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

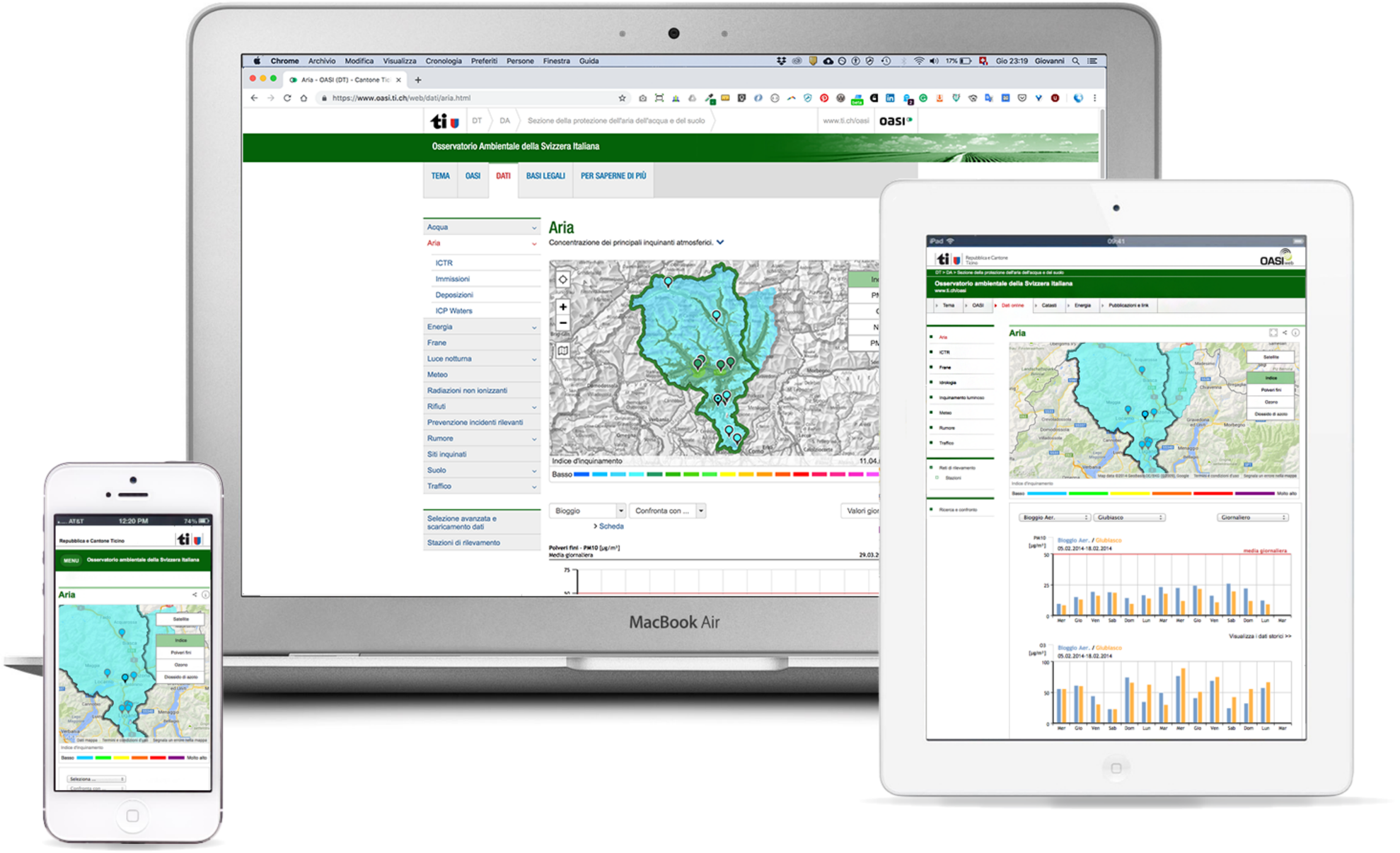
Il mio lavoro

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

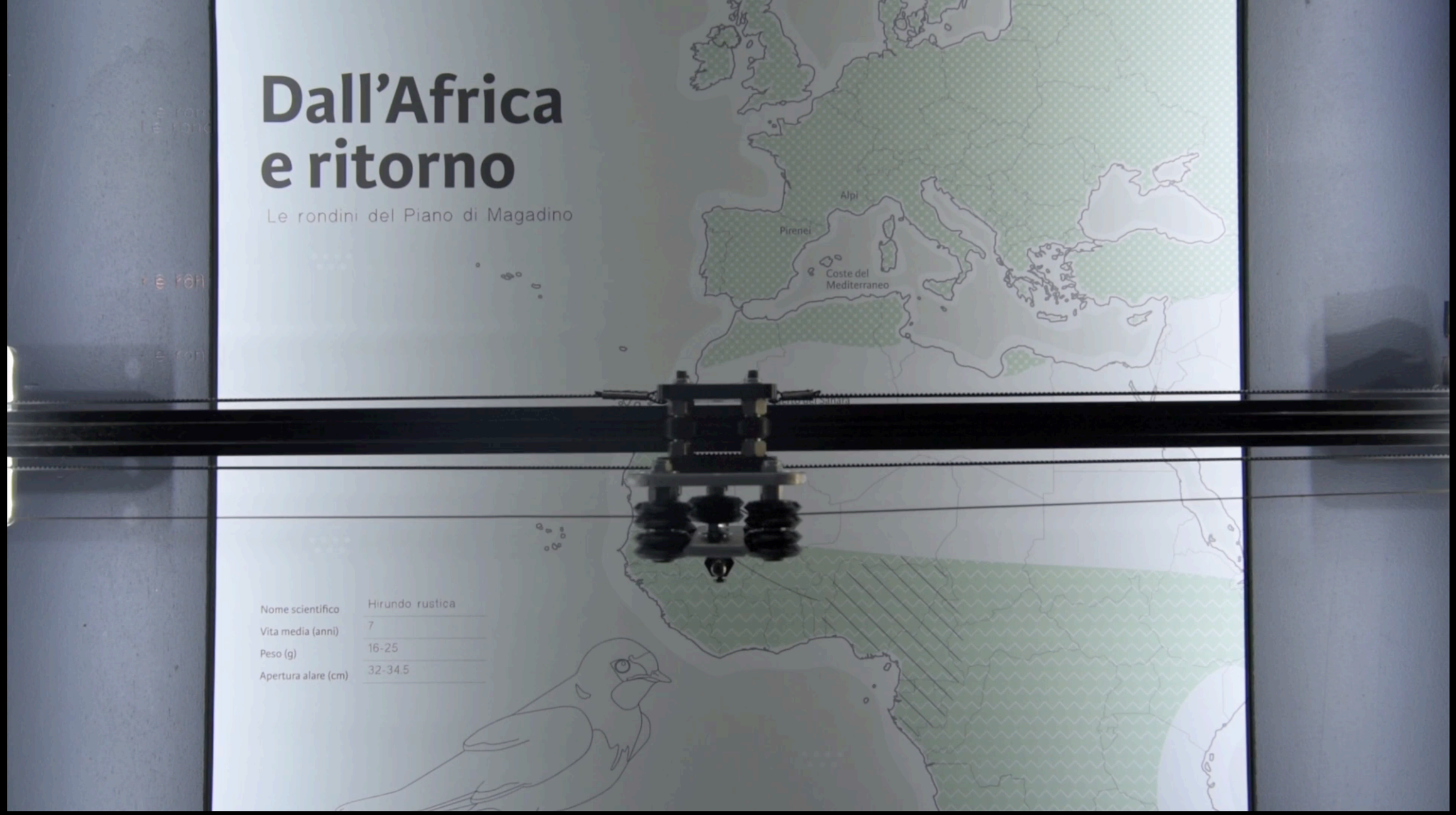
Docente presso il bachelor in Comunicazione Visiva SUPSI

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



Marco Lurati, Giovanni Profeta, Serena Cangiano
Dall’Africa e ritorno. Plotter interattivo. Museo Cantonale di Storia Naturale, Lugano (2015)

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

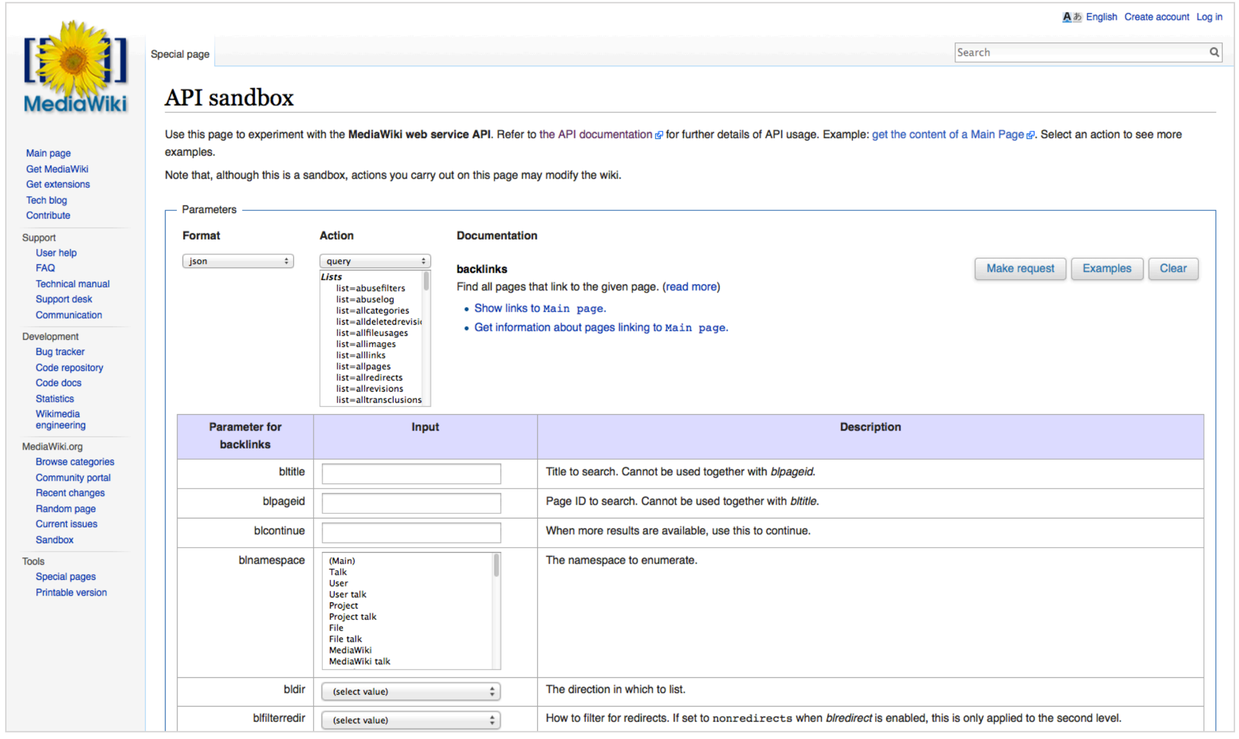
Stato finale
Conclusioni
Inverno 2016

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

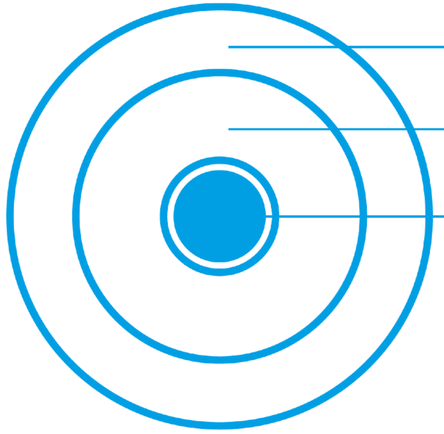
Alcoholism, Animal husbandry and wild animal maintenance, Animal shelters, Animal shelters, Biological sciences, Bullying, Circuits, Crop and stock farming, Discrimination, stereotypes and bias, Domestic violence, Economy - political system, Ecosystem, natural resources, Fair trading, Gangster, Gender 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

https://www.wikipedia.org/



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

Valutazioni dirette e indirette



Valutazioni indirette
Portali, wikiprojects

Valutazioni dirette
Qualità articoli

Articoli
Referenze, wikilinks, dimensione, ecc.

Caratteristiche indagate

Trovabilità

intertestualizzazione degli articoli

Attrattività

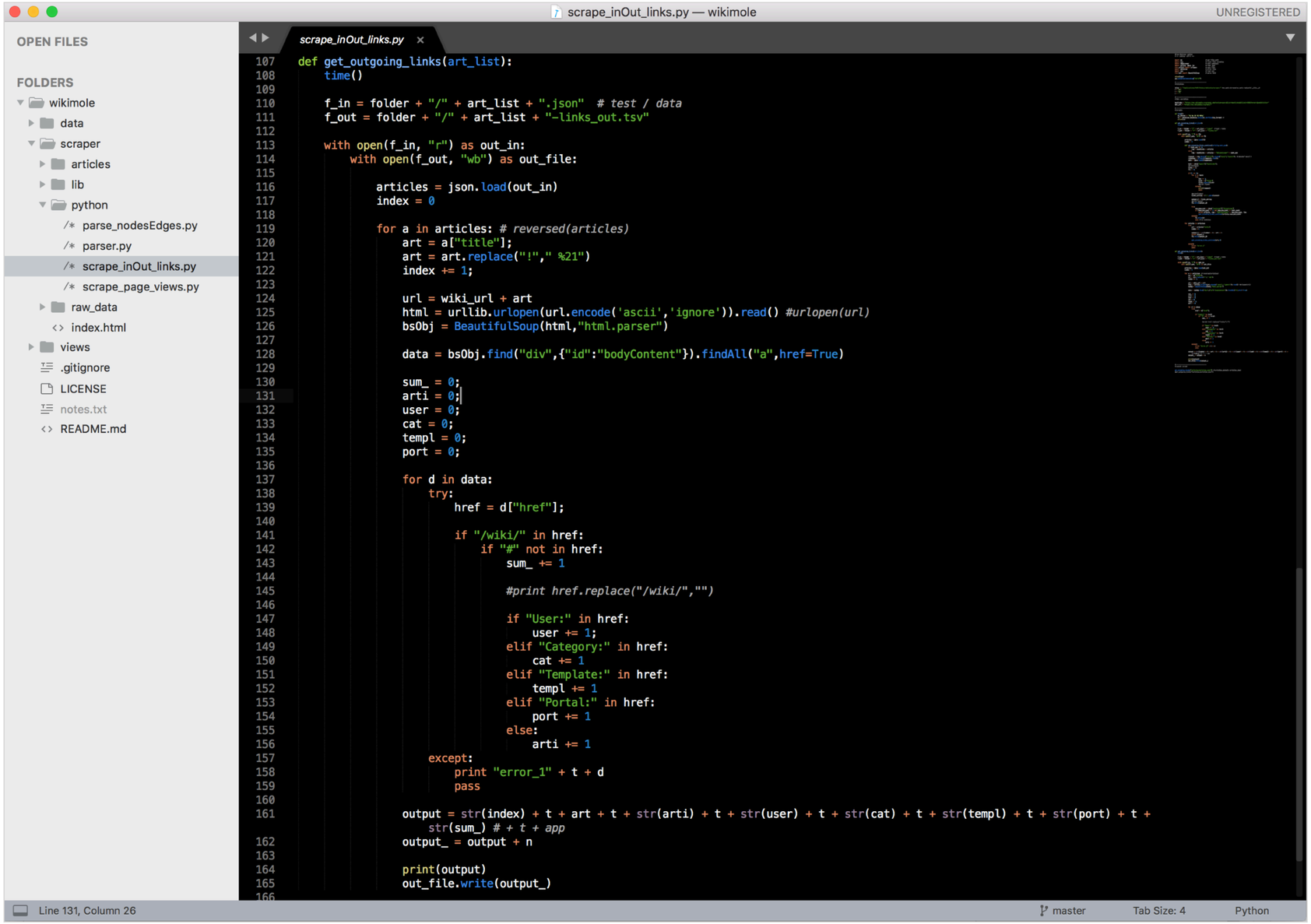
modalità di redazione degli articoli

Specificità

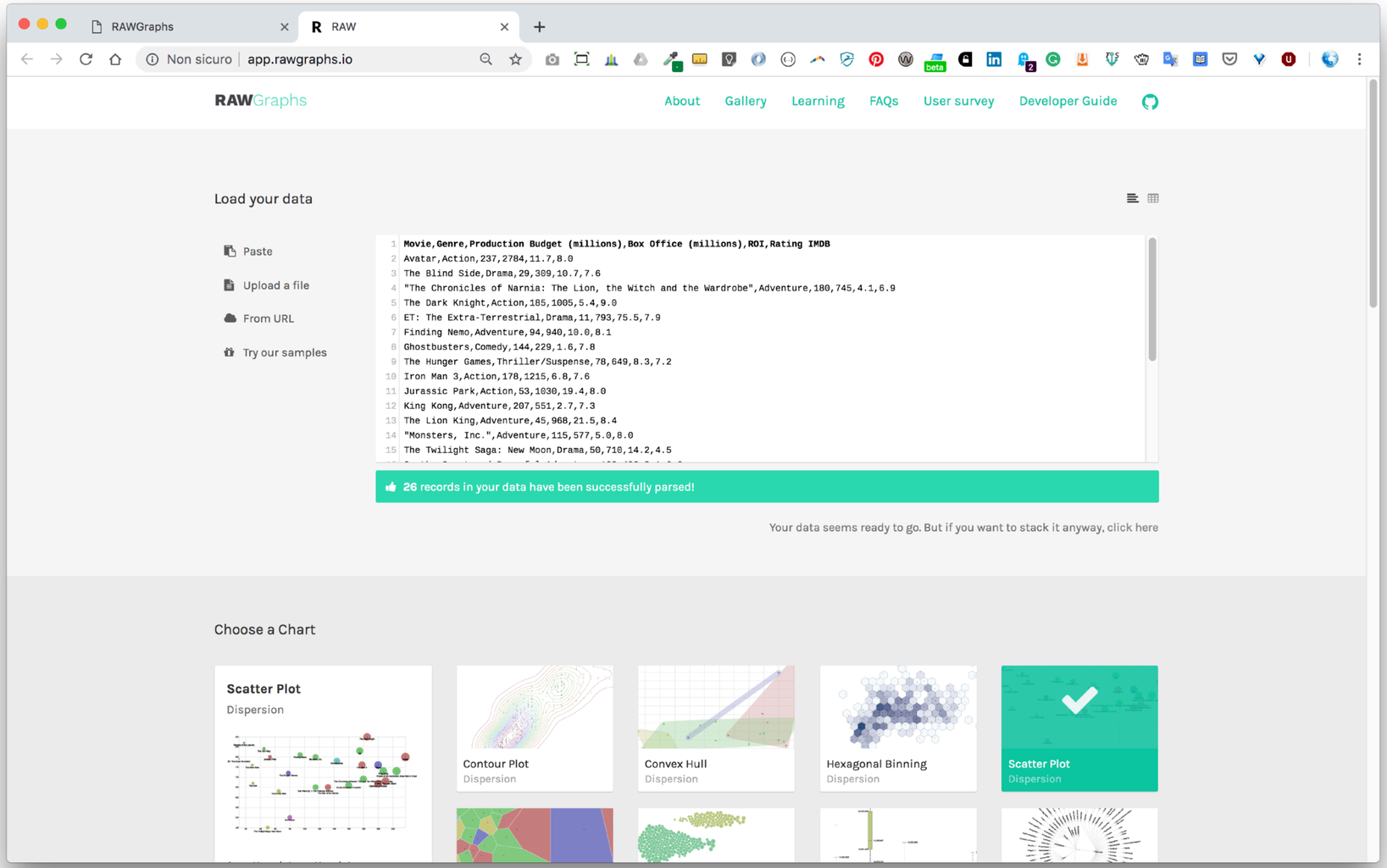
integrazione di risorse negli articoli

link nell’articolo

Ottenimento e elaborazione dati con script in Python

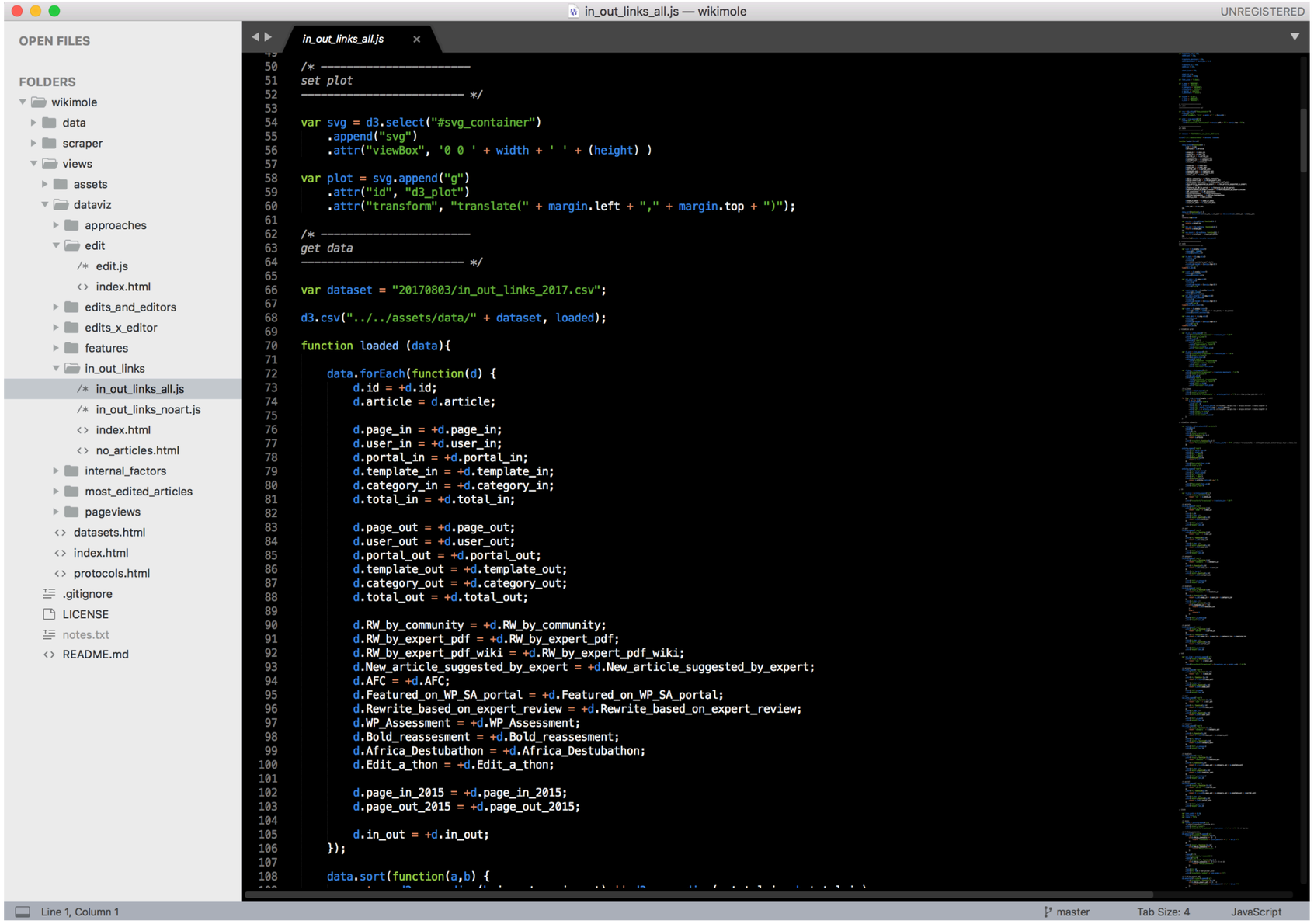


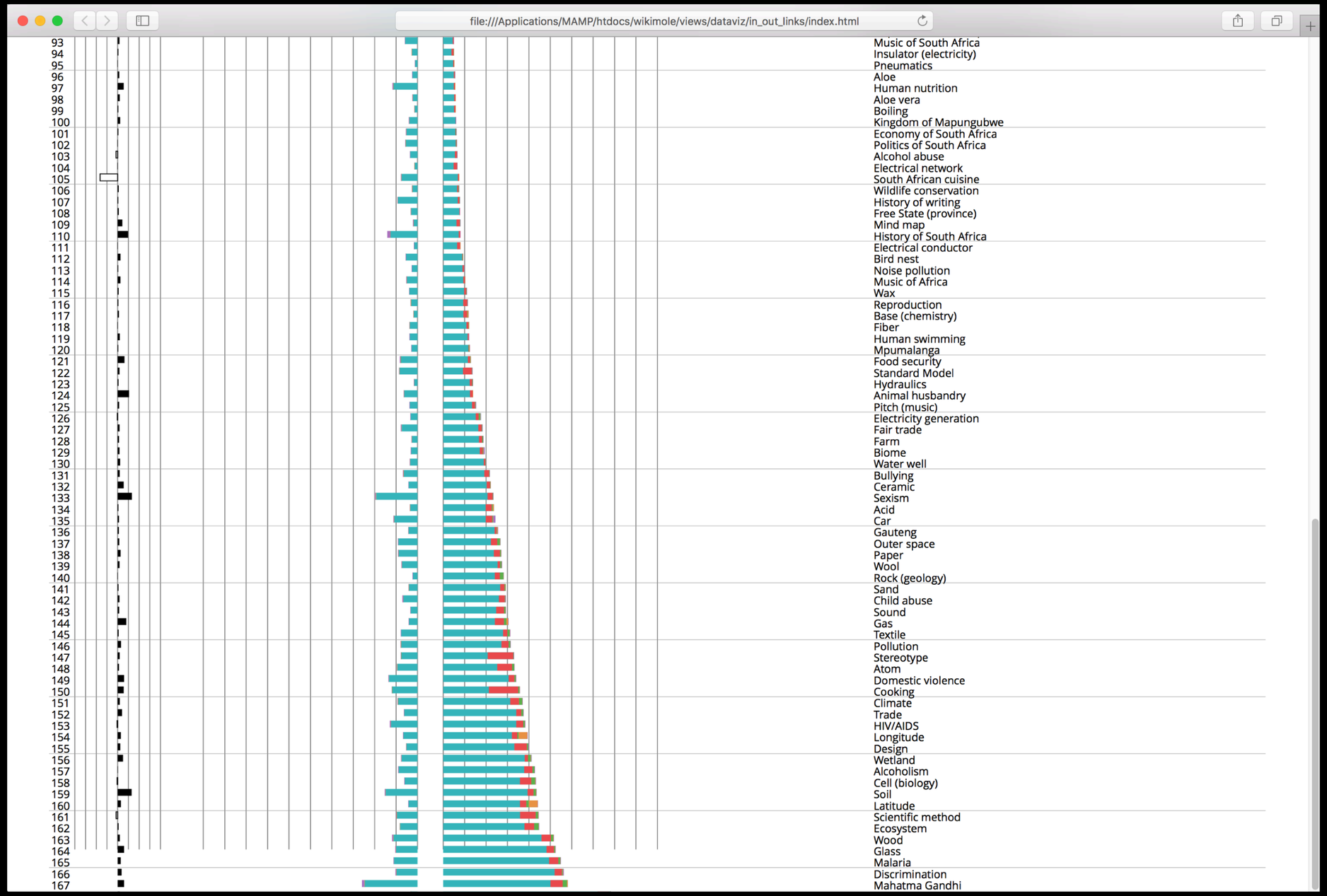
Esplorazione grafica
con RawGraphs



DensityDesign Research Lab, Studio calibro
RawGraphs (2013). Website: https://rawgraphs.io/

Visualizzazione dati con D3.js




file:///Applications/MAMP/htdocs/wikimole/views/dataviz/in_out_links/index.html
93 Music of South Africa
94 Insulator (electricity)
95 Pneumatics
96 Aloe
97 Human nutrition
98 Aloe vera
99 Boiling
100 Kingdom of Mapungubwe
101 Economy of South Africa
102 Politics of South Africa
103 Alcohol abuse
104 Electrical network
105 South African cuisine
106 Wildlife conservation
107 History of writing
108 Free State (province)
109 Mind map
110 History of South Africa
111 Electrical conductor
112 Bird nest
113 Noise pollution
114 Music of Africa
115 Wax
116 Reproduction
117 Base (chemistry)
118 Fiber
119 Human swimming
120 Mpumalanga
121 Food security
122 Standard Model
123 Hydraulics
124 Animal husbandry
125 Pitch (music)
126 Electricity generation
127 Fair trade
128 Farm
129 Biome
130 Water well
131 Bullying
132 Ceramic
133 Sexism
134 Acid
135 Car
136 Gauteng
137 Outer space
138 Paper
139 Wool
140 Rock (geology)
141 Sand
142 Child abuse
143 Sound
144 Gas
145 Textile
146 Pollution
147 Stereotype
148 Atom
149 Domestic violence
150 Cooking
151 Climate
152 Trade
153 HIV/AIDS
154 Longitude
155 Design
156 Wetland
157 Alcoholism
158 Cell (biology)
159 Soil
160 Latitude
161 Scientific method
162 Ecosystem
163 Wood
164 Glass
165 Malaria
166 Discrimination
167 Mahatma Gandhi

Wikipedia Primary School

Incoming and outgoing links

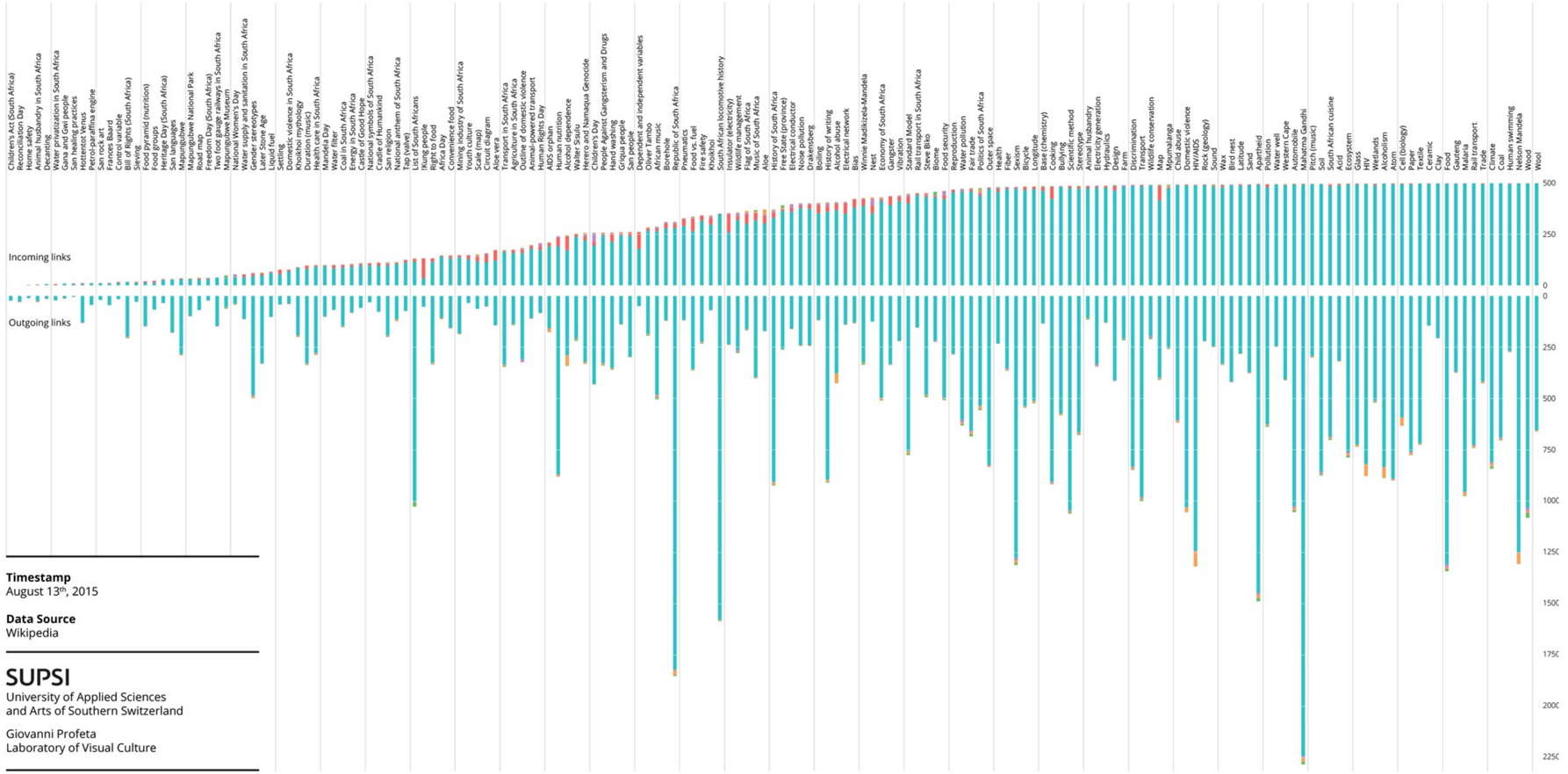
The visualization shows the balance between incoming links (in Wikipedia: "What links here") and wikilinks for every article (171).

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

Agosto 2015

tra la fase 1 e 2

Definizione metodi per il miglioramento degli articoli

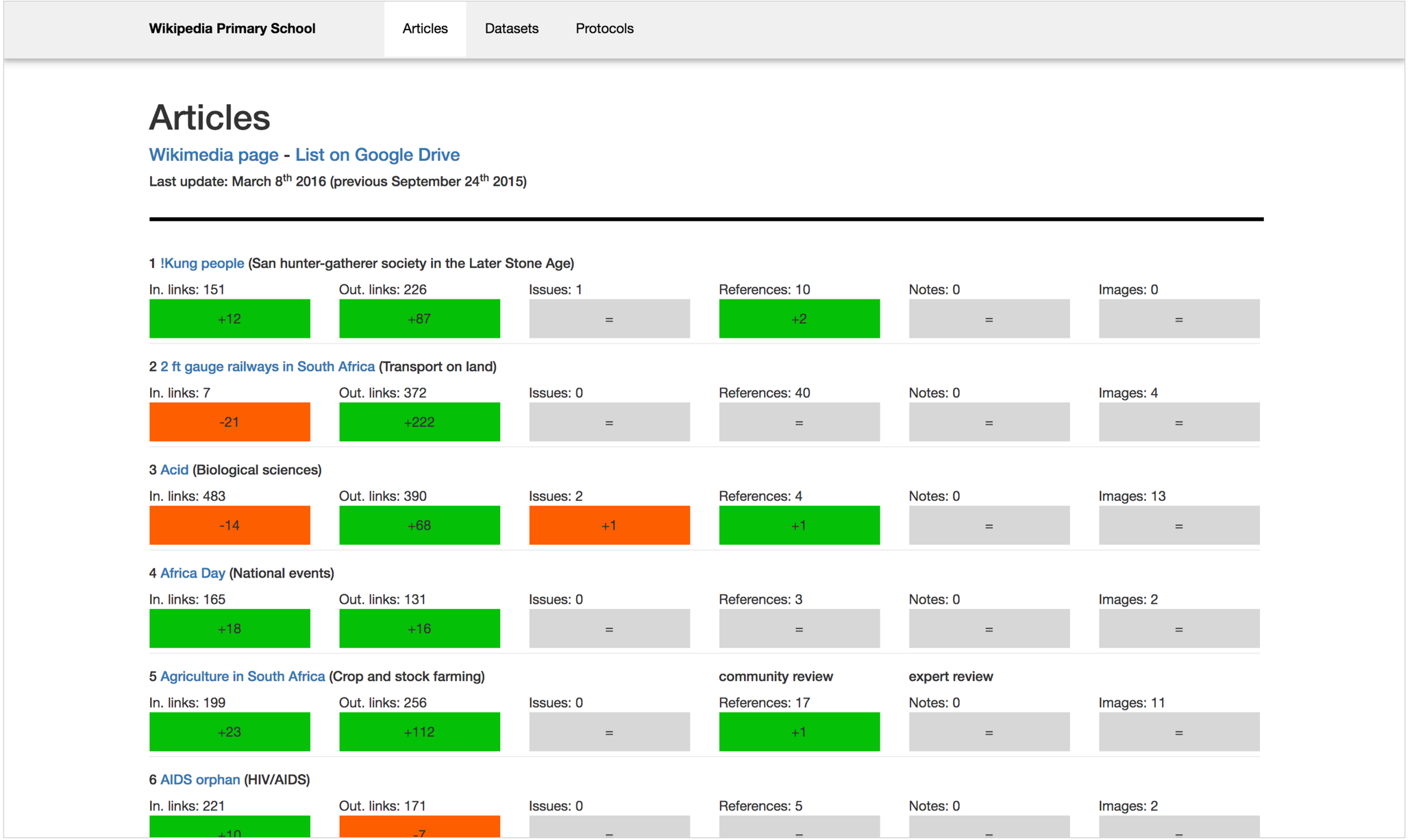
Review by community
Review by expert pdf
Review by expert pdf and wiki
New article suggested by expert
New article created AFC
Featured on WikiProject South Africa
Rewrite based on expert review
Article assessment
Bold reassesment
Africa Destubathon
Edit-a-thon

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
…

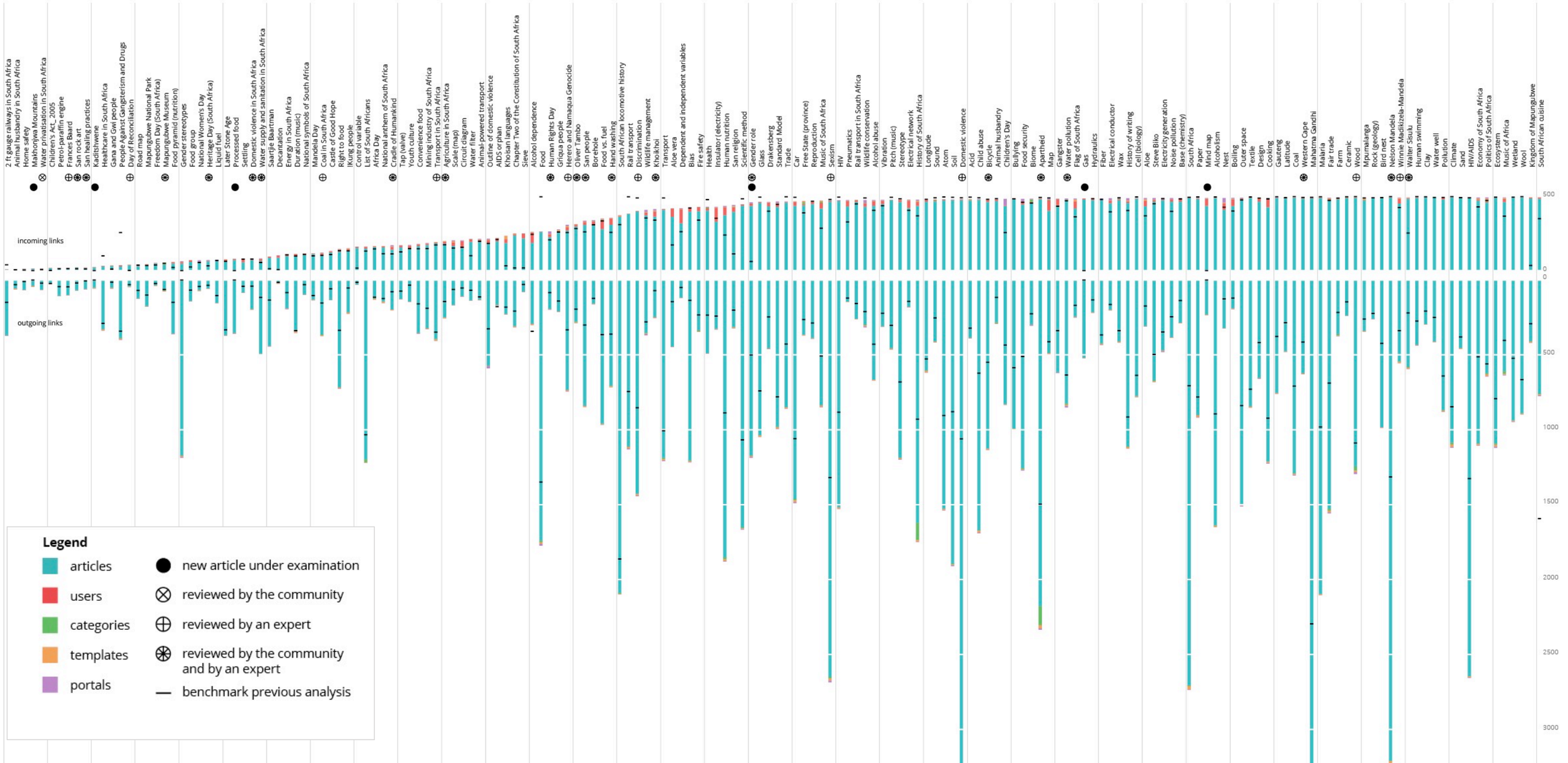
Articoli che hanno cambiato titolo, sono stati aggiunti e cancellati

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

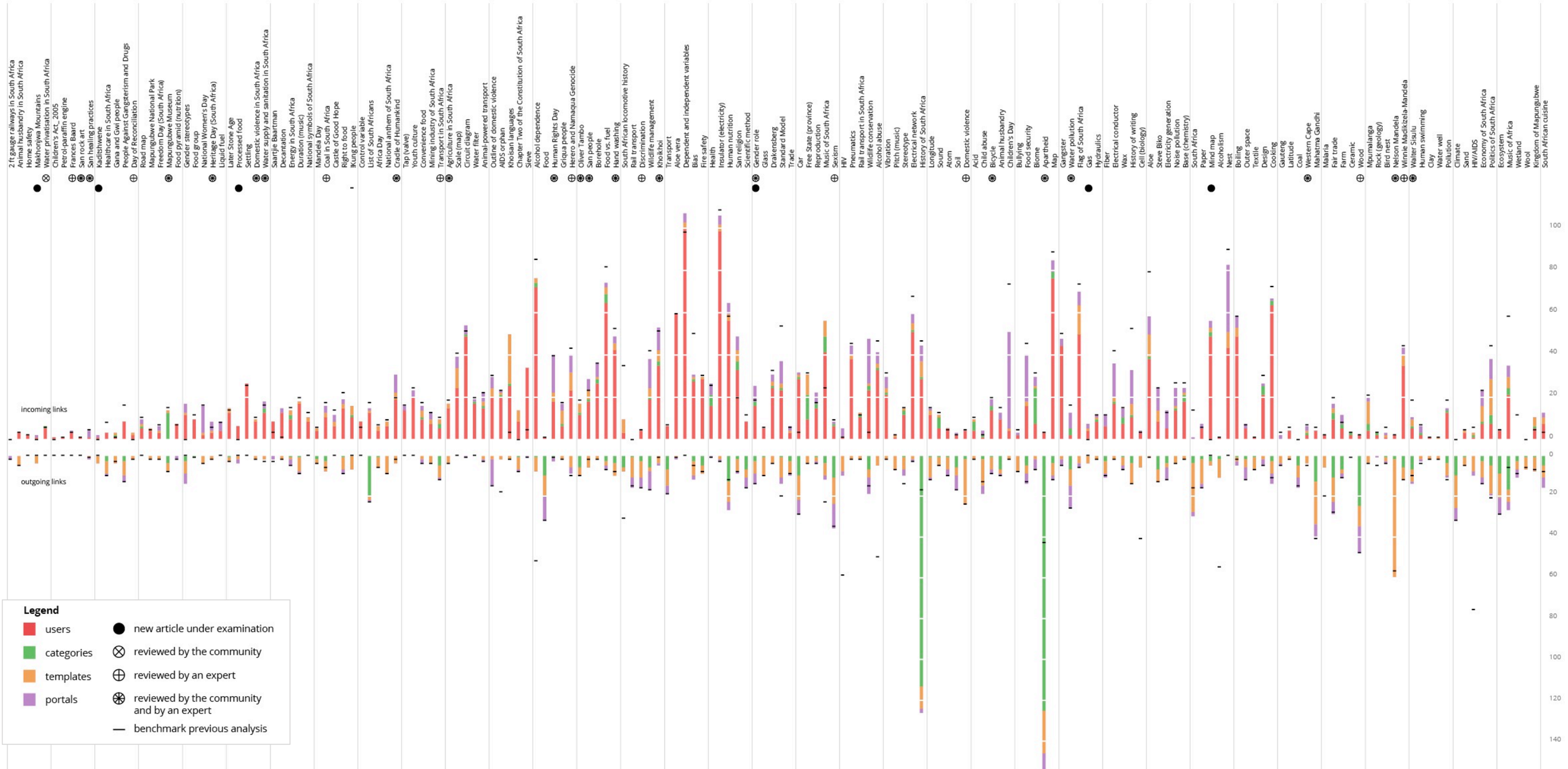


https://giovannipro.github.io/wikipedia-primary-school/

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



Link in entrata e in uscita (esclusi gli articoli)



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
Water filter
Animal-powered transport
Outline of domestic violence
AIDS orphan
Khoisan languages
Chapter Two of the Constitution of South A
Sieve
Alcohol dependence
Food
Human Rights Day
Griqua people
Herero and Namaqua Genocide
Oliver Tambo
San people
Borehole
Food vs. fuel
Hand washing
South African locomotive history
Rail transport
Discrimination
Wildlife management
Khoikhoi
Transport
Aloe vera
Dependent and independent variables
Bias
Fire safety
Health
Insulator (electricity)
Human nutrition
San religion
Scientific method
Gender role
Glass
Drakensberg
Standard Model
Trade
Car
Free State (province)
Reproduction
Music of South Africa
Sexism
HIV
Pneumatics
Rail transport in South Africa
Wildlife conservation
Alcohol abuse
Vibration
Pitch (music)
Stereotype
Electrical network
History of South Africa
Longitude
Sound
Atom
Soil
Domestic violence
Acid
Child abuse
Bicycle
Animal husbandry
Children's Day
Bullying
Food security
Biome
Apartheid
Map
Gangster
Water pollution
Flag of South Africa
Gas
Hydraulics
Fiber
Electrical conductor
Wax
History of writing
Cell (biology)
Aloe
Steve Biko
Electricity generation
Noise pollution
Base (chemistry)
South Africa
Paper
Mind map
Alcoholism
Nest
Boiling
Outer space
Textile
Design
Cooking
Gauteng
Latitude
Coal
Western Cape
Mahatma Gandhi
Malaria
Fair trade
Farm
Ceramic
Wood
Mpumalanga
Rock (geology)
Bird nest
Nelson Mandela
Winnie Madikizela-Mandela
Walter Sisulu
Human swimming
Clay
Water well
Pollution
Climate
Sand
HIV/AIDS
Economy of South Africa
Politics of South Africa
Ecosystem
Music of Africa
Wetland
Wool
Kingdom of Mapungubwe
South African cuisine
80
60
40
20
0
0
20
40
60
80
100

Wikipedia Primary School

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 for Environment Constructions and Design
Laboratory of visual culture

Project by

Giovanni Profeta
giovanni.profeta@supsi.ch

Timestamp: 05-03-2016; benchmark: 13-08-2015;

Data Source: Wikipedia (page scraping); Wikipedia API

Visualization tool: D3.js

This work is licensed under the Creative Commons Attribution-Share Alike 4.0 International license.

Marzo 2016

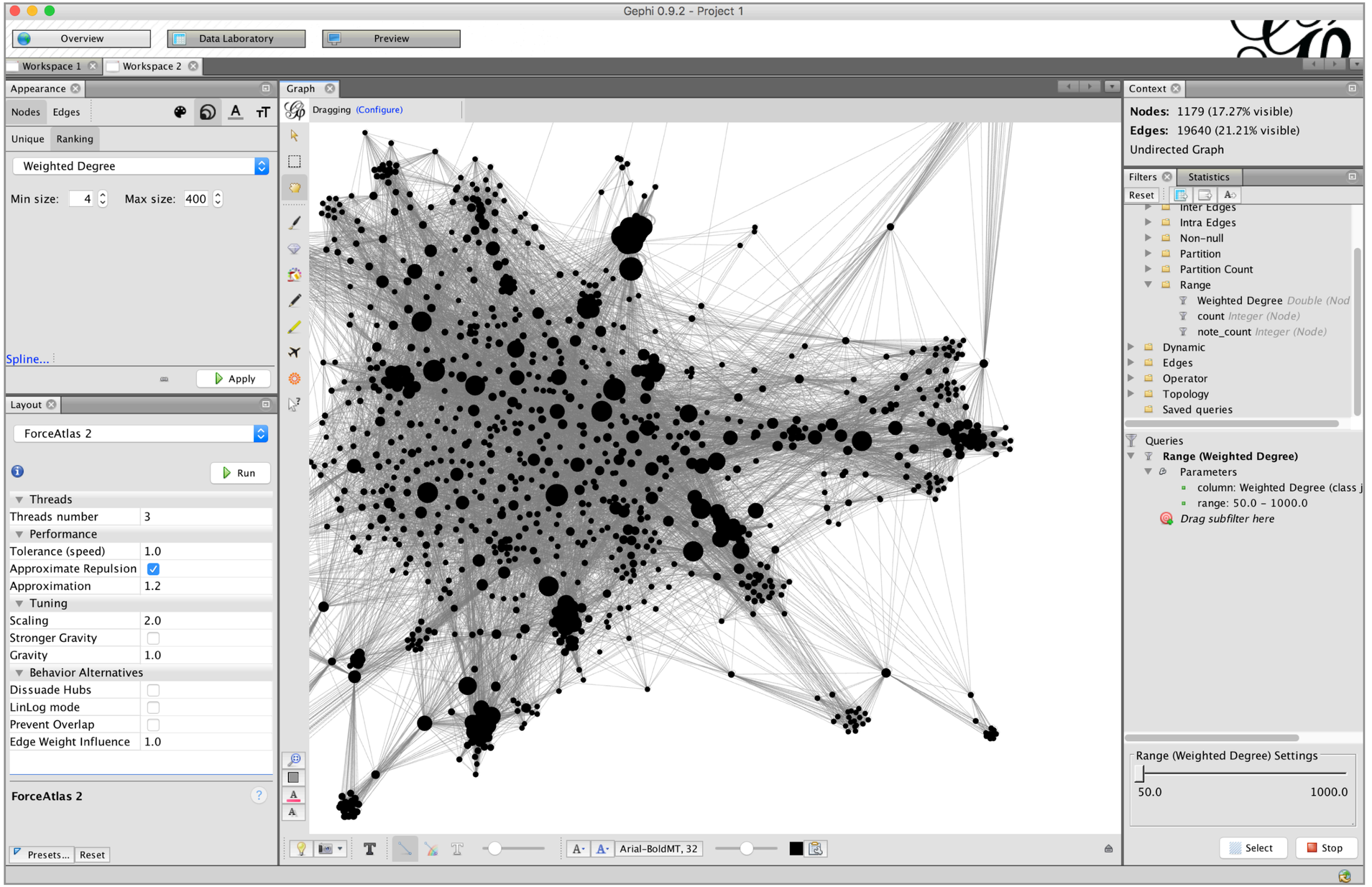
link tra gli articoli

Rete di pagine
basata sui link
in entrata

edges_2017.tsv

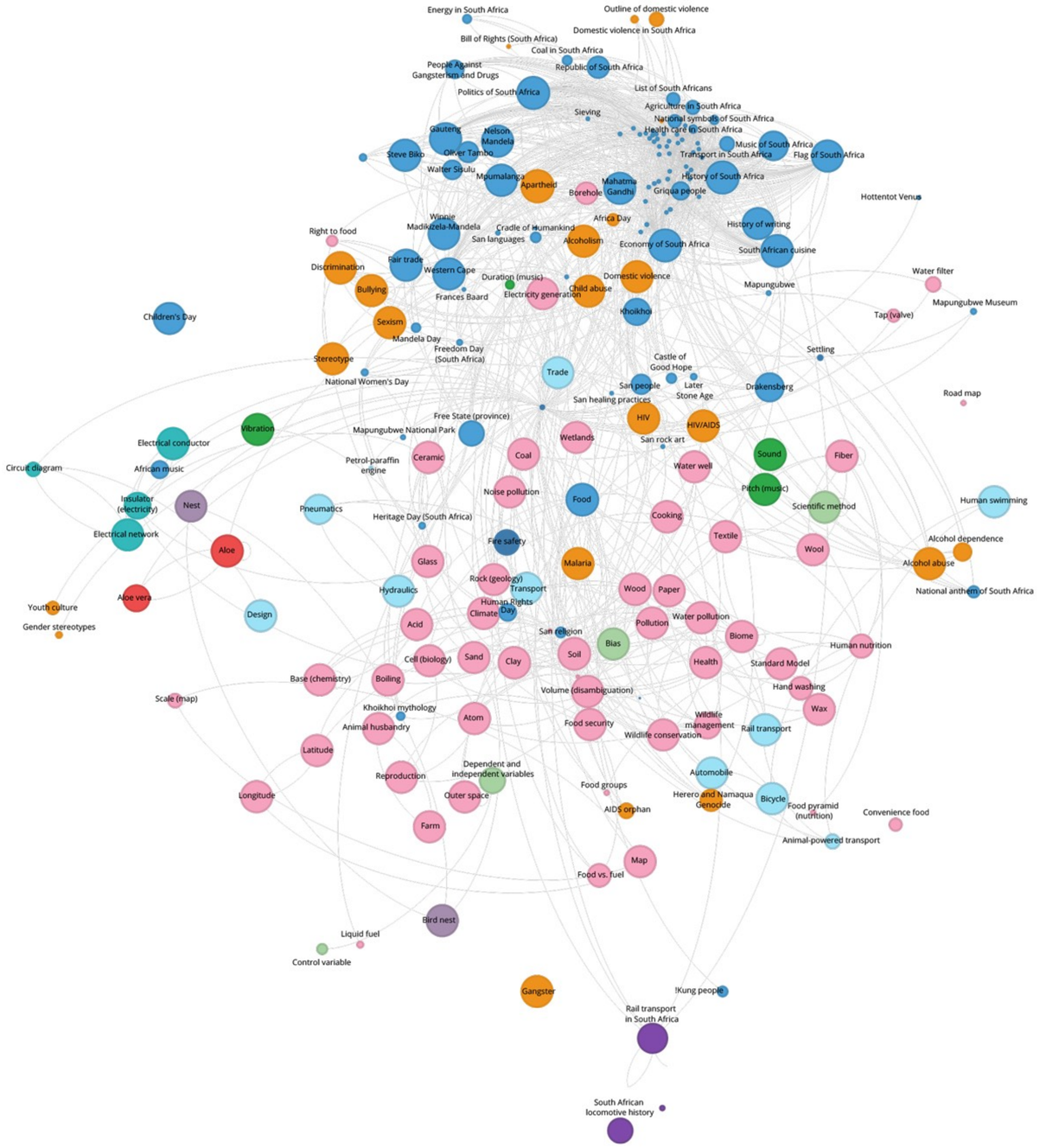
	source	target
2	Acid	Alkane
3	Acid	Arsenic
4	Acid	Americium
5	Acid	Aluminium
6	Acid	Alcohol
7	Acid	Ammonia
8	Acid	Amine
9	Acid	Amide
10	Acid	Ascorbic acid (molecular aspects)
11	Acid	Alkaloid
12	Acid	Aspartame
13	Acid	Acid–base reaction
14	Acid	Acid rain
15	Acid	Boron
16	Acid	Bromine
17	Acid	Bicarbonate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

Gephi

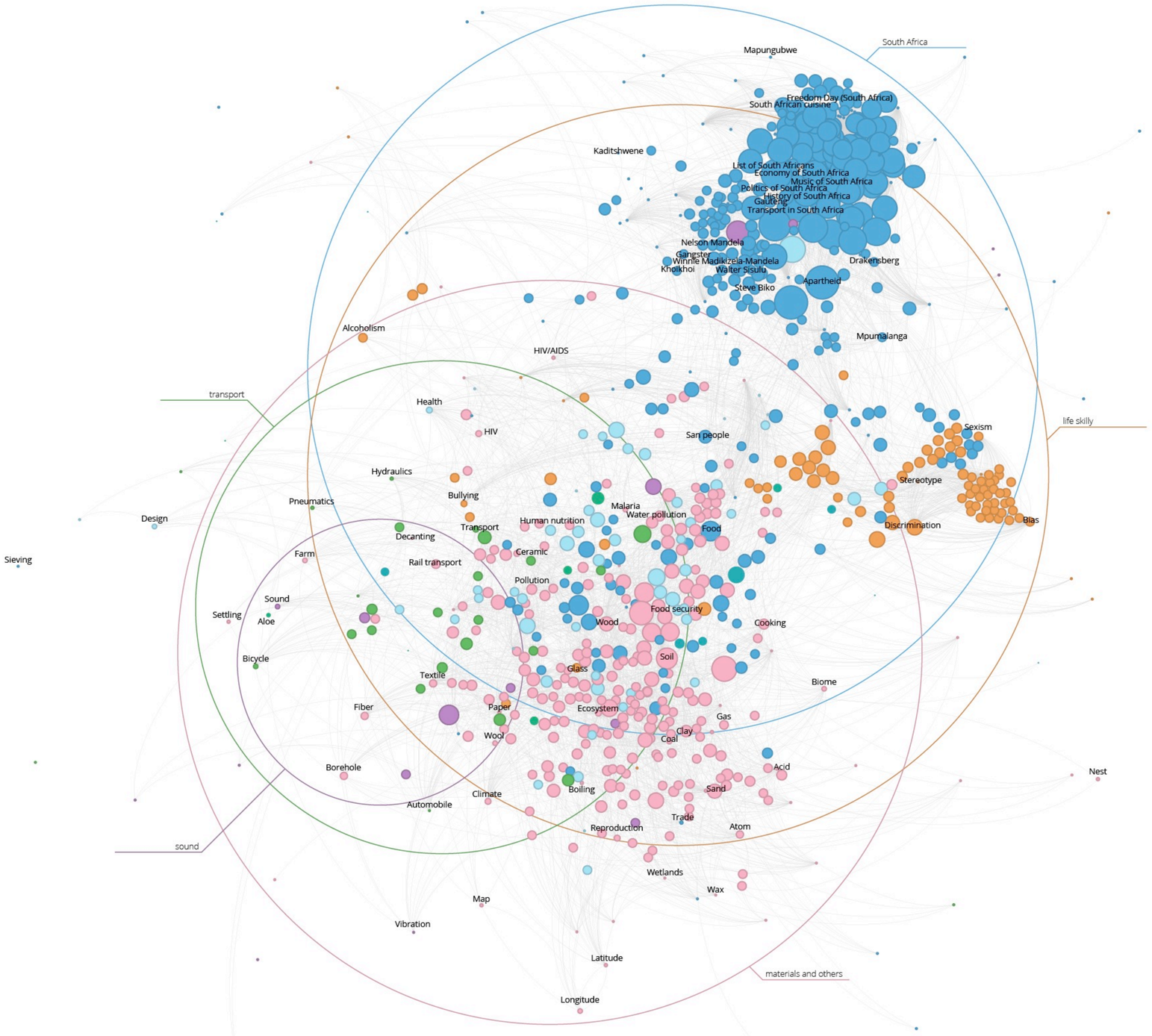


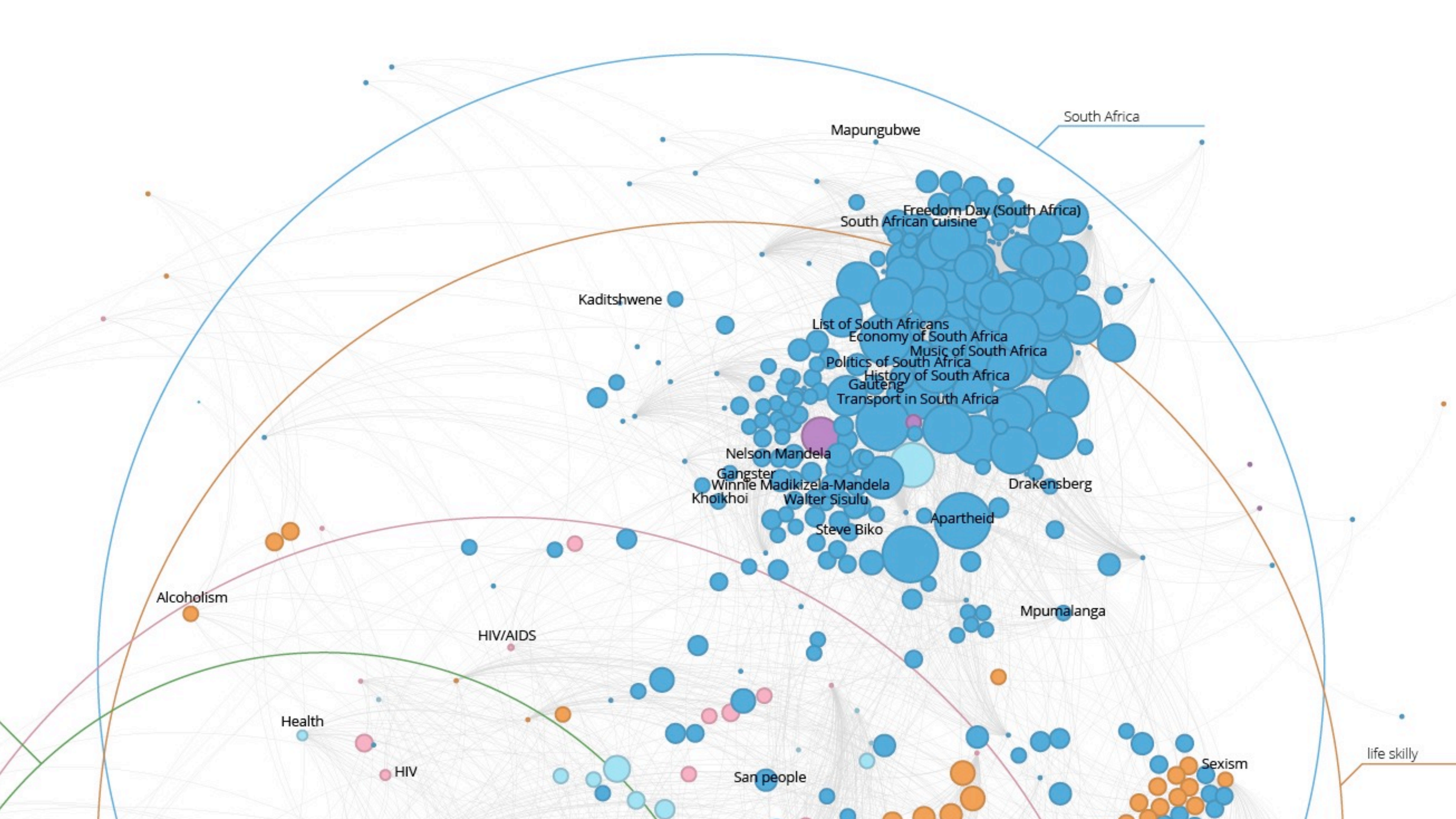
https://gephi.org/

Rete di articoli
agosto 2015

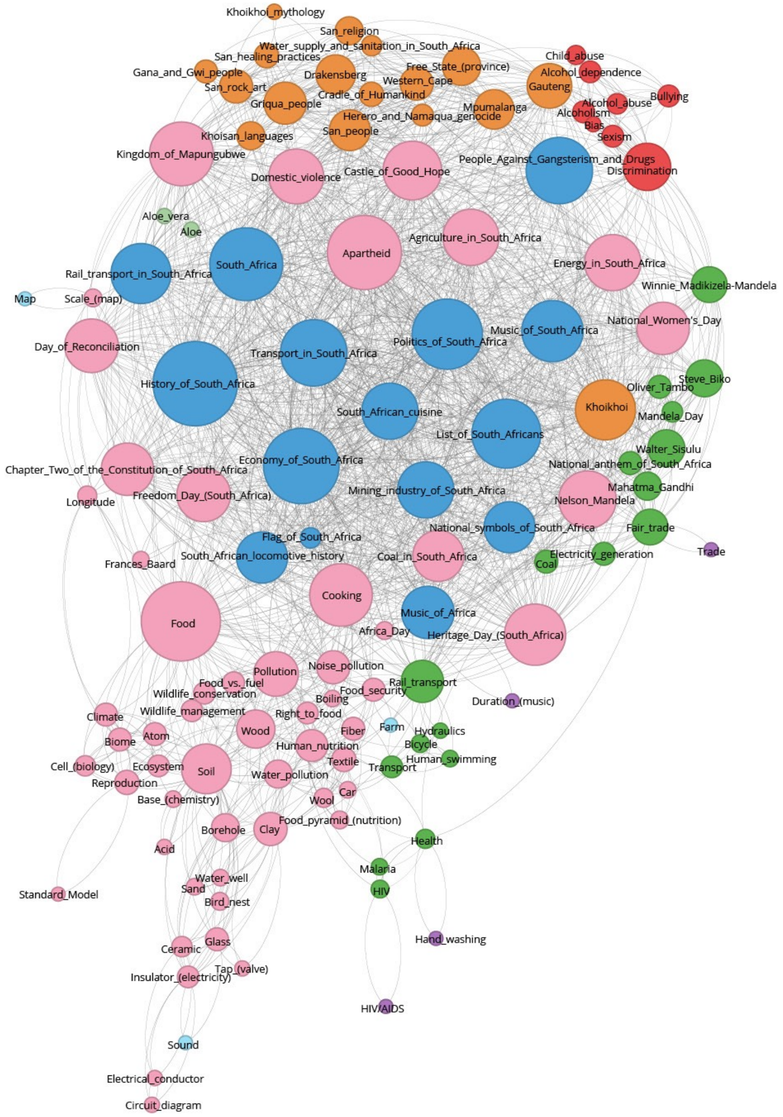


Rete di articoli
marzo 2016

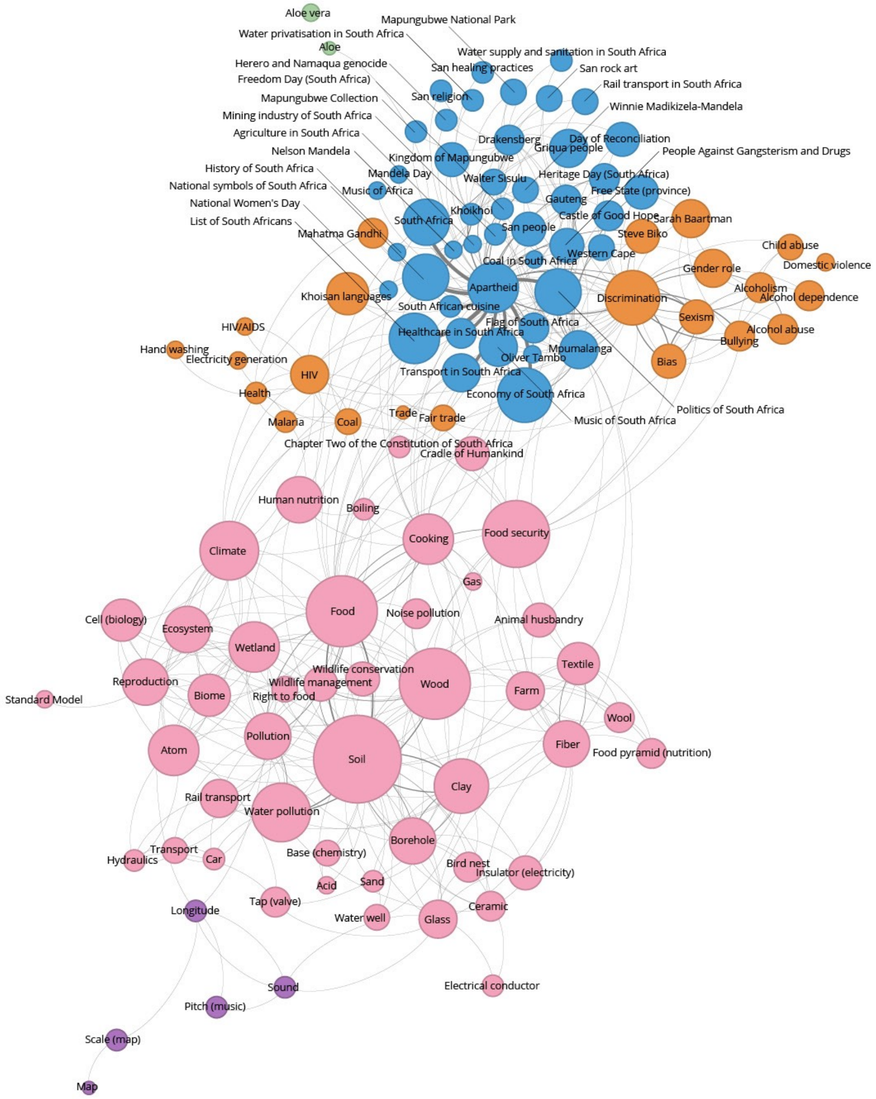



South Africa
Mapungubwe
Freedom Day (South Africa)
South African cuisine
Kaditshwene
List of South Africans
Economy of South Africa
Music of South Africa
Politics of South Africa
History of South Africa
Gauteng
Transport in South Africa
Nelson Mandela
Gangster
Winnie Madikizela-Mandela
Khoikhoi
Walter Sisulu
Drakensberg
Apartheid
Steve Biko
Mpumalanga
Alcoholism
HIV/AIDS
Health
HIV
San people
Sexism
life skilly

Rete di articoli



August 2015



August 2017

altri dati

Features degli articoli
ottenuti con jQuery

- issues
- references
- notes
- images
- see also

features.js

```
function get_images(url) {

    $.ajax({
        type: 'GET',
        url: my_proxy + wikilink + url, //my_proxy + url,
        processData: true,
    })
    .done (function (get_ref) {

        var parsedata_func = $.parseHTML(get_ref)

        get = [];
        get1 = [];
        get2 = [];

        get1 = $(parsedata_func).find('.thumbimage')
        get3 = ($(parsedata_func).find('.thumb').find('a').find('img'));

        sum = 0;

        var url_clean = url.replace('https://en.wikipedia.org/wiki/','').replace(

        container = $('#output')

        $.each( get1, function( i, val ) {
            txt = $(this).prop('href');
            if (txt !== undefined) {
                console.log(txt);
                sum++;
            }
            else{}
            //sum++;
        })

        jQuery.each( get3, function( i, val ) {
            if (val !== undefined) {
                sum++;
            }
        })

        container.append(url_clean + ',' + sum + '</br>')

    })
    .error (function (xhr, ajaxOptions, thrownError) {
        console.log(xhr.status);
        console.log(thrownError);
    });
}

function get_all_images() {
    container = $('#output')
    container.append('article,images</br>')
    jQuery.each( articles_a, function( i, val ) {  //  list; articles_a;
```

Wikipedia Primary School

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.

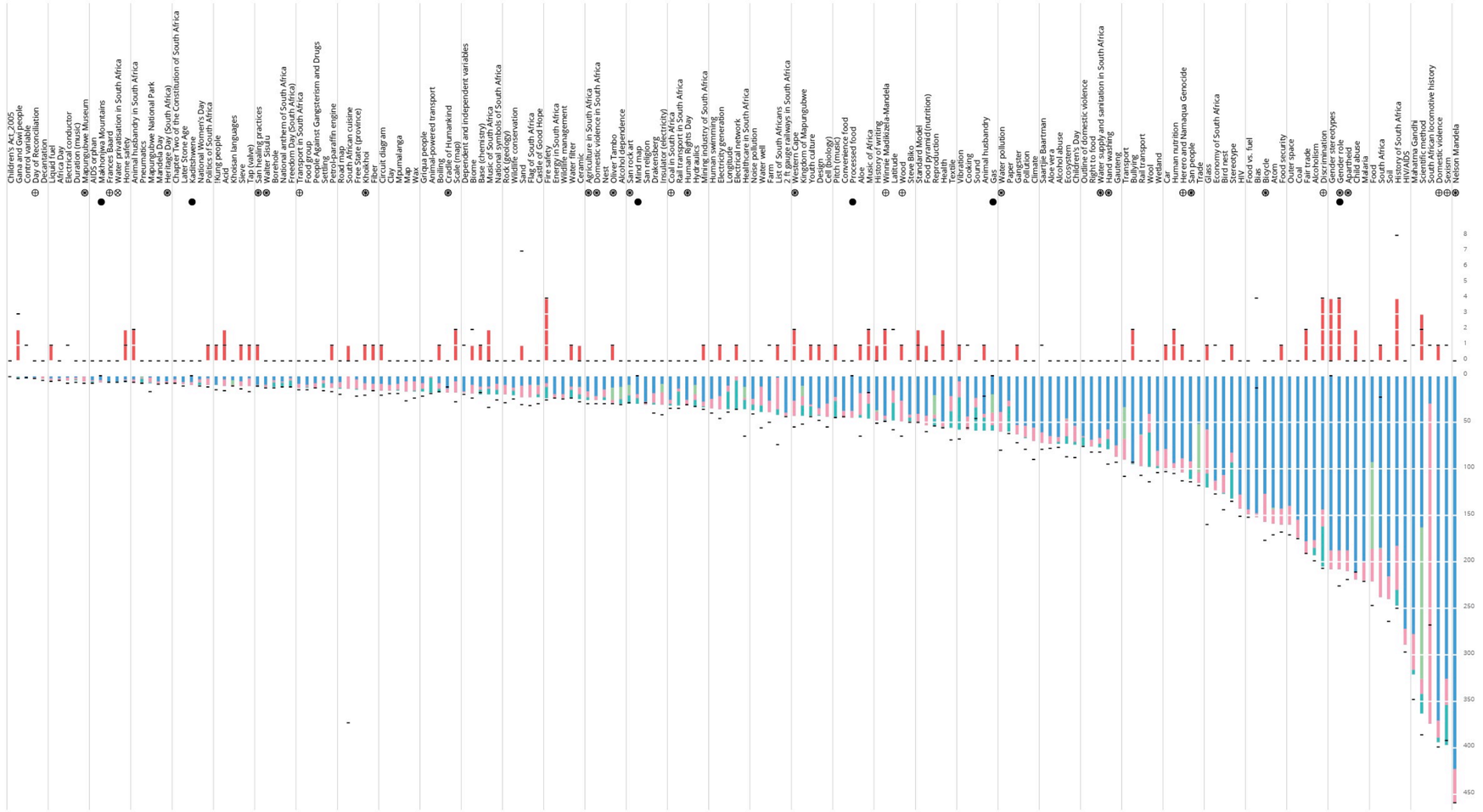
Please note: issues and the other features are shown at different scales.

SUPSI

Department for Environment Constructions and Design
Laboratory of visual culture

Project by

Giovanni Profeta
giovanni.profeta@supsi.ch



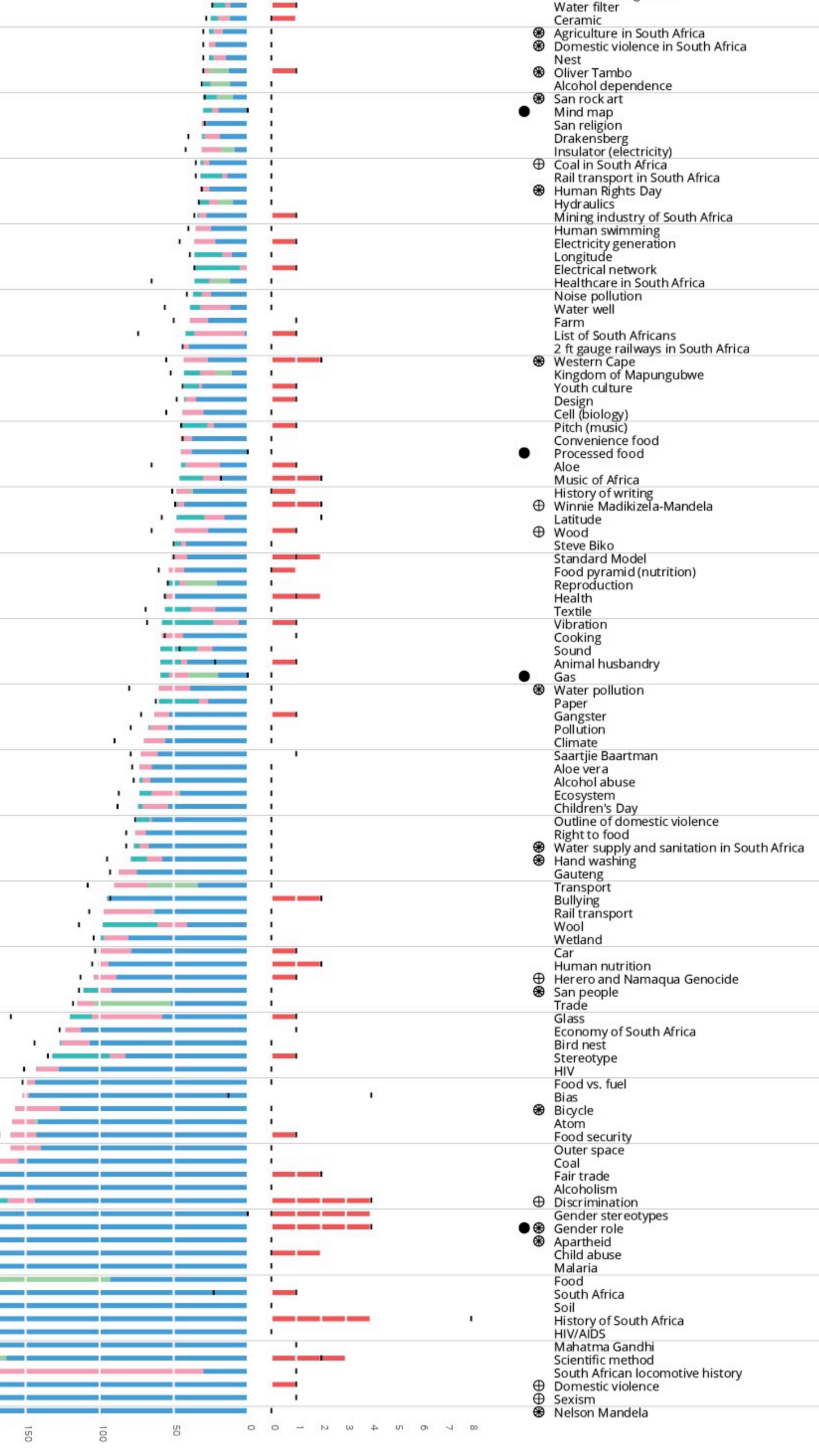
Timestamp: 05-03-2016; benchmark: 13-08-2015;

Data Source: Wikipedia (page scraping); Wikipedia API

Visualization tool: D3.js

This work is licensed under the Creative Commons Attribution-Share Alike 4.0 international license.

Marzo 2016


Water filter
Ceramic
Agriculture in South Africa
Domestic violence in South Africa
Nest
Oliver Tambo
Alcohol dependence
San rock art
Mind map
San religion
Drakensberg
Insulator (electricity)
Coal in South Africa
Rail transport in South Africa
Human Rights Day
Hydraulics
Mining industry of South Africa
Human swimming
Electricity generation
Longitude
Electrical network
Healthcare in South Africa
Noise pollution
Water well
Farm
List of South Africans
2 ft gauge railways in South Africa
Western Cape
Kingdom of Mapungubwe
Youth culture
Design
Cell (biology)
Pitch (music)
Convenience food
Processed food
Aloe
Music of Africa
History of writing
Winnie Madikizela-Mandela
Latitude
Wood
Steve Biko
Standard Model
Food pyramid (nutrition)
Reproduction
Health
Textile
Vibration
Cooking
Sound
Animal husbandry
Gas
Water pollution
Paper
Gangster
Pollution
Climate
Saartjie Baartman
Aloe vera
Alcohol abuse
Ecosystem
Children's Day
Outline of domestic violence
Right to food
Water supply and sanitation in South Africa
Hand washing
Gauteng
Transport
Bullying
Rail transport
Wool
Wetland
Car
Human nutrition
Herero and Namaqua Genocide
San people
Trade
Glass
Economy of South Africa
Bird nest
Stereotype
HIV
Food vs. fuel
Bias
Bicycle
Atom
Food security
Outer space
Coal
Fair trade
Alcoholism
Discrimination
Gender stereotypes
Gender role
Apartheid
Child abuse
Malaria
Food
South Africa
Soil
History of South Africa
HIV/AIDS
Mahatma Gandhi
Scientific method
South African locomotive history
Domestic violence
Sexism
Nelson Mandela
150
100
50
0
0
1
2
3
4
5
6
7
8

Wikipedia Primary School

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 picks, the page views are displayed as horizon charts. Thus it is possible to understan 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 reassesment
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 Schönwetter (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 founded by

FNSNF Fonds national suisse Schweizerischer Nationalfonds Fondo nazionale svizzero Swiss National Science Foundation

NRF National Research Foundation

Timespan: July 2015—June 2017

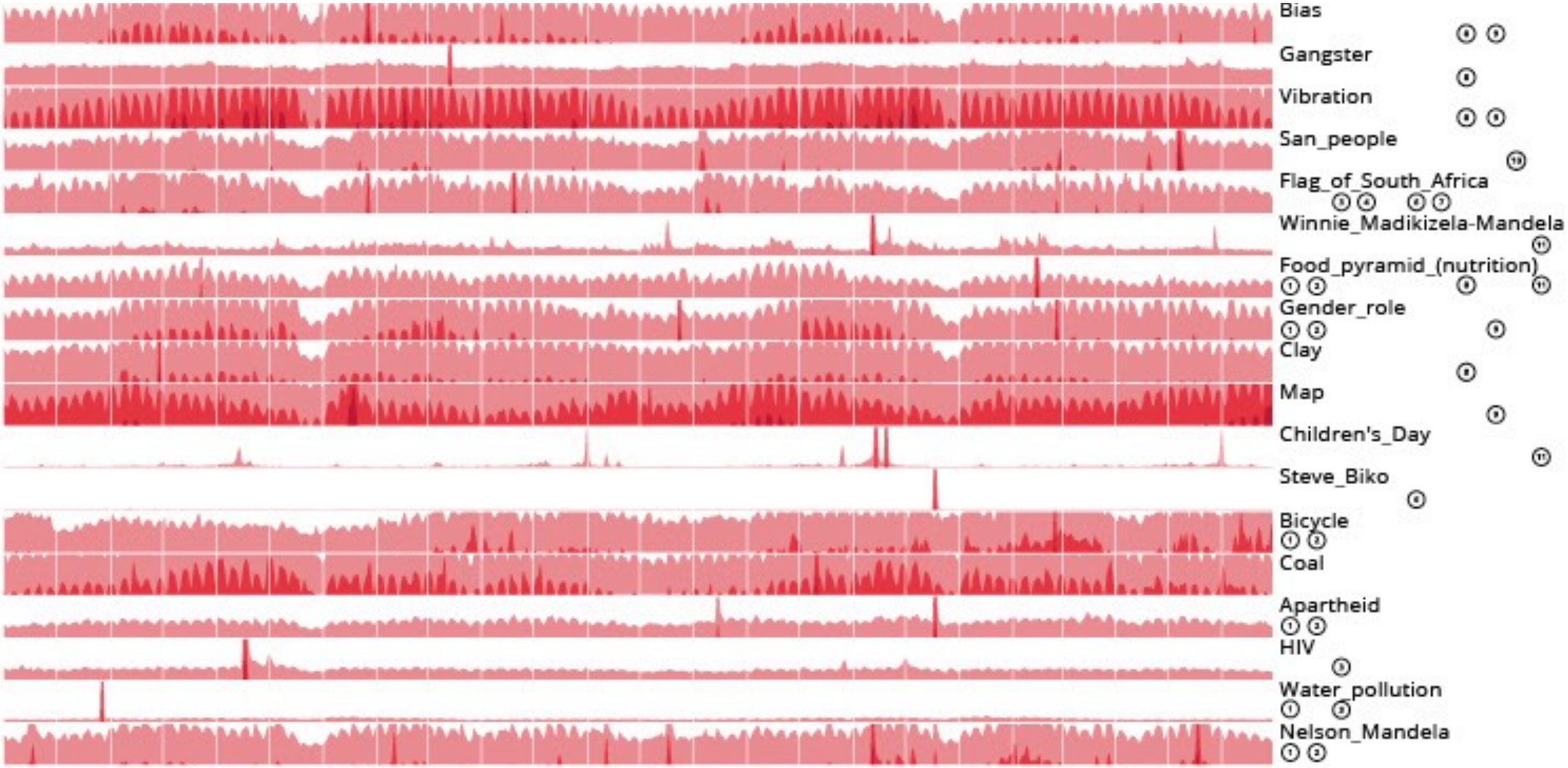
Data Source: Wikipedia API

Visualization tool: D3.js

This work is licensed under the Creative Commons Attribution-Share Alike 4.0 International license.

Dati del 2015 al 2017

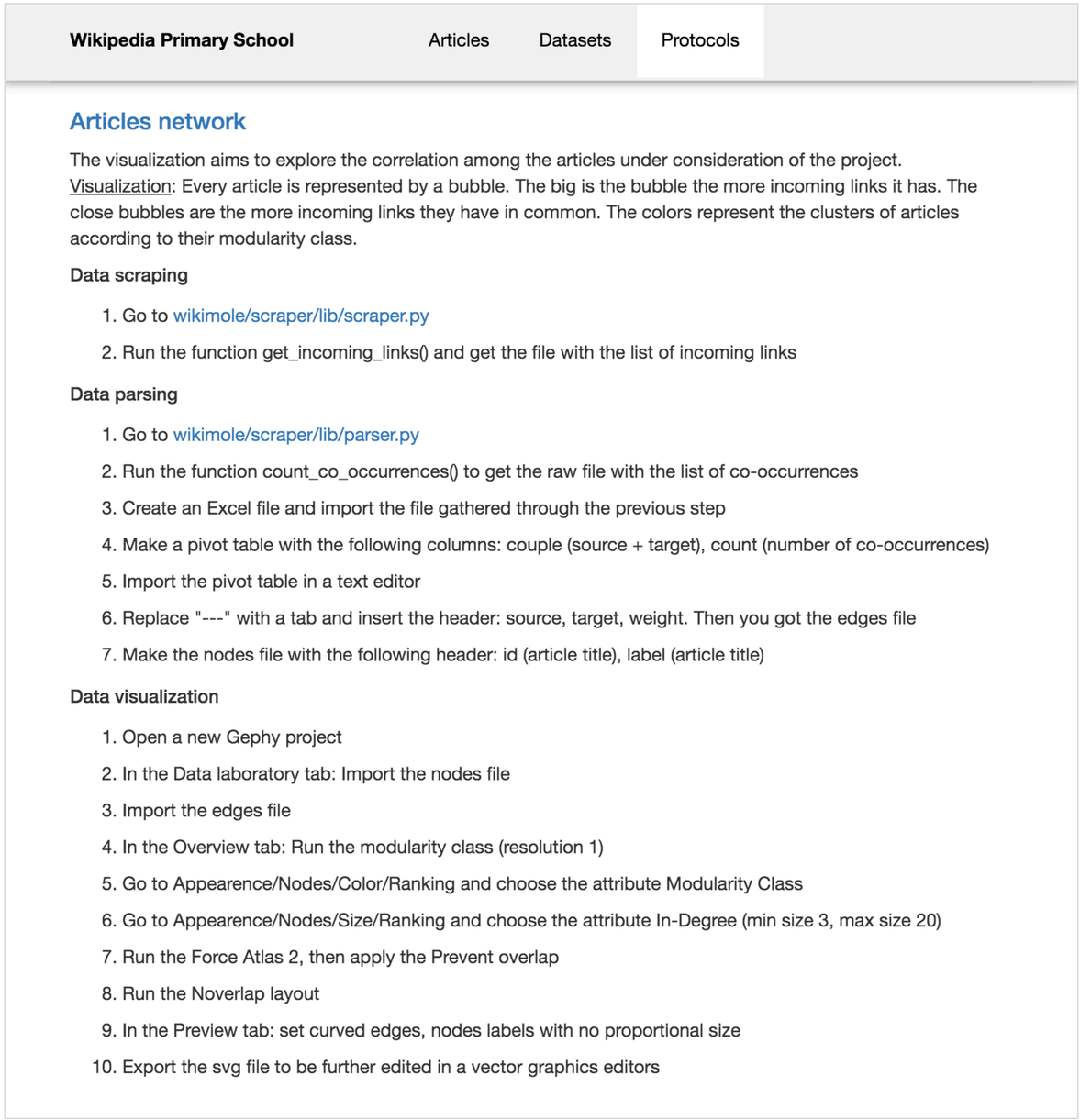
Pagine viste



Dati del 2015 al 2017

strumenti per proseguire le analisi

Protocolli di ottenimento e visualizzazione dati



Wikipedia Primary School Articles Datasets Protocols

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/scraper/lib/scraper.py
2. Run the function get_incoming_links() and get the file with the list of incoming links

Data parsing

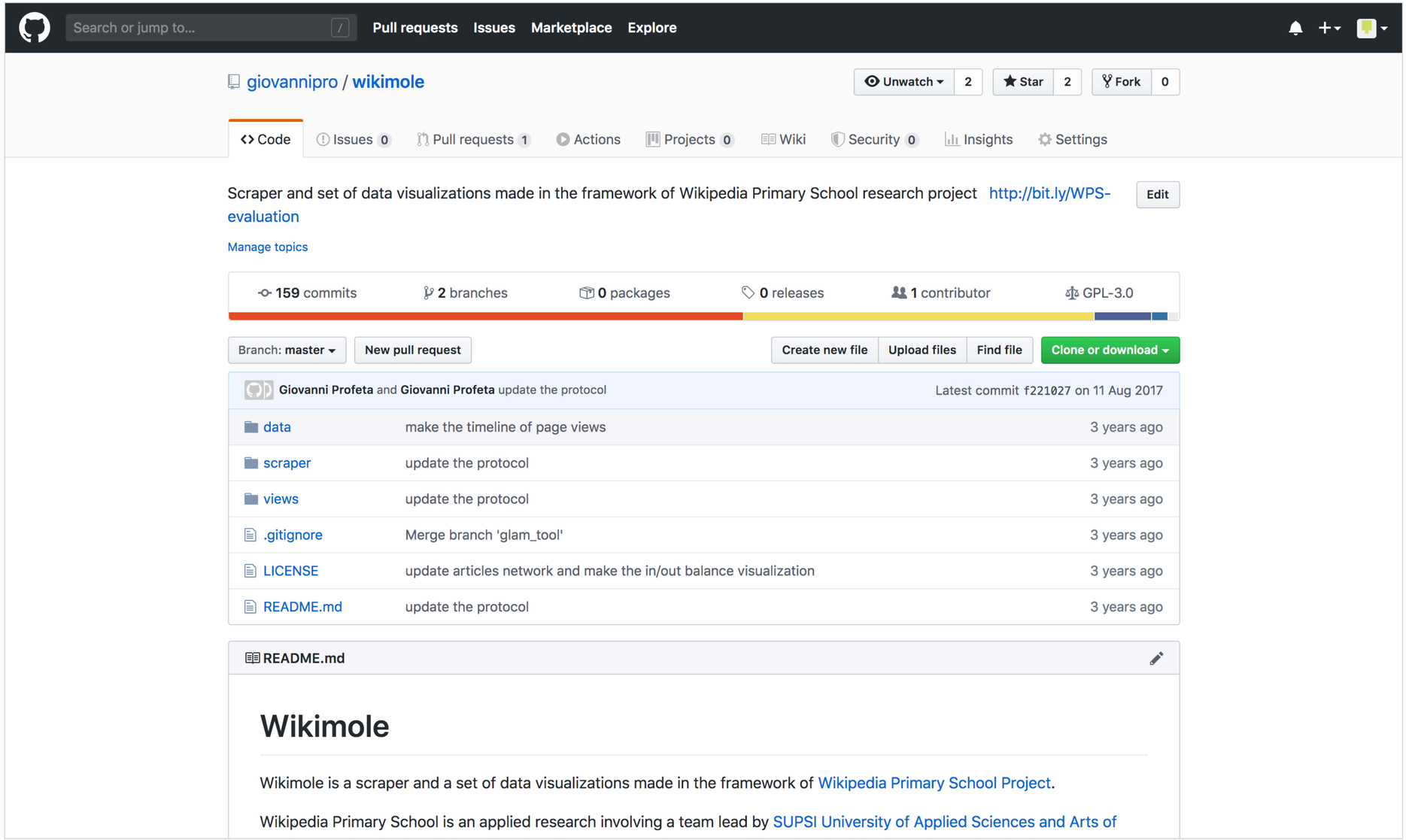
1. Go to wikimole/scraper/lib/parser.py
2. Run the function count_co_occurrences() to get the raw file with the list of co-occurrences
3. Create an Excel file and import the file gathered through the previous step
4. Make a pivot table with the following columns: couple (source + target), count (number 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 Appearence/Nodes/Color/Ranking and choose the attribute Modularity Class
6. Go to Appearence/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

https://giovannipro.github.io/wikipedia-primary-school/protocols.html

Wikimole scraper



https://github.com/giovannipro/wikimole

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

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