation in South Africa, Water well, Wax, Western Cape, Wetland, Wildlife conservation, Wildlife management, Winnie Madikizela-Mandela, Wood, Wool, Youth culture, !Kung people

Alcoholism, Animal husbandry and wild animal maintenance, Animal shelters, Animal shelters, Biological sciences, Bullying, Circuits, Crop and stock farming, Discrimination, stereotypes and bias, Domestic violence, Economy - political system, Ecosystem, natural resources, Fair trading, Gangster, Gender stereotyping, Geography, Health and hygiene topics and impacts, Heritage from each province, History of written communication, HIV/AIDS, How people get access to water, Human rights, Khoikhoi herder society in the Later Stone Age, Life stories of leaders, Mathematics, Mineral and coal resources in South Africa, Movement and energy in a system, National events, National history, Nutrients in food / Balanced diets, Physics and engineering, Processes to purify water, Safety measures, San hunter-gatherer society in the Later Stone Age, Solid materials, Stored energy in fuels, Transport on land, Vibration and sound, Youth culture

Wikipedia, the free encyclopedia x Giovanni

Sicuro https://en.wikipedia.org/wiki/Main_Page

Main Page Talk Read View source View history Search Wikipedia

WIKIPEDIA The Free Encyclopedia

Welcome to **Wikipedia**, the free encyclopedia that anyone can edit. 5,411,027 articles in English

- Arts
- History
- Society
- Biography
- Mathematics
- Technology
- Geography
- Science
- All portals

From today's featured article

Béla Lugosi as Dracula

A **vampire** is a being from **folklore** who subsists on the blood or life essence of the living. In European folklore, vampires were shroud-wearing **undead** beings who often visited loved ones and caused mischief in the neighbourhoods they inhabited when they were alive. Before the early 19th century, they were described as bloated and of ruddy or dark countenance, markedly different from today's gaunt, pale vampire. The term *vampire* was popularised in the West in the early 18th century, after vampire legends from oral traditions of ethnic groups of the Balkans and Eastern Europe were recorded and published. The charismatic and sophisticated vampire of modern fiction was born in 1819 with the publication of *The Vampyre* by John Polidori; the story was highly successful and arguably the most influential vampire work of the early 19th century. The modern basis of the vampire legend comes from Bram Stoker's 1897 novel *Dracula*, considered the quintessential *vampire novel*. The success of this book spawned a distinctive vampire *genre*, still popular in the 21st century. **(Full article...)**

Recently featured: Brabham · "Here We Go Again" (Ray Charles song) · Wood Siding railway station Archive · By email · More featured articles...

Did you know...

- ... that the **white-barred piculet** hybridises with the **ochre-collared (pictured)**, **varzea**, **ocellated**, and **white-wedged** piculets where their ranges overlap?
- ... that **Vidhu Vincent** is the first woman to win the **Kerala State Film**

In the news

- In cricket, **Mumbai Indians win the Indian Premier League, defeating Rising Pune Supergiant in the final.**
- Hassan Rouhani (pictured) is re-elected President of Iran.**
- Salvador Sobral, representing Portugal, wins the Eurovision Song Contest with the song "Amar pelos dois".**
- A large-scale cyber attack involving ransomware causes severe disruption around the world.**
- At least 28 people are killed in Mastung, Pakistan, in **a suicide attack targeting Abdul Ghafoor Haideri, the Senate Deputy Chairman.**
- The **Free National Movement, led by Hubert Minnis, wins the Bahamian general election.**

Ongoing: Battle of Mosul
Recent deaths: Roger Ailes · Chris Cornell · Ian Brady

On this day...

May 22: **Victoria Day** in Canada (2017)

- 1826 – **HMS Beagle** departed on her first voyage from Plymouth for a hydrographic survey of the Patagonia and Tierra del Fuego regions of South America.
- 1897 – The first **Blackwall Tunnel (construction pictured)** under the **River Thames** was opened to

Special page Search

API sandbox

Use this page to experiment with the **MediaWiki web service API**. Refer to the API documentation for further details of API usage. Example: [get the content of a Main Page](#). Select an action to see more examples.

Note that, although this is a sandbox, actions you carry out on this page may modify the wiki.

Parameters

Format json

Action query

Documentation

backlinks

Find all pages that link to the given page. (read more)

- Show links to Main page.
- Get information about pages linking to Main page.

Make request Examples Clear

Parameter for backlinks	Input	Description
bltitle	<input type="text"/>	Title to search. Cannot be used together with <i>blpageid</i> .
blpageid	<input type="text"/>	Page ID to search. Cannot be used together with <i>bltitle</i> .
blcontinue	<input type="text"/>	When more results are available, use this to continue.
blnamespace	(Main) Talk User User talk Project Project talk File File talk MediaWiki MediaWiki talk	The namespace to enumerate.
blidir	(select value)	The direction in which to list.
blfilterredir	(select value)	How to filter for redirects. If set to <i>nonredirects</i> when <i>blredirect</i> is enabled, this is only applied to the second level.

Main page
Get MediaWiki
Get extensions
Tech blog
Contribute

Support
User help
FAQ
Technical manual
Support desk
Communication

Development
Bug tracker
Code repository
Code docs
Statistics
Wikimedia engineering

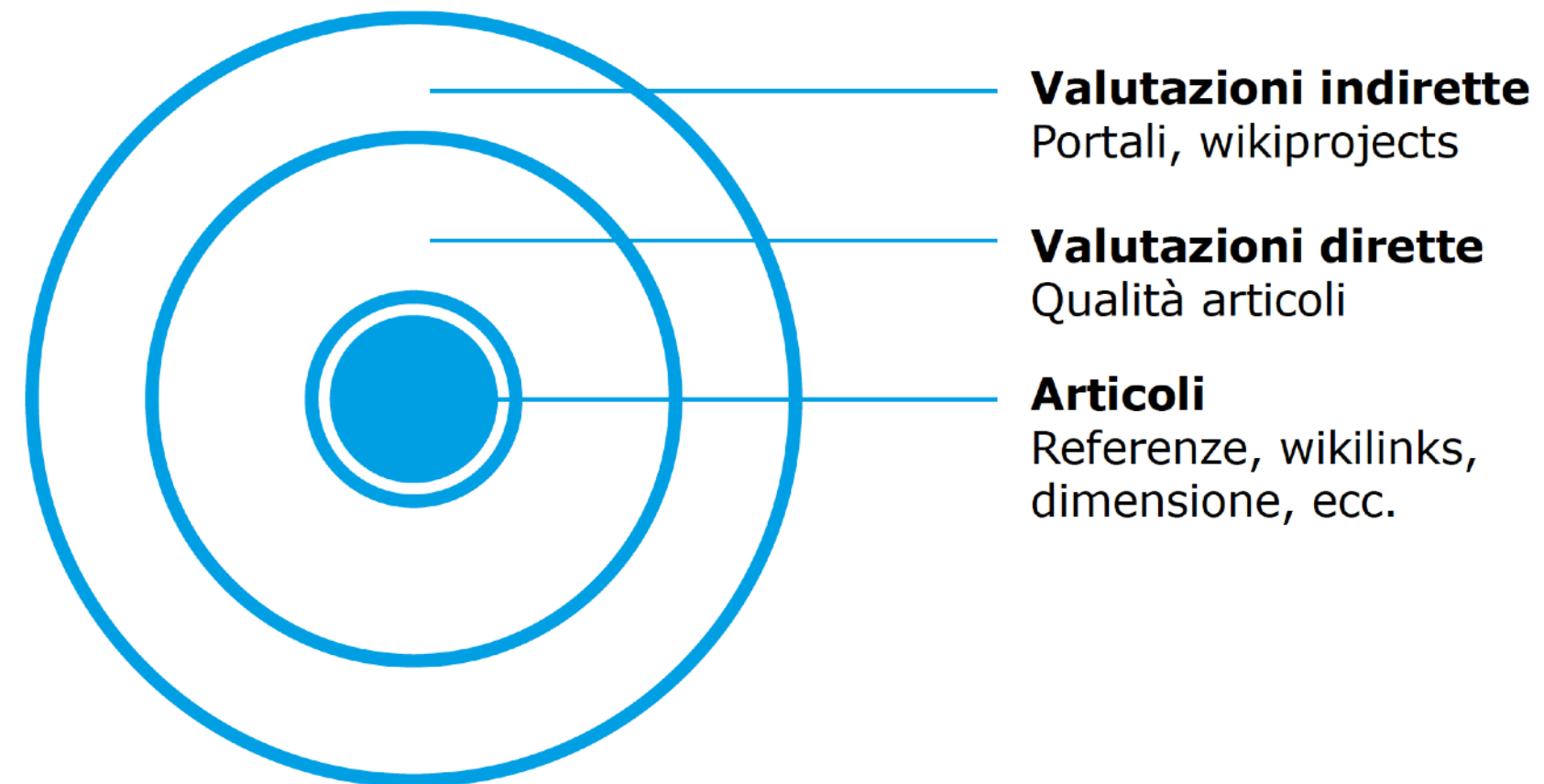
MediaWiki.org
Browse categories
Community portal
Recent changes
Random page
Current issues
Sandbox

Tools
Special pages
Printable version

<https://www.wikipedia.org/>

<https://www.mediawiki.org/wiki/Special:ApiSandbox?action=languagesearch&format=json&search=Maratih&typos=1>

Valutazioni dirette e indirette



Caratteristiche indagate

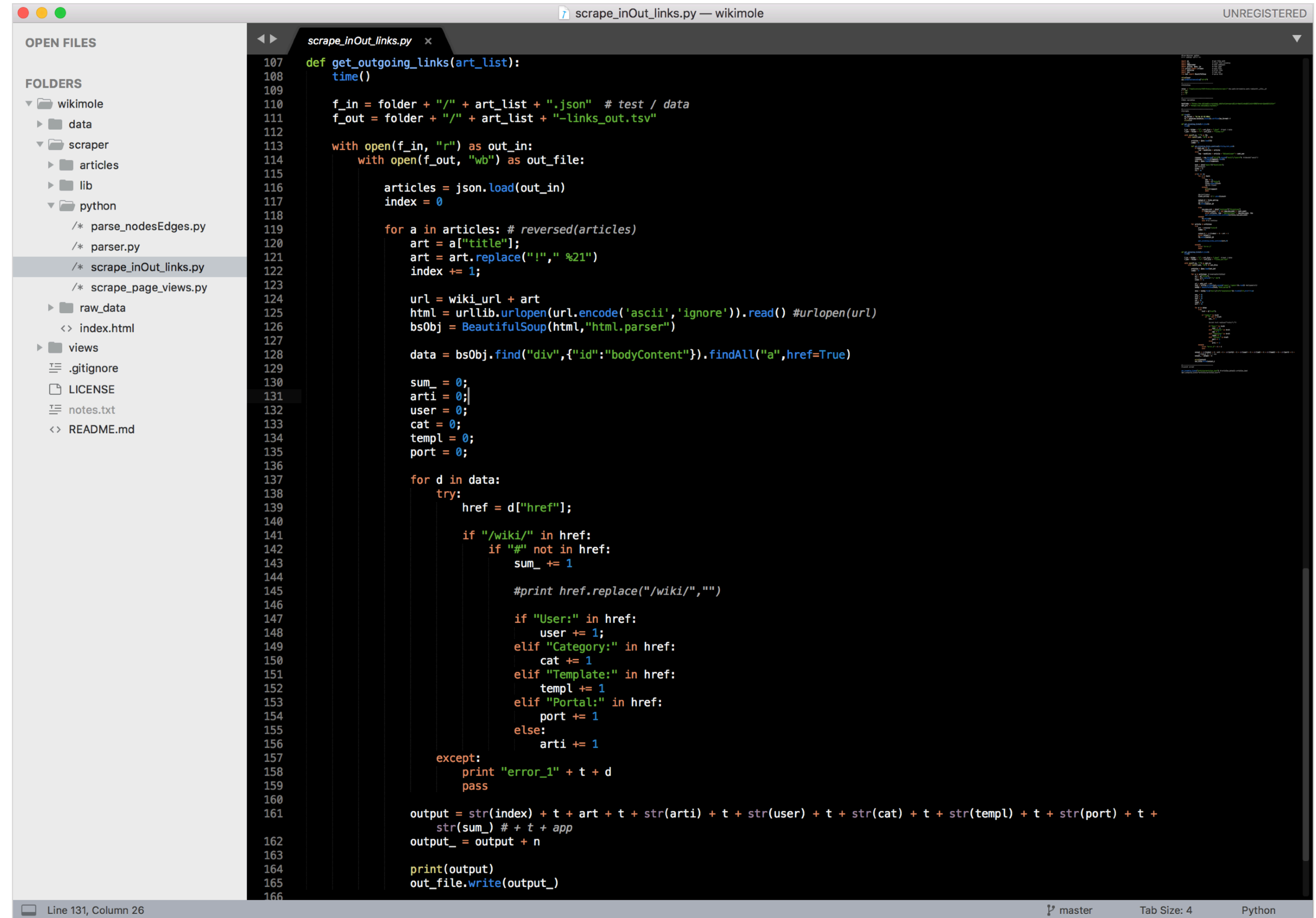
Trovabilità
intertestualizzazione
degli articoli

Attrattività
modalità di
redazione degli
articoli

Specificità
integrazione
di risorse negli
articoli

link nell'articolo

Ottenimento e elaborazione dati con script in Python



```
scrape_inOut_links.py — wikimole UNREGISTERED

OPEN FILES
FOLDERS
  wikimole
    data
    scraper
      articles
      lib
      python
        /* parse_nodesEdges.py
        /* parser.py
        /* scrape_inOut_links.py
        /* scrape_page_views.py
    raw_data
      <> index.html
    views
  .gitignore
  LICENSE
  notes.txt
  <> README.md

scrape_inOut_links.py x
107 def get_outgoing_links(art_list):
108     time()
109
110     f_in = folder + "/" + art_list + ".json" # test / data
111     f_out = folder + "/" + art_list + "-links_out.tsv"
112
113     with open(f_in, "r") as out_in:
114         with open(f_out, "wb") as out_file:
115
116             articles = json.load(out_in)
117             index = 0
118
119             for a in articles: # reversed(articles)
120                 art = a["title"];
121                 art = art.replace("!", "%21")
122                 index += 1;
123
124                 url = wiki_url + art
125                 html = urllib.urlopen(url.encode('ascii', 'ignore')).read() #urlopen(url)
126                 bsObj = BeautifulSoup(html, "html.parser")
127
128                 data = bsObj.find("div", {"id": "bodyContent"}).findAll("a", href=True)
129
130                 sum_ = 0;
131                 arti = 0;
132                 user = 0;
133                 cat = 0;
134                 templ = 0;
135                 port = 0;
136
137                 for d in data:
138                     try:
139                         href = d["href"];
140
141                         if "/wiki/" in href:
142                             if "#" not in href:
143                                 sum_ += 1
144
145                                 #print href.replace("/wiki/", "")
146
147                                 if "User:" in href:
148                                     user += 1;
149                                 elif "Category:" in href:
150                                     cat += 1
151                                 elif "Template:" in href:
152                                     templ += 1
153                                 elif "Portal:" in href:
154                                     port += 1
155                                 else:
156                                     arti += 1
157                     except:
158                         print "error_1" + t + d
159                         pass
160
161                 output = str(index) + t + art + t + str(arti) + t + str(user) + t + str(cat) + t + str(templ) + t + str(port) + t +
162                     str(sum_) # + t + app
163                 output_ = output + n
164
165                 print(output)
166                 out_file.write(output_)
```

Line 131, Column 26

master Tab Size: 4 Python

Esplorazione grafica con RawGraphs

The screenshot shows the RawGraphs web application interface. The browser address bar displays "Non sicuro | app.rawgraphs.io". The page header includes the "RAWGraphs" logo and navigation links: "About", "Gallery", "Learning", "FAQs", "User survey", and "Developer Guide".

The main section is titled "Load your data" and offers four options: "Paste", "Upload a file", "From URL", and "Try our samples". A text area contains a list of 15 movie records with columns for "Movie", "Genre", "Production Budget (millions)", "Box Office (millions)", "ROI", and "Rating IMDB".

Movie	Genre	Production Budget (millions)	Box Office (millions)	ROI	Rating IMDB
Avatar	Action	237	2784	11.7	8.0
The Blind Side	Drama	29	309	10.7	7.6
"The Chronicles of Narnia: The Lion, the Witch and the Wardrobe"	Adventure	180	745	4.1	6.9
The Dark Knight	Action	185	1005	5.4	9.0
ET: The Extra-Terrestrial	Drama	11	793	75.5	7.9
Finding Nemo	Adventure	94	940	10.0	8.1
Ghostbusters	Comedy	144	229	1.6	7.8
The Hunger Games	Thriller/Suspense	78	649	8.3	7.2
Iron Man 3	Action	178	1215	6.8	7.6
Jurassic Park	Action	53	1030	19.4	8.0
King Kong	Adventure	207	551	2.7	7.3
The Lion King	Adventure	45	968	21.5	8.4
"Monsters, Inc."	Adventure	115	577	5.0	8.0
The Twilight Saga: New Moon	Drama	50	710	14.2	4.5

A green notification bar at the bottom of the data area states: "26 records in your data have been successfully parsed!". Below this, a link says: "Your data seems ready to go. But if you want to stack it anyway, click here".

The "Choose a Chart" section displays a grid of chart options:

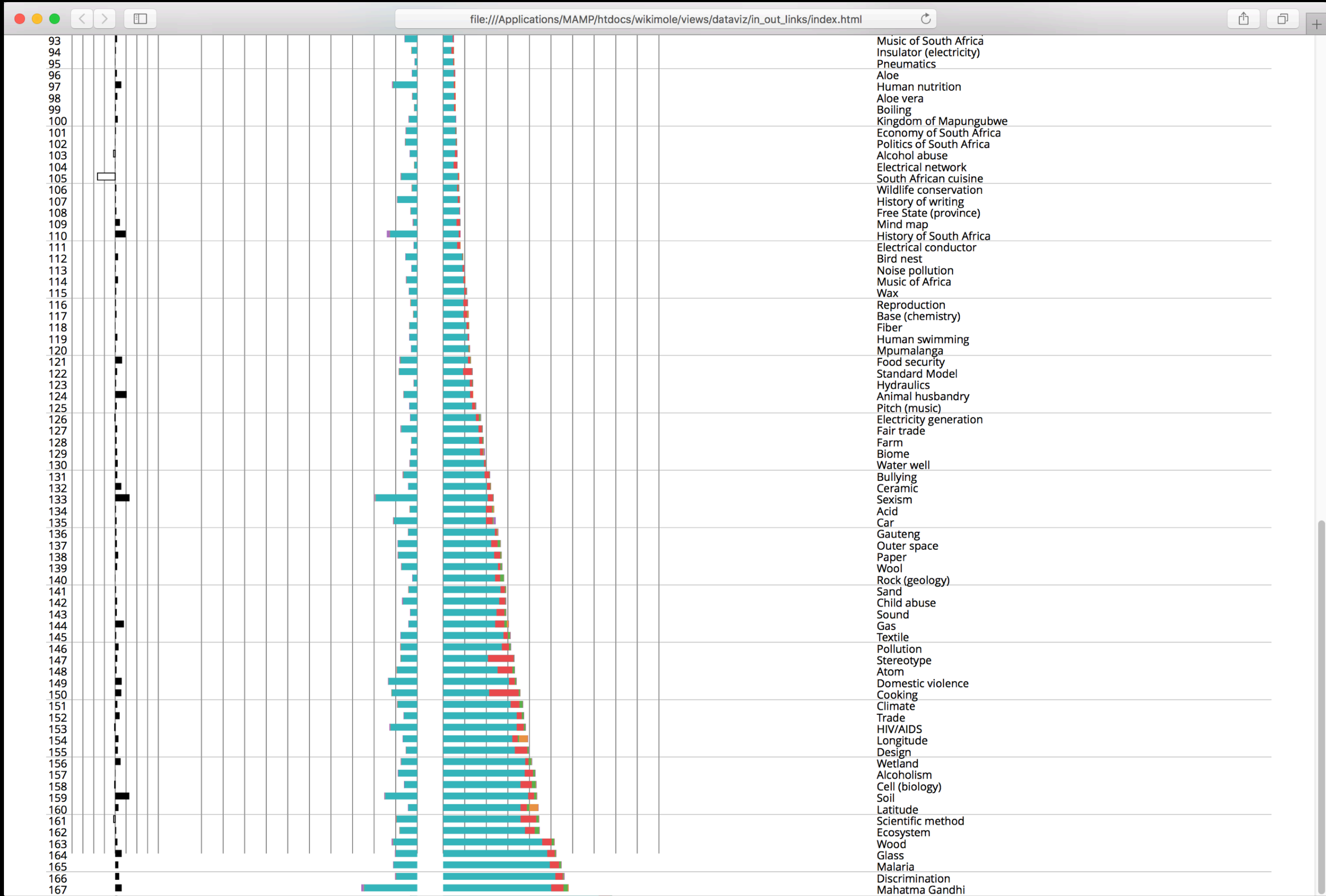
- Scatter Plot Dispersion
- Contour Plot Dispersion
- Convex Hull Dispersion
- Hexagonal Binning Dispersion
- Scatter Plot Dispersion (highlighted with a green checkmark)

Visualizzazione dati con D3.js

The image shows a code editor window with a file explorer on the left and a code editor on the right. The file explorer shows a project structure with folders like 'wikimole', 'data', 'scraper', 'views', 'assets', 'dataviz', 'approaches', 'edit', 'edits_and_editors', 'edits_x_editor', 'features', 'in_out_links', and files like 'index.html', 'no_articles.html', 'datasets.html', 'protocols.html', '.gitignore', 'LICENSE', 'notes.txt', and 'README.md'. The code editor shows the file 'in_out_links_all.js' with the following code:

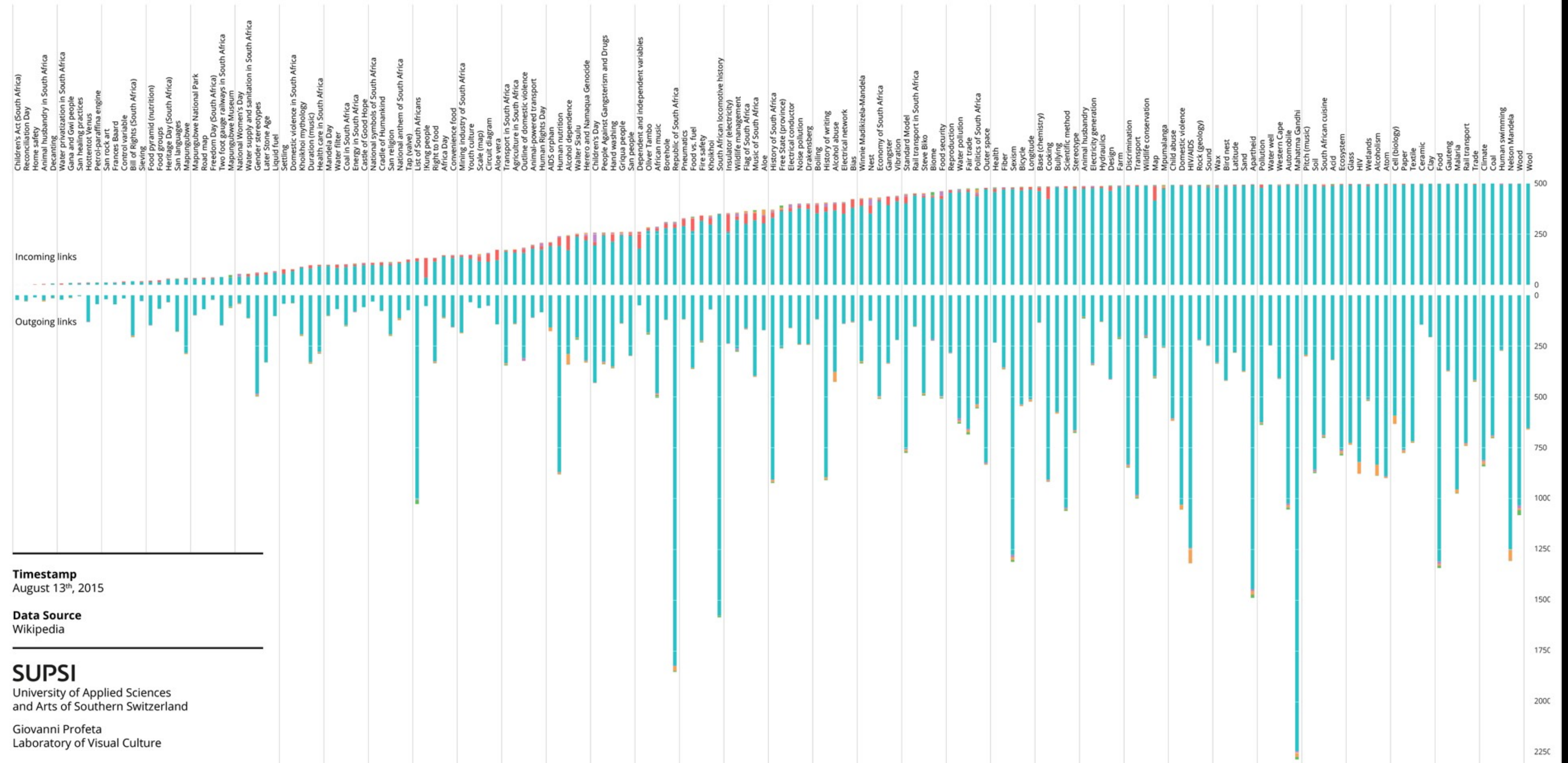
```
49
50 /* -----
51 set plot
52 ----- */
53
54 var svg = d3.select("#svg_container")
55   .append("svg")
56   .attr("viewBox", '0 0 ' + width + ' ' + (height) )
57
58 var plot = svg.append("g")
59   .attr("id", "d3_plot")
60   .attr("transform", "translate(" + margin.left + "," + margin.top + ")");
61
62 /* -----
63 get data
64 ----- */
65
66 var dataset = "20170803/in_out_links_2017.csv";
67
68 d3.csv("../assets/data/" + dataset, loaded);
69
70 function loaded (data){
71
72   data.forEach(function(d) {
73     d.id = +d.id;
74     d.article = d.article;
75
76     d.page_in = +d.page_in;
77     d.user_in = +d.user_in;
78     d.portal_in = +d.portal_in;
79     d.template_in = +d.template_in;
80     d.category_in = +d.category_in;
81     d.total_in = +d.total_in;
82
83     d.page_out = +d.page_out;
84     d.user_out = +d.user_out;
85     d.portal_out = +d.portal_out;
86     d.template_out = +d.template_out;
87     d.category_out = +d.category_out;
88     d.total_out = +d.total_out;
89
90     d.RW_by_community = +d.RW_by_community;
91     d.RW_by_expert_pdf = +d.RW_by_expert_pdf;
92     d.RW_by_expert_pdf_wiki = +d.RW_by_expert_pdf_wiki;
93     d.New_article_suggested_by_expert = +d.New_article_suggested_by_expert;
94     d.AFC = +d.AFC;
95     d.Featured_on_WP_SA_portal = +d.Featured_on_WP_SA_portal;
96     d.Rewrite_based_on_expert_review = +d.Rewrite_based_on_expert_review;
97     d.WP_Assessment = +d.WP_Assessment;
98     d.Bold_reassessment = +d.Bold_reassessment;
99     d.Africa_Destubathon = +d.Africa_Destubathon;
100    d.Edit_a_thon = +d.Edit_a_thon;
101
102    d.page_in_2015 = +d.page_in_2015;
103    d.page_out_2015 = +d.page_out_2015;
104
105    d.in_out = +d.in_out;
106  });
107
108  data.sort(function(a,b) {
```

The status bar at the bottom shows 'Line 1, Column 1', 'master', 'Tab Size: 4', and 'JavaScript'.



Incoming and outgoing links

The visualization shows the balance between incoming links (in Wikipedia: "What links here") and wikilinks for every article (171).



How to read it

On the top the amount of incoming links. On the bottom the amount of outgoing links. Both consist of links to other articles, users' pages, categories, templates and portals. From left to right, articles are in ascending order of incoming links. Please note: 6 articles (4% of the total) has 500 or more incoming links (Wikipedia API does not allow to get more than 500 links for article).

Legenda

- articles
- users
- categories
- templates
- portals

Timestamp
August 13th, 2015

Data Source
Wikipedia

SUPSI
University of Applied Sciences
and Arts of Southern Switzerland

Giovanni Profeta
Laboratory of Visual Culture

tra la fase 1 e 2

Definizione metodi per il miglioramento degli articoli

Review by community
Review by expert pdf
Review by expert pdf and wiki
New article suggested by expert
New article created AFC
Featured on WikiProject South Africa
Rewrite based on expert review
Article assessment
Bold reassessment
Africa Destubathon
Edit-a-thon

Applicazione dei metodi ad un set ristretto di articoli

Water privatisation in South Africa
Home safety
Frances Baard
Petrol-paraffin engine
Kaditshwene
San healing practices
Makhonjwa Mountains
Children's Act; 2005
Day of Reconciliation
...

Articoli che hanno cambiato titolo, sono stati aggiunti e cancellati

Music of Africa
Car
Chapter Two of the Constitution of South Africa
Decantation
Saartjie Baartman
Kingdom of Mapungubwe
Gas
Gender role
Kaditshwene
...

Articles

[Wikimedia page](#) - [List on Google Drive](#)

Last update: March 8th 2016 (previous September 24th 2015)

1 !Kung people (San hunter-gatherer society in the Later Stone Age)

In. links: 151

Out. links: 226

Issues: 1

References: 10

Notes: 0

Images: 0

+12

+87

=

+2

=

=

2 2 ft gauge railways in South Africa (Transport on land)

In. links: 7

Out. links: 372

Issues: 0

References: 40

Notes: 0

Images: 4

-21

+222

=

=

=

=

3 Acid (Biological sciences)

In. links: 483

Out. links: 390

Issues: 2

References: 4

Notes: 0

Images: 13

-14

+68

+1

+1

=

=

4 Africa Day (National events)

In. links: 165

Out. links: 131

Issues: 0

References: 3

Notes: 0

Images: 2

+18

+16

=

=

=

=

5 Agriculture in South Africa (Crop and stock farming)

In. links: 199

Out. links: 256

Issues: 0

community review

References: 17

expert review

Notes: 0

Images: 11

+23

+112

=

+1

=

=

6 AIDS orphan (HIV/AIDS)

In. links: 221

Out. links: 171

Issues: 0

References: 5

Notes: 0

Images: 2

+10

-7

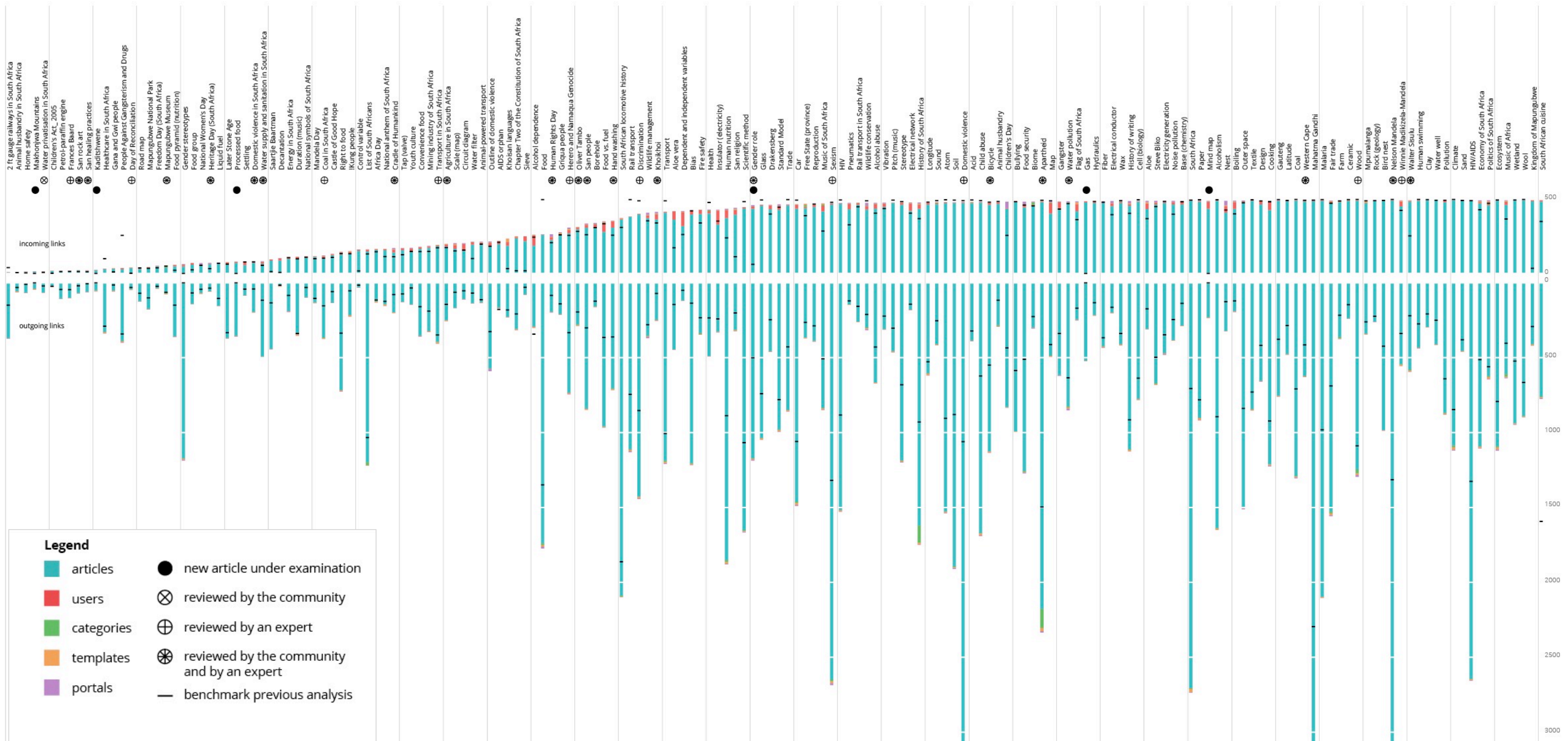
-

-

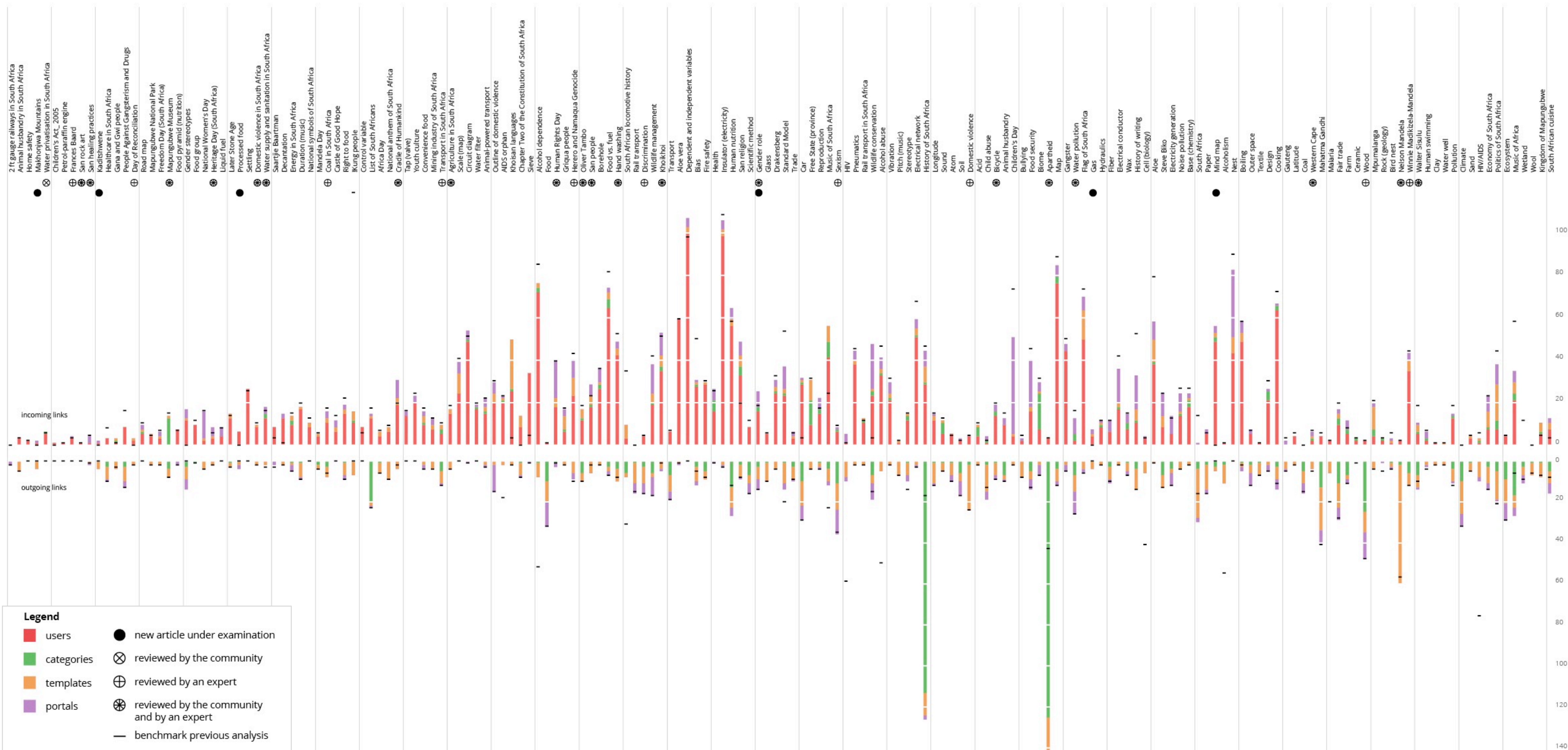
-

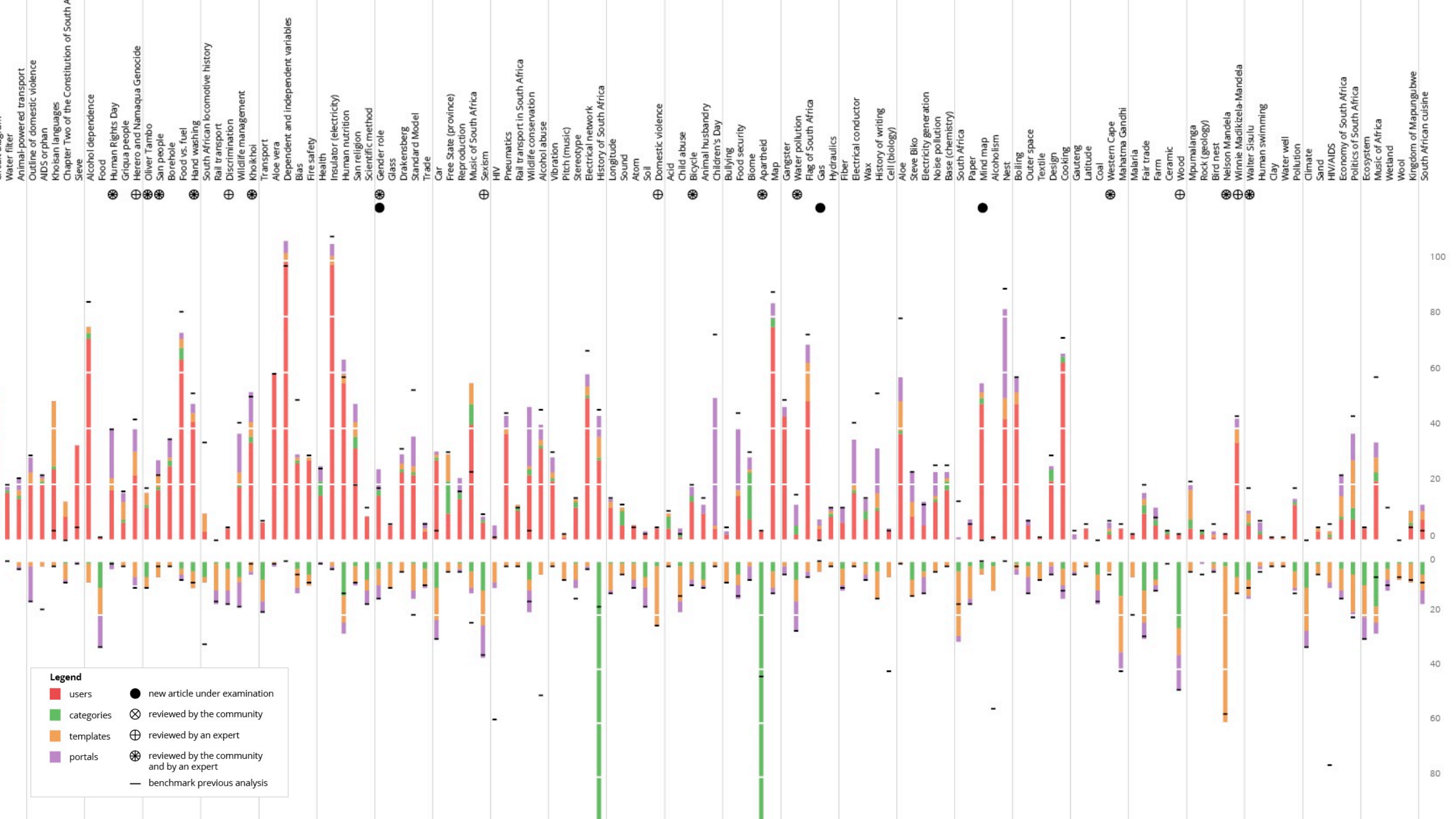
-

Link in entrata e in uscita (agosto 2015/marzo 2016)



Link in entrata e in uscita (esclusi gli articoli)





Incoming and outgoing links (except articles)

The visualization shows the balance between incoming links (in Wikipedia: "What links here") and links to other Wikipedia pages (Wikilinks). Link from and to other articles are not shown.

Legend

- users
- categories
- templates
- portals
- new article under examination
- ⊗ reviewed by the community
- ⊕ reviewed by an expert
- ⊗ reviewed by the community and by an expert
- benchmark previous analysis

How to read the visualization

On the top the bars show the amount of incoming links. On the bottom the bars show the amount of outgoing links. Both consist of links to other, users pages, categories, templates and portals.

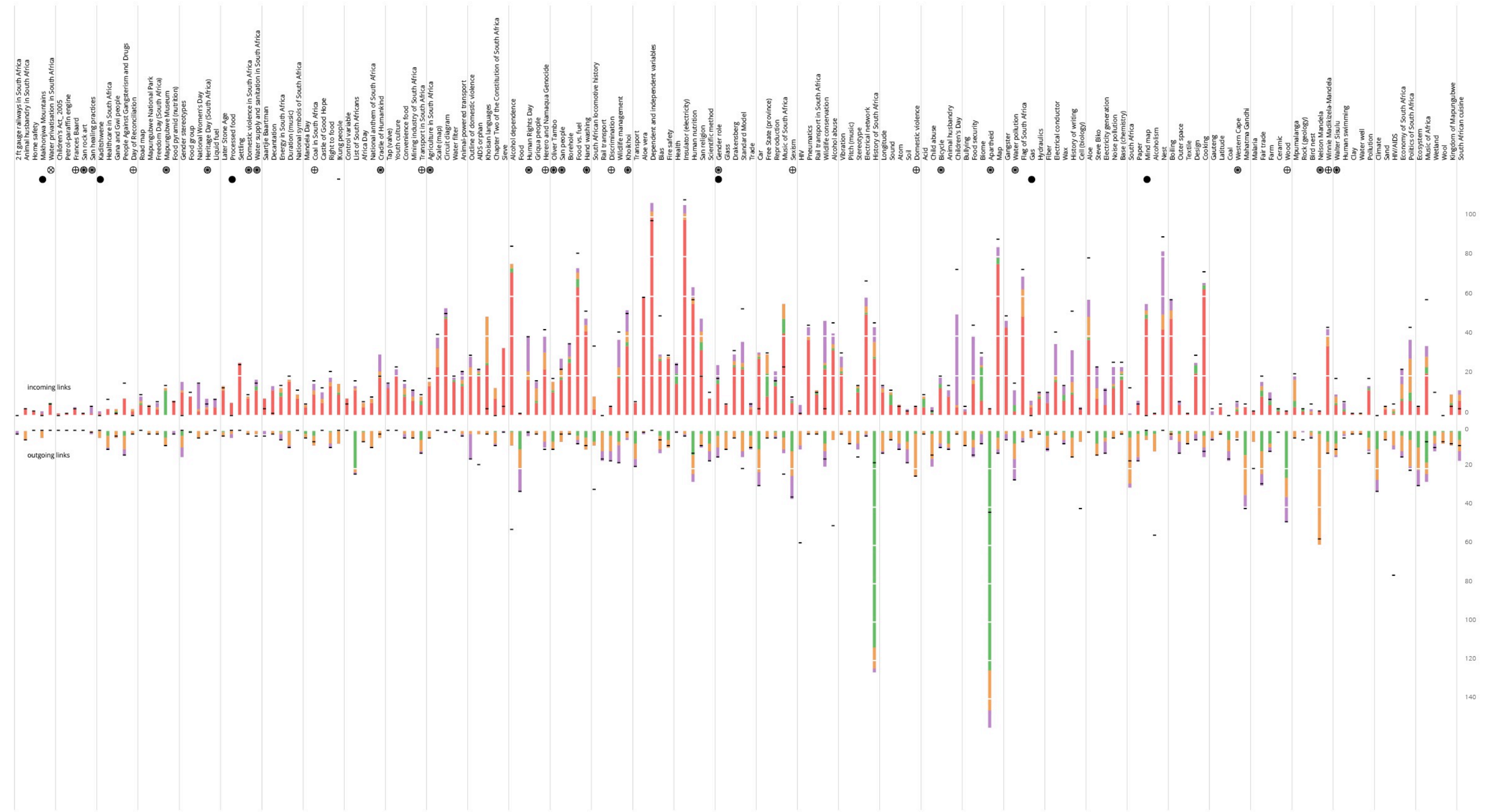
Please note: from left to right, articles are in ascending order of total amount of incoming links (link to and from other articles included).

SUPSI

Department of Environment Constructions and Design
Laboratory of visual culture

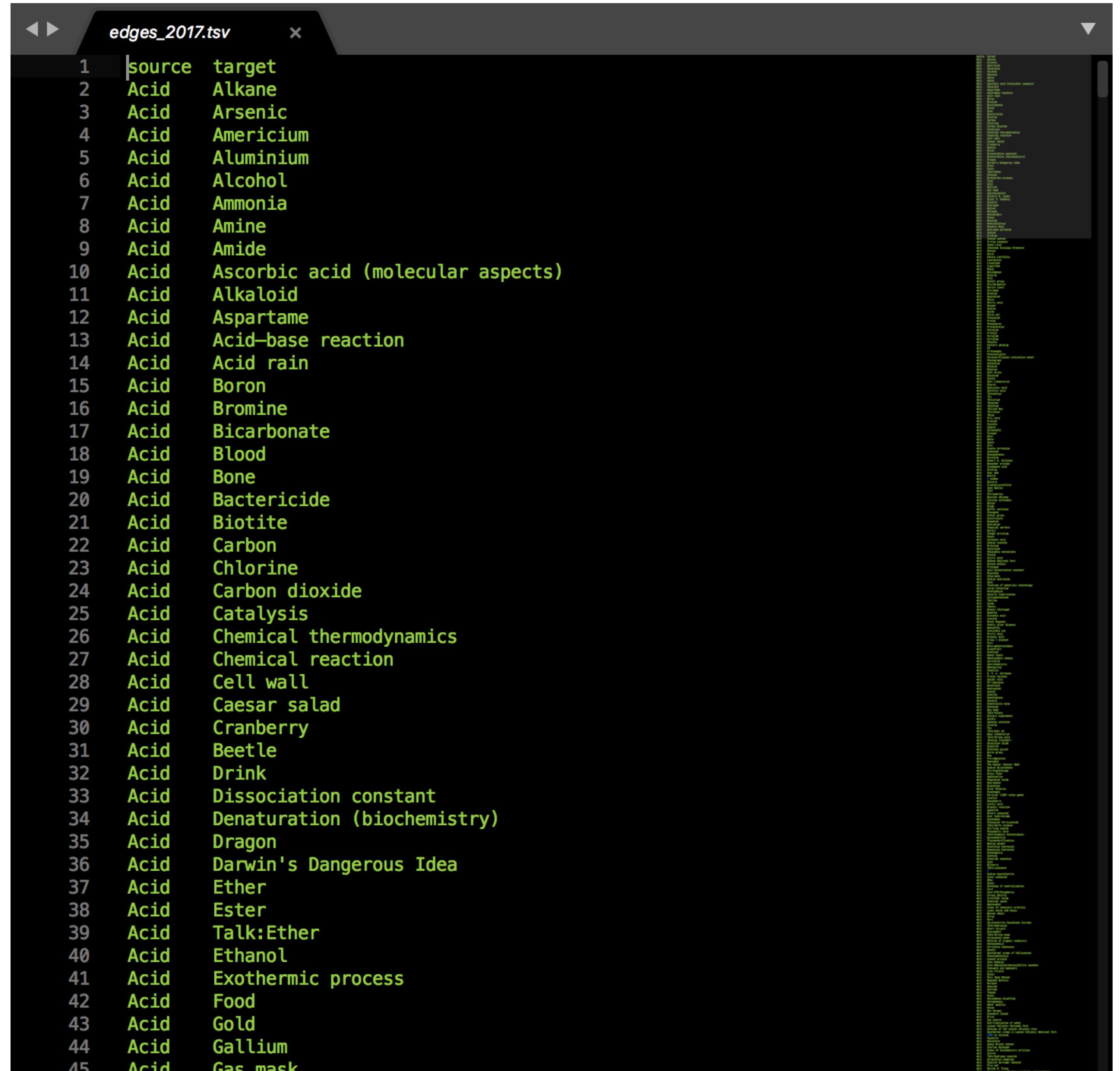
Project by

Giovanni Profeta
giovanni.profeta@supsi.ch



link tra gli articoli

Rete di pagine basata sui link in entrata



	source	target
1	Acid	Alkane
2	Acid	Arsenic
3	Acid	Americium
4	Acid	Aluminium
5	Acid	Alcohol
6	Acid	Ammonia
7	Acid	Amine
8	Acid	Amide
9	Acid	Ascorbic acid (molecular aspects)
10	Acid	Alkaloid
11	Acid	Aspartame
12	Acid	Acid–base reaction
13	Acid	Acid rain
14	Acid	Boron
15	Acid	Bromine
16	Acid	Bicarbonate
17	Acid	Blood
18	Acid	Bone
19	Acid	Bactericide
20	Acid	Biotite
21	Acid	Carbon
22	Acid	Chlorine
23	Acid	Carbon dioxide
24	Acid	Catalysis
25	Acid	Chemical thermodynamics
26	Acid	Chemical reaction
27	Acid	Cell wall
28	Acid	Caesar salad
29	Acid	Cranberry
30	Acid	Beetle
31	Acid	Drink
32	Acid	Dissociation constant
33	Acid	Denaturation (biochemistry)
34	Acid	Dragon
35	Acid	Darwin's Dangerous Idea
36	Acid	Ether
37	Acid	Ester
38	Acid	Talk:Ether
39	Acid	Ethanol
40	Acid	Exothermic process
41	Acid	Food
42	Acid	Gold
43	Acid	Gallium
44	Acid	Gas mask
45	Acid	

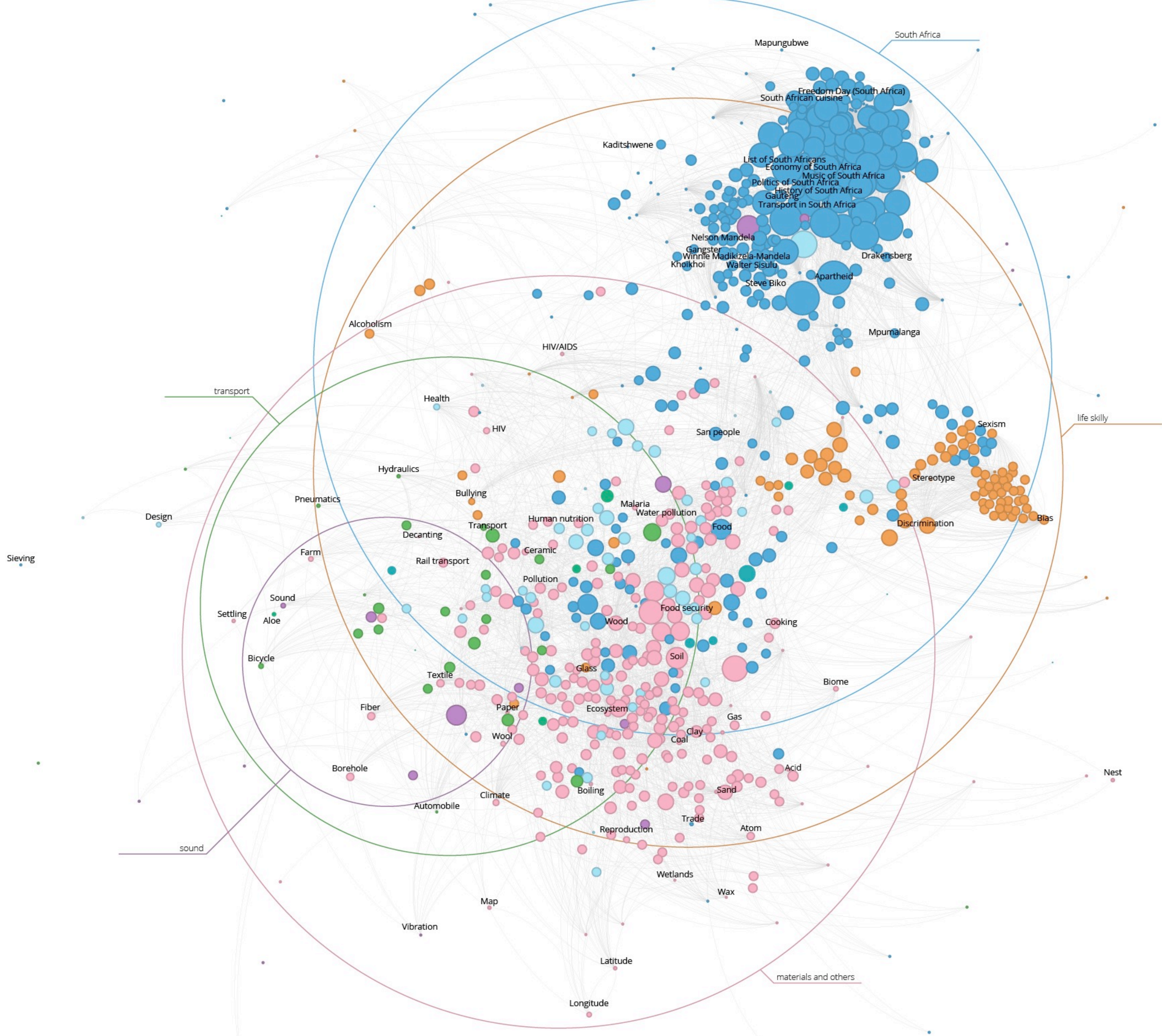
Gephi

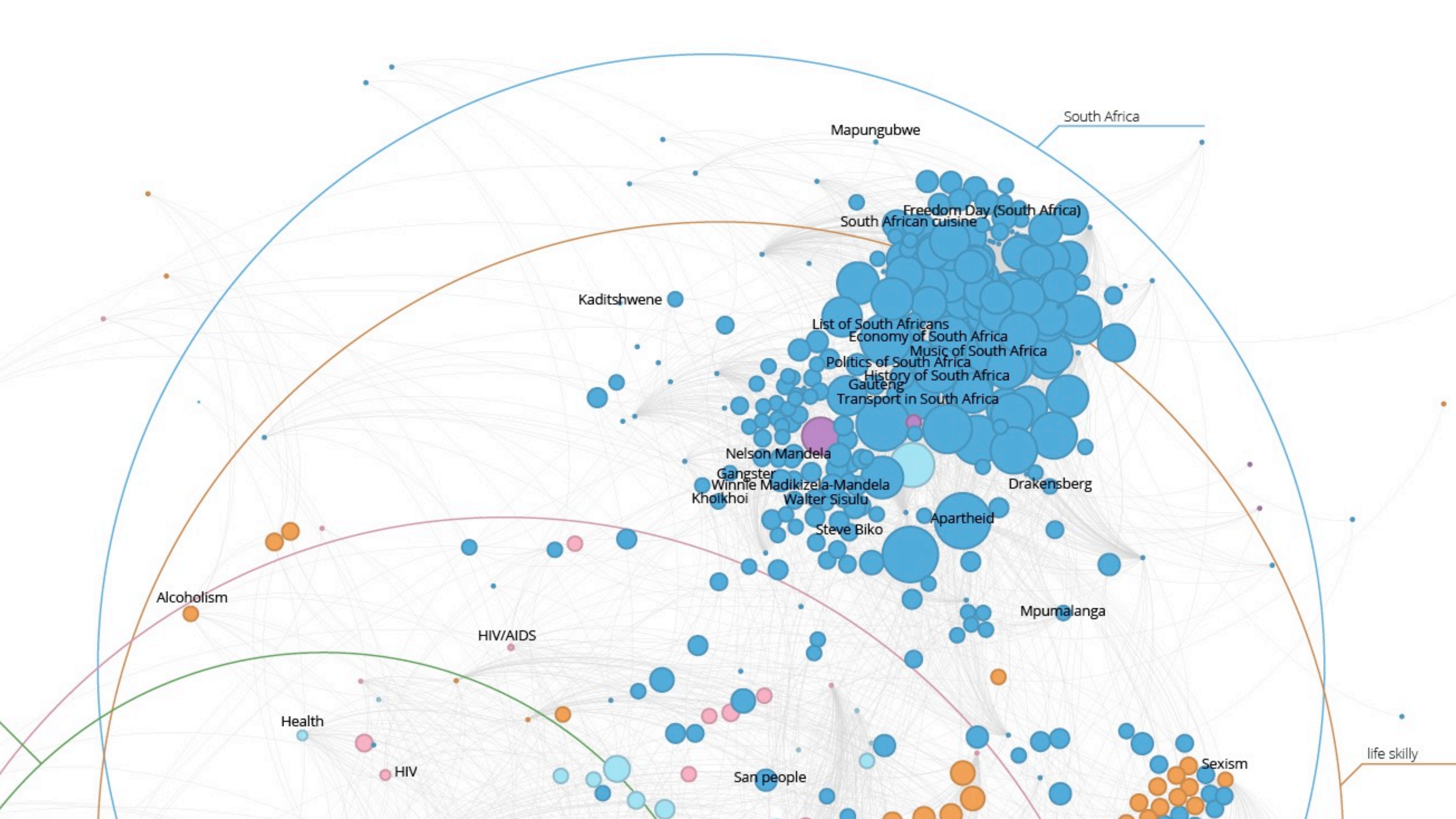
The screenshot displays the Gephi 0.9.2 interface with a network graph visualization. The central canvas shows a dense network of nodes and edges, with node sizes varying significantly. The interface is divided into several panels:

- Appearance Panel (Left):** Shows settings for Nodes and Edges. Under the Nodes tab, the 'Ranking' section is active, with 'Weighted Degree' selected. The 'Min size' is set to 4 and 'Max size' is 400. An 'Apply' button is visible.
- Layout Panel (Left):** Shows the 'ForceAtlas 2' layout algorithm selected. It includes a 'Run' button and a list of parameters:
 - Threads: 3
 - Performance: Tolerance (speed) 1.0, Approximate Repulsion , Approximation 1.2
 - Tuning: Scaling 2.0, Stronger Gravity , Gravity 1.0
 - Behavior Alternatives: Dissuade Hubs , LinLog mode , Prevent Overlap , Edge Weight Influence 1.0
- Context Panel (Right):** Displays graph statistics: Nodes: 1179 (17.27% visible), Edges: 19640 (21.21% visible), Undirected Graph. It also shows a 'Filters' tree with categories like Inter Edges, Intra Edges, Non-null, Partition, Partition Count, Range, Dynamic, Edges, Operator, Topology, and Saved queries. A 'Queries' section shows a 'Range (Weighted Degree)' query with parameters: column: Weighted Degree (class j), range: 50.0 - 1000.0. Below this is a 'Range (Weighted Degree) Settings' slider set from 50.0 to 1000.0.
- Bottom Panel:** Contains a toolbar with icons for selection, zooming, and styling. The font is set to 'Arial-BoldMT, 32'.

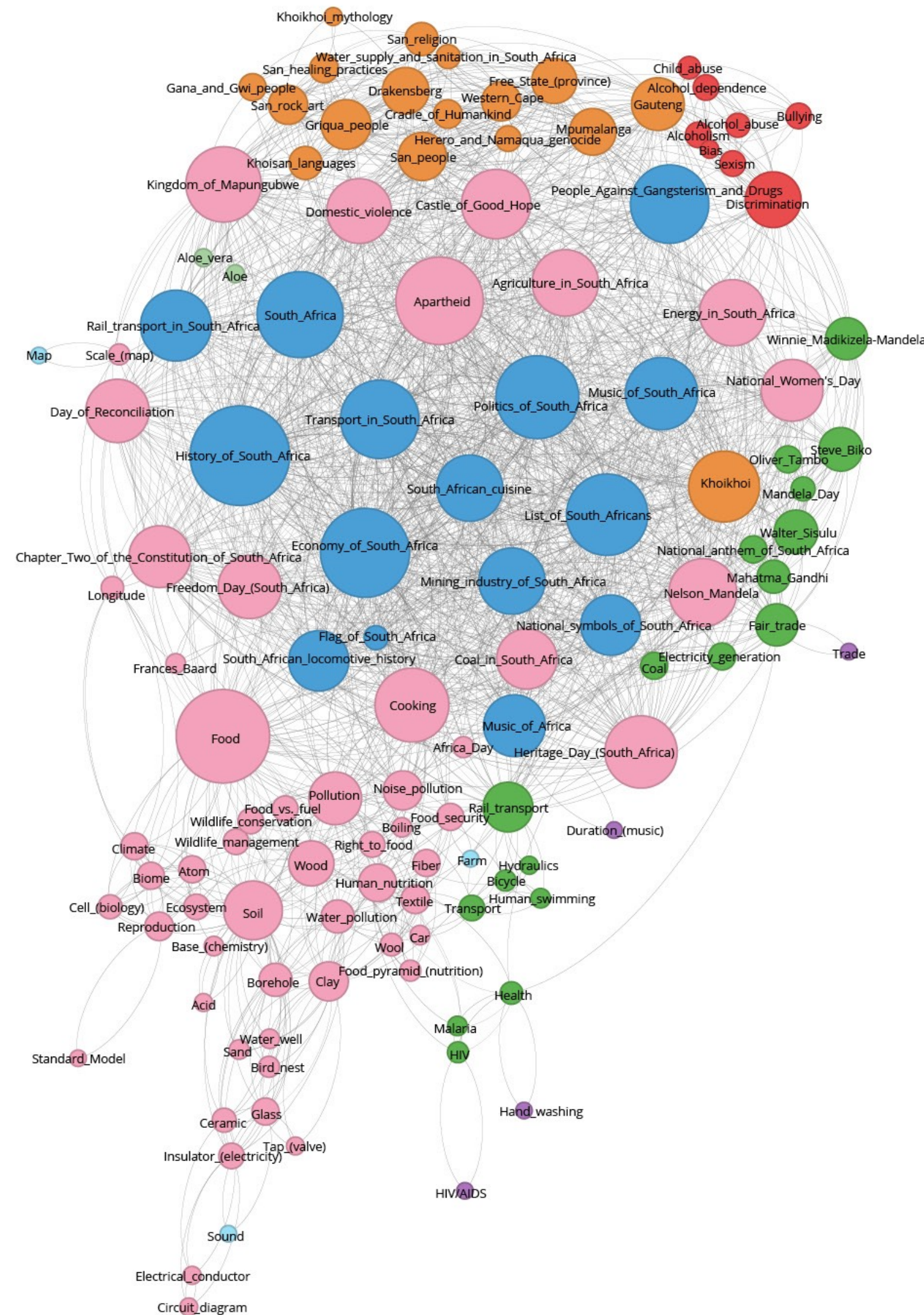
Rete di articoli

marzo 2016

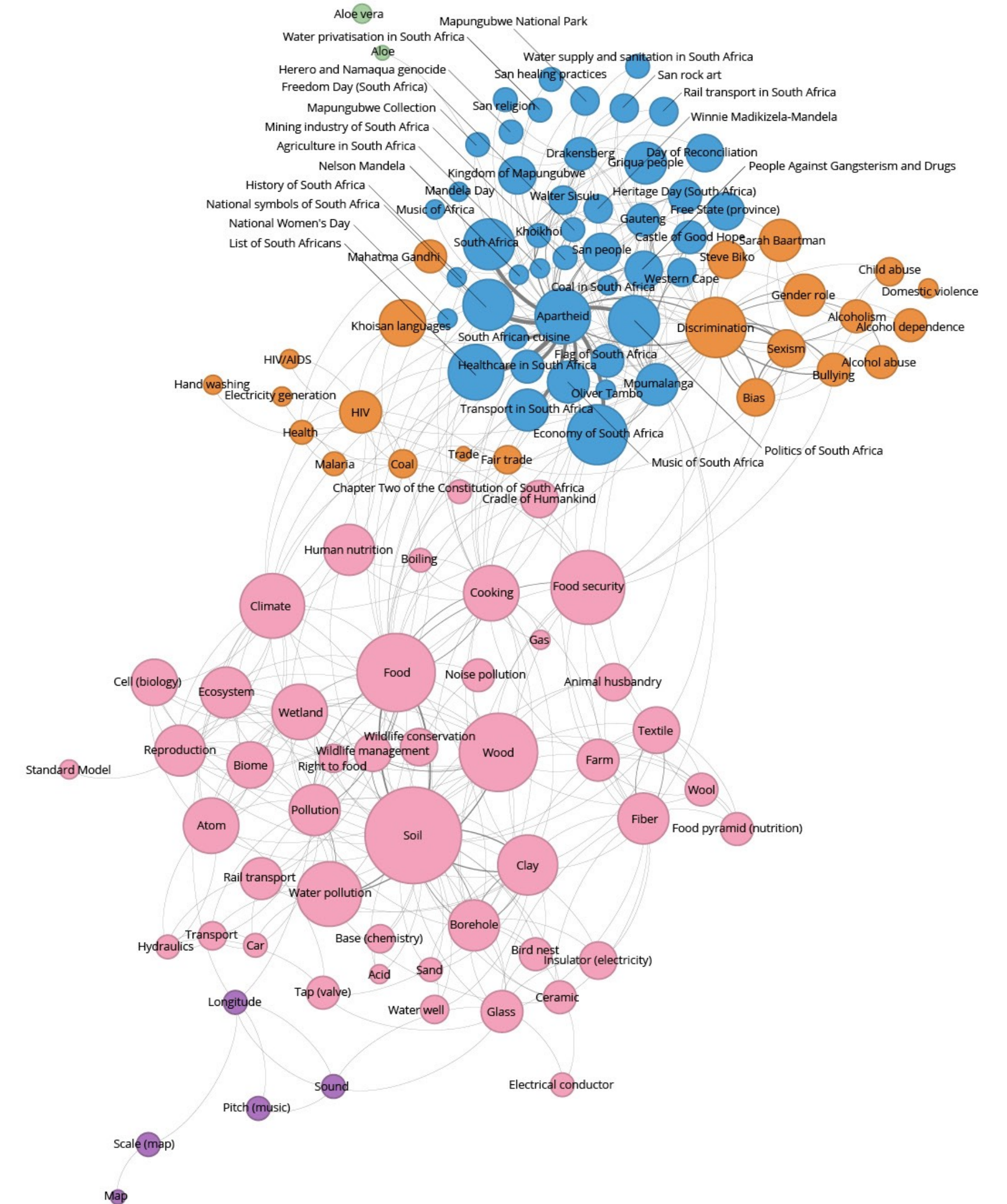




Rete di articoli



August 2015



August 2017

altri dati

Features degli articoli ottenuti con jQuery

- issues
- references
- notes
- images
- see also

```
features.js
542 function get_images(url) {
543
544     $.ajax({
545         type: 'GET',
546         url: my_proxy + wikilink + url, //my_proxy + url,
547         processData: true,
548     })
549     .done (function (get_ref) {
550
551         var parsedata_func = $.parseHTML(get_ref)
552
553         get = [];
554         get1 = [];
555         get2 = [];
556
557         get1 = $(parsedata_func).find('.thumbimage')
558         get3 = ($(parsedata_func).find('.thumb').find('a').find('img'));
559
560         sum = 0;
561
562         var url_clean = url.replace('https://en.wikipedia.org/wiki/', '').replace(
563
564         container = $('#output')
565
566         $.each( get1, function( i, val ) {
567             txt = $(this).prop('href');
568             if (txt !== undefined) {
569                 console.log(txt);
570                 sum++;
571             }
572             else{}
573             //sum++;
574         })
575
576         jQuery.each( get3, function( i, val ) {
577             if (val !== undefined) {
578                 sum++;
579             }
580         })
581
582         container.append(url_clean + ',' + sum + '</br>')
583
584     })
585     .error (function (xhr, ajaxOptions, thrownError) {
586         console.log(xhr.status);
587         console.log(thrownError);
588     });
589 }
590
591 function get_all_images() {
592     container = $('#output')
593     container.append('article,images</br>')
594     jQuery.each( articles_a, function( i, val ) { // list; articles_a;
```


Features of the articles

The visualization shows the amount of issues, references, notes, images and see also for every article (176).

Legend

- issues
- references
- notes
- images
- see also
- new article under examination
- ⊗ reviewed by the community
- ⊕ reviewed by an expert
- ⊗⊕ reviewed by the community and by an expert
- benchmark previous analysis

How to read the visualization

On the top the bars show the number of issues. On the bottom the bars show the amount of references, notes, images and see also. From left to right, articles are in ascending order of features.

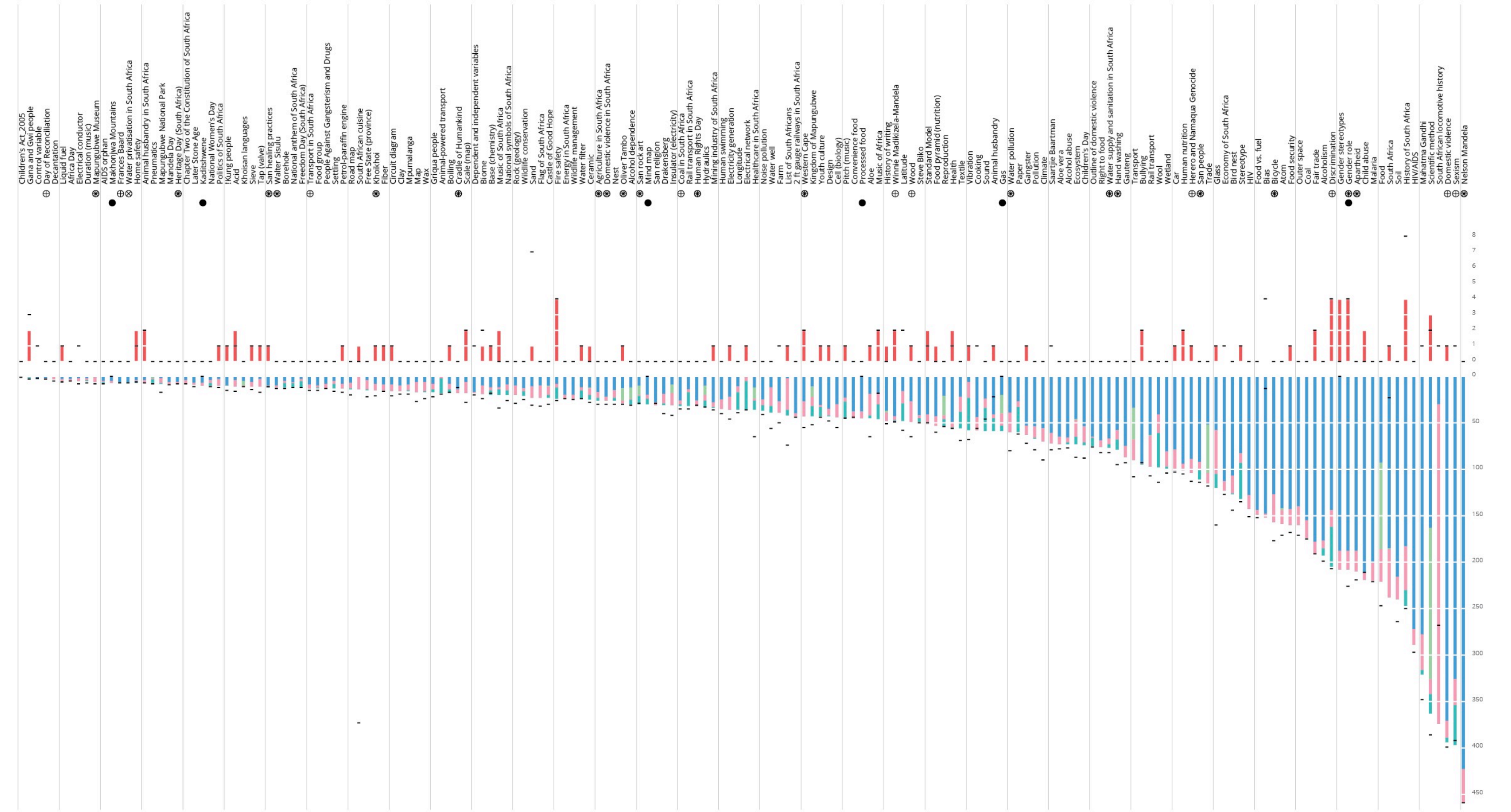
Please note: issues and the other features are shown at different scales.

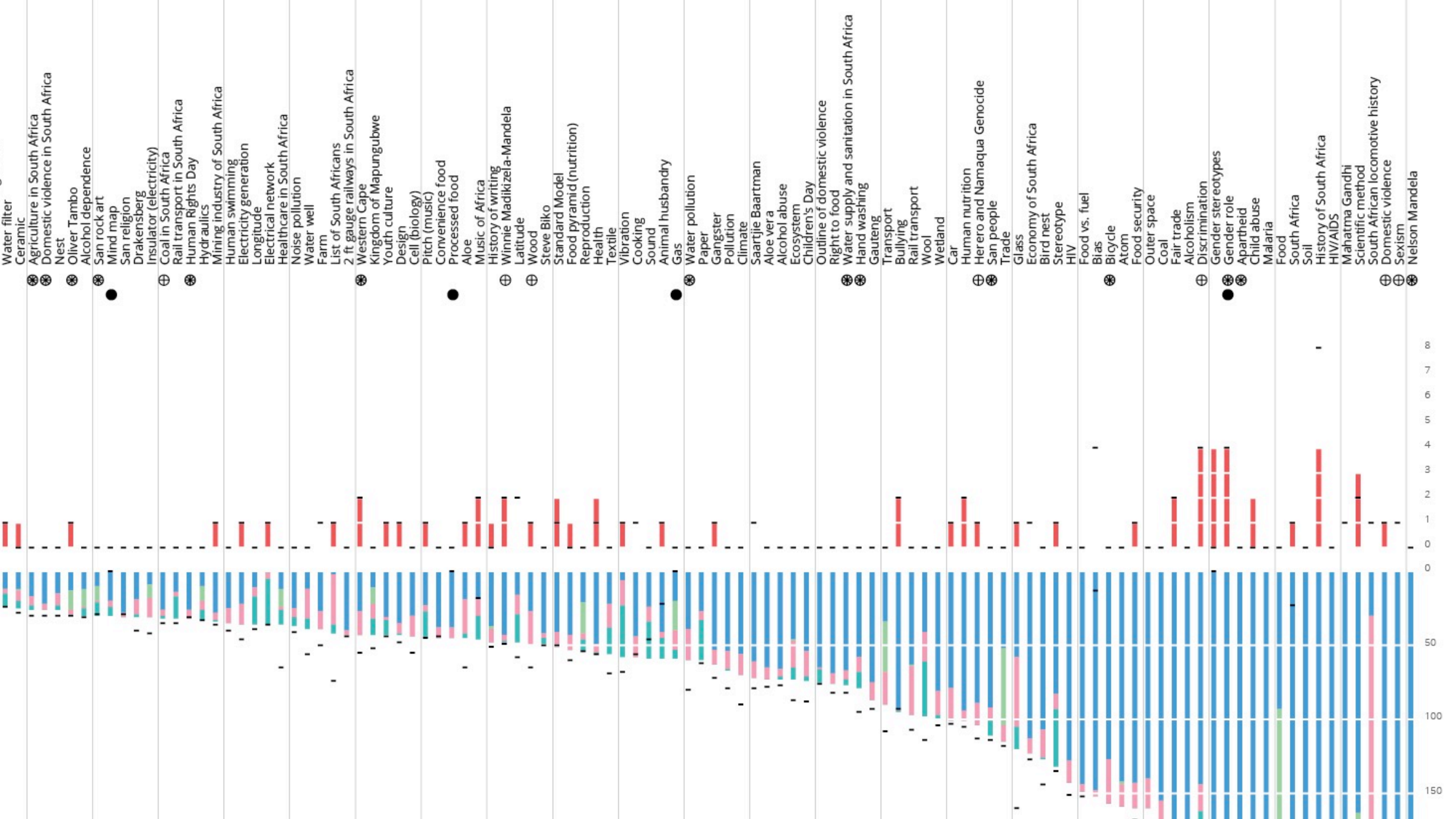
SUPSI

Department for Environment Constructions and Design
Laboratory of visual culture

Project by

Giovanni Profeta
giovanni.profeta@supsi.ch





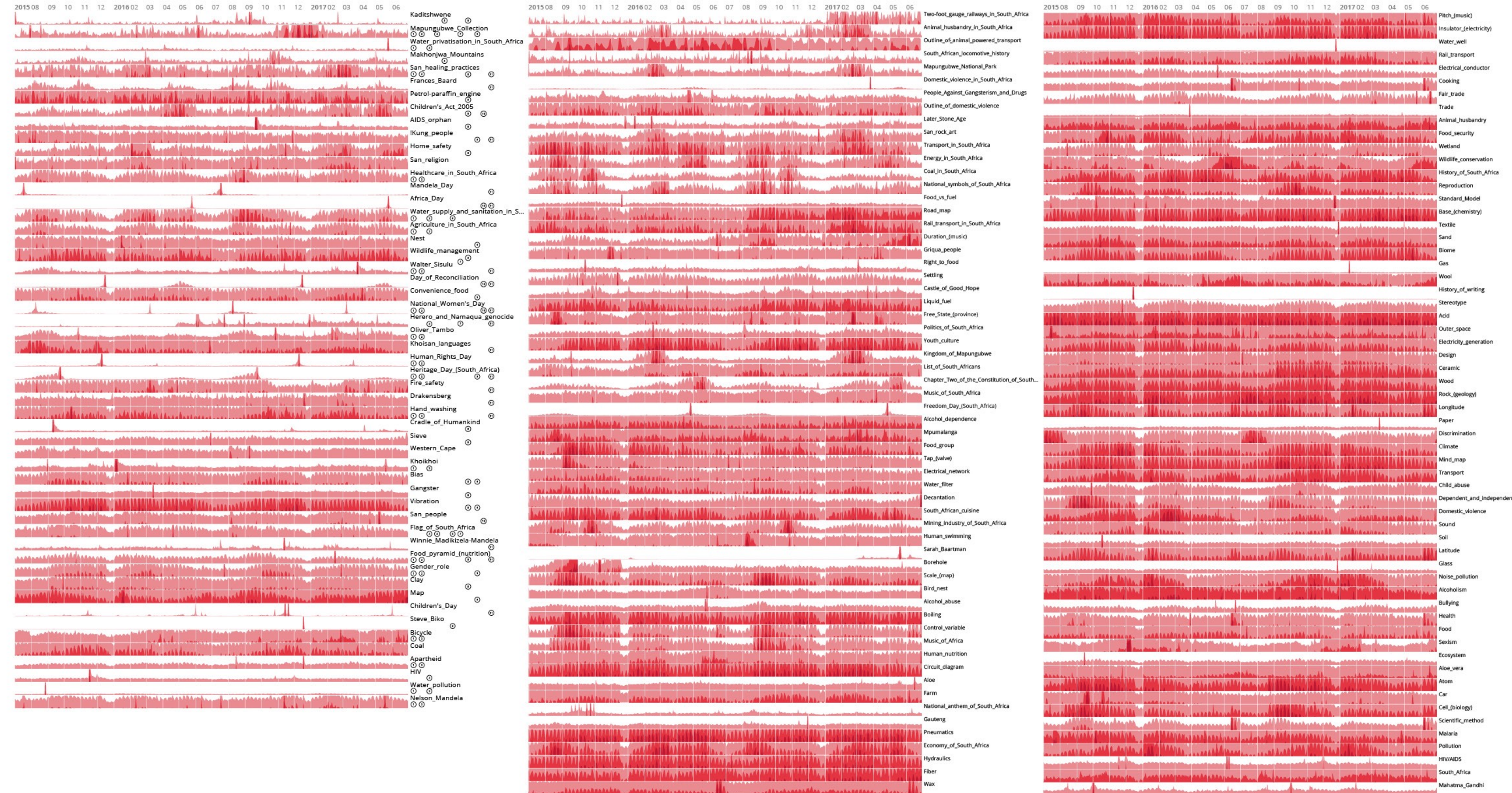
Page views

The visualization shows the page views for every article from July 2015 to June 2017.

How to read the visualization

The timelines show the amount of daily page views for every article. From top to bottom, from left to right, articles are in ascending order of total page views.

Please note: due to very high peaks, the page views are displayed as horizon charts. Thus it is possible to understand the trends of the page views.



- Legend**
- ① Review by community
 - ② Review by expert (pdf)
 - ③ Review by expert (pdf and wiki)
 - ④ New article suggested by expert
 - ⑤ New article created (AFC)
 - ⑥ Featured on WikiProject South Africa
 - ⑦ Rewrite based on expert review
 - ⑧ Article assessment
 - ⑨ Bold reassessment
 - ⑩ Africa Destubathon
 - ⑪ Edit-a-thon

SUPSI

University of Applied Sciences and Arts of Southern Switzerland - Laboratory of visual culture

Research team

Iolanda Pensa (principal investigator, Switzerland), Tobias Schonwetter (principal investigator, South Africa), Luca Botturi, Florence Devouard, Giancarlo Gianocca, Erica Litrenta, Giovanni Profeta, Marta Pucciarelli, Kelsey Wiens

Data visualization

Giovanni Profeta, giovanni.profeta@supsi.ch

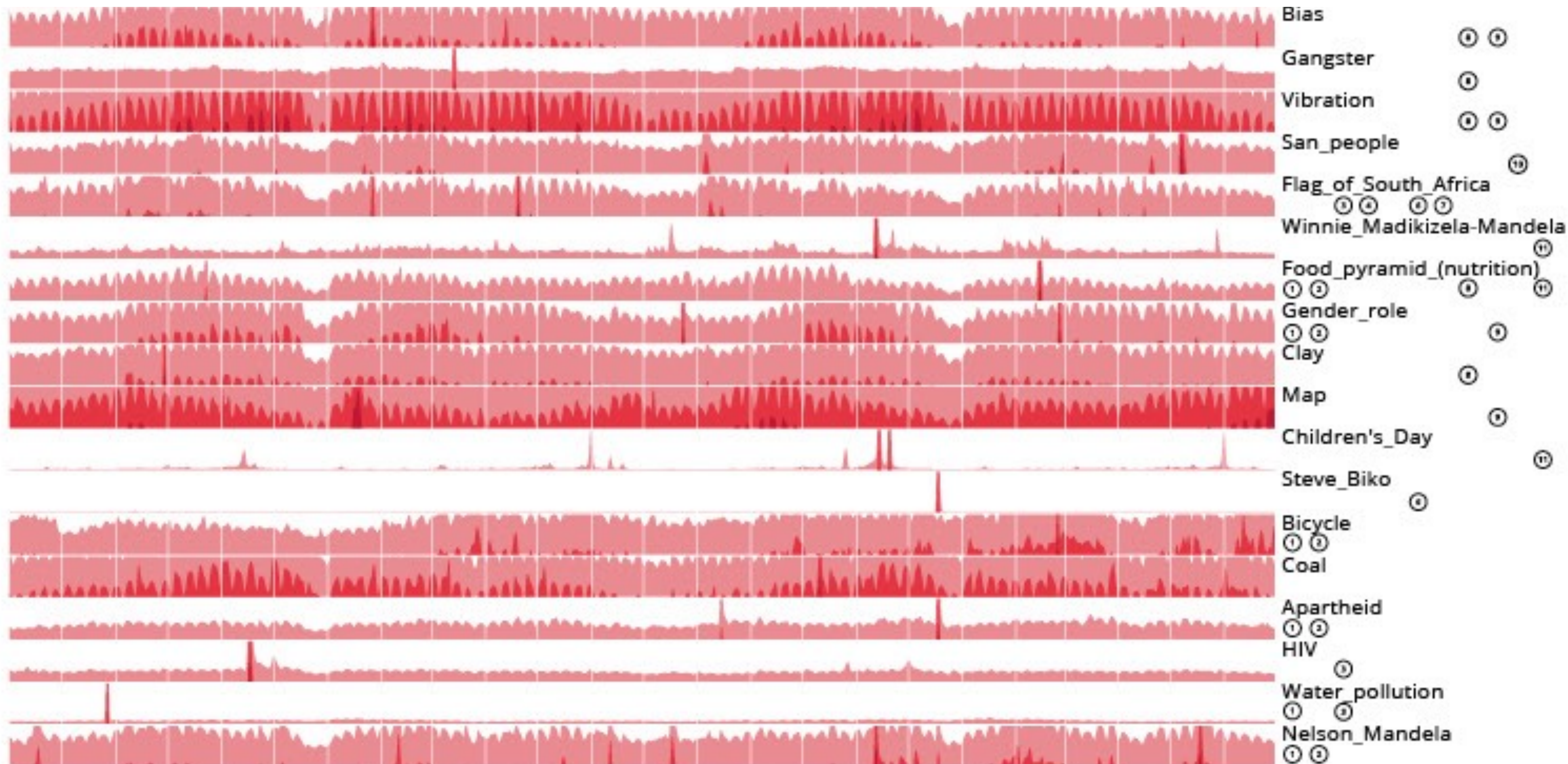
For more information

<http://bit.ly/WikipediaPS>

Project funded by



Pagine viste



Dati del 2015 al 2017

strumenti per proseguire le analisi

Protocolli di ottenimento e visualizzazione dati

Articles network

The visualization aims to explore the correlation among the articles under consideration of the project.

Visualization: Every article is represented by a bubble. The big is the bubble the more incoming links it has. The close bubbles are the more incoming links they have in common. The colors represent the clusters of articles according to their modularity class.

Data scraping

1. Go to [wikimole/scrapper/lib/scrapper.py](#)
2. Run the function `get_incoming_links()` and get the file with the list of incoming links

Data parsing

1. Go to [wikimole/scrapper/lib/parser.py](#)
2. Run the function `count_co_occurrences()` to get the raw file with the list of co-occurrences
3. Create an Excel file and import the file gathered through the previous step
4. Make a pivot table with the following columns: couple (source + target), count (number of co-occurrences)
5. Import the pivot table in a text editor
6. Replace "---" with a tab and insert the header: source, target, weight. Then you got the edges file
7. Make the nodes file with the following header: id (article title), label (article title)

Data visualization

1. Open a new Gephy project
2. In the Data laboratory tab: Import the nodes file
3. Import the edges file
4. In the Overview tab: Run the modularity class (resolution 1)
5. Go to Appearance/Nodes/Color/Ranking and choose the attribute Modularity Class
6. Go to Appearance/Nodes/Size/Ranking and choose the attribute In-Degree (min size 3, max size 20)
7. Run the Force Atlas 2, then apply the Prevent overlap
8. Run the Noverlap layout
9. In the Preview tab: set curved edges, nodes labels with no proportional size
10. Export the svg file to be further edited in a vector graphics editors

Wikimole scraper

The screenshot shows the GitHub repository page for 'giovannipro / wikimole'. The repository is described as a 'Scraper and set of data visualizations made in the framework of Wikipedia Primary School research project'. It has 159 commits, 2 branches, 0 packages, 0 releases, and 1 contributor. The repository is licensed under GPL-3.0. The commit history shows several updates to the protocol and data files. The README.md file is visible at the bottom of the screenshot.

Search or jump to... Pull requests Issues Marketplace Explore

giovannipro / wikimole Unwatch 2 Star 2 Fork 0

Code Issues 0 Pull requests 1 Actions Projects 0 Wiki Security 0 Insights Settings

Scraper and set of data visualizations made in the framework of Wikipedia Primary School research project <http://bit.ly/WPS-evaluation> Edit

Manage topics

159 commits 2 branches 0 packages 0 releases 1 contributor GPL-3.0

Branch: master New pull request Create new file Upload files Find file Clone or download

Giovanni Profeta and Giovanni Profeta update the protocol Latest commit f221027 on 11 Aug 2017

data	make the timeline of page views	3 years ago
scraper	update the protocol	3 years ago
views	update the protocol	3 years ago
.gitignore	Merge branch 'glam_tool'	3 years ago
LICENSE	update articles network and make the in/out balance visualization	3 years ago
README.md	update the protocol	3 years ago

README.md

Wikimole

Wikimole is a scraper and a set of data visualizations made in the framework of [Wikipedia Primary School Project](#).

Wikipedia Primary School is an applied research involving a team lead by [SUPSI University of Applied Sciences and Arts of](#)

Conclusioni

Le visualizzazioni che più hanno rappresentato l'impatto del progetto sugli articoli sono quelle sui link e sulle features

L'impatto di pochi utenti può essere determinate nel successivo sviluppo di articoli appartenenti ad un'area tematica

Per un ristretto numero di articoli è possibile guidare l'interesse nell'editare Wikipedia

Grazie!

✉ giovanni.profeta@supsi.ch

W G.prof

🐦 profeta_g



Giovanni Profeta, Lugano 2020

Wikipedia Primary School (2014-2017)

Team di progetto: Iolanda Pensa, Tobias Schönwetter, Luca Botturi, Florence Devouard, Giancarlo Gianocca, Erica Litrenta, Giovanni Profeta, Marta Pucciarelli, Isla Haddow-Flood, Kelsey Wiens

Enti finanziatori: Fondo Nazionale Svizzero e Sud Africa National Research Foundation