

Dall'opendata alla physicalization

Andrea Trentini

Aprile 2020

Arcobaleno *Cittadinanza Digitale e Tecnocivismo*

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

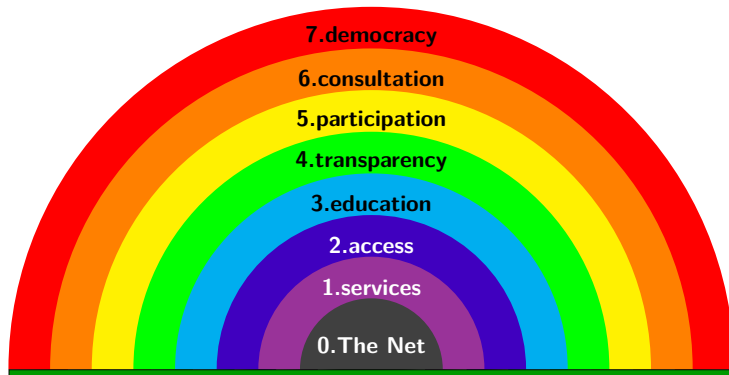
Barra prestiti

Lampada

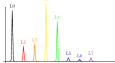
Dekatron

Physiradio

References



Toccheremo 'trasparenza', 'education' e 'participation' e faremo ponte con SciVis e DataPhys.

Spettrogrammi: 

Trasparenza e CDT

- ▶ stato interno (nel senso MdT) di un'organizzazione o della P.A.
- ▶ *civic accountability*
- ▶ controllo “sociale” (sulle spese ad esempio)
- ▶ verifica policy (*ex-ante* e *ex-post*)
- ▶ approccio critico al governo (in senso lato)
- ▶ aspetti tecnici ma anche squisitamente politici
- ▶ ponte con crowdsourcing e intelligenza collettiva

Flusso *top-down*.

Chi c'è all'altro estremo del flusso?
(inizia per “c”)

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

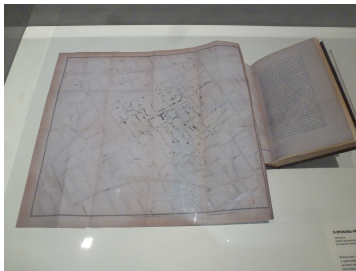
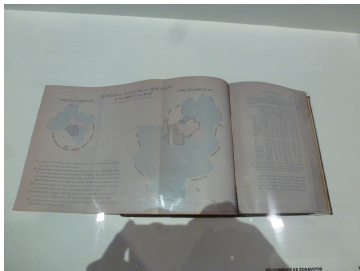
Dekatron

Physiradio

References

Intelligenza collettiva (crowdsourcing)

Nightingale e Snow



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Trasparenza concreta

Chi:

- ▶ amministrazioni ed enti per servizi di pubblica utilità
- ▶ pubblici, partecipati, no profit
- ▶ privati (?), tema IP

Cosa (rendere pubbliche le informazioni “interne”):

- ▶ informazioni di “stato” relative ai processi
- ▶ informazioni su cui si basano le decisioni (documentazione istruttoria)

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Oggetti della trasparenza, “ontologia”

- ▶ audio/video di sedute pubbliche: *live stream*, *podcast*, archivio
- ▶ delibere (di giunta, di consiglio): OdG, mozioni, votazioni ecc.
- ▶ documentazione istruttoria
- ▶ dati sulla partecipazione degli eletti alle sedute degli organi (<http://openpolis.it>)
- ▶ bilancio (“semanticamente” leggibile)
- ▶ bandi di gara, capitolati e appalti (cfr. albo pretorio)
- ▶ contratti, affidamenti, determine amministrative ecc.
- ▶ leggi (!!!): 241/1990, FOIA, CAD
- ▶ ...

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Sunlight...

“Publicity is justly commended as a remedy for social and industrial diseases.

Sunlight is said to be the best of disinfectants, electric light the most efficient policeman.”

[Louis Brandeis (US Supreme Court, 1916)]

Teoreticamente possibile?

C'è sempre la c.d. “politica di corridoio” che non è necessariamente negativa, ma che in ogni caso non è controllabile.

(esempio dello “stato coerente di un oggetto”)

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Opendata? I

“A piece of data or content is open if anyone is free to use, reuse, and redistribute it - subject only, at most, to the requirement to attribute and/or share-alike.”

*[The Open Knowledge Foundation, [8]]
<http://opendefinition.org>*

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

In [6] si aggiunge:

- ▶ Open Government Data: “Public Sector Information that has been made available to the public as open data”
- ▶ Public Sector Information (PSI): “data and information produced, collected or held by public authorities, as part of their public task”, data that should be accessible (via web) at marginal cost and without discrimination, available in a digital and machine-readable format, and provided free of restrictions on use or redistribution

Opendata? III

Cioè:

- ▶ dati che possono essere usati “liberamente” (in termini di licenza)
- ▶ magari raggiungibili “facilmente” (web)
- ▶ magari elaborabili facilmente

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Classificazioni

- ▶ Berners-Lee [3]: formats, technical
- ▶ Davies [5]: pragmatics
- ▶ Berners-Lee and others [2, 4] 'linked open data'

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

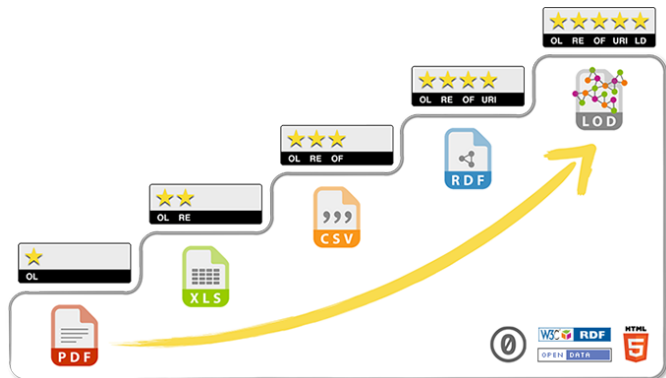
Lampada

Dekatron

Physiradio

References

Tim Berners-Lee



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenzal

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Altri aspetti

- ▶ licenze (nel dettaglio)
- ▶ aggiornamenti
- ▶ granularità del dato
- ▶ utilità
- ▶ commentabilità
- ▶ possibilità di chiedere/prioritizzare l'apertura

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

“Publishing open data is not just about technology. Tim Berners-Lee’s five stars of linked open data set out a series of approaches that open data initiatives [...] can take to publish data on the web. The following five star model seeks to add to this, highlighting key steps that open data initiatives can take to engage with data users. Each star includes a set of questions to unpack what might be involved in taking that step towards engagement.”

[http://www.timdavies.org.uk/2012/01/21/
5-stars-of-open-data-engagement](http://www.timdavies.org.uk/2012/01/21/5-stars-of-open-data-engagement)

<http://www.opendataimpacts.net/engagement>

1. Be demand driven (utile)
2. Provide context (metadati)
3. Support conversation (“interagibile”, commenti, feedback)
4. Build capacity & skills (strumenti)
5. Collaborate with the community (ponte L5, crowd)

Adoption:

- ▶ *de facto* (based on use)
- ▶ *de iure* (IETF, ECMA, RFC, ...)
“well” documented: es. Office XML: 7200 pagine...

Freedom:

- ▶ liberi (licenze!!!)
- ▶ proprietari/vincolati (brevetti?!?)

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

*Di più avrei potuto fare, specialmente nel campo statistico, se non ci fosse nel nostro paese, e nei privati e negli enti morali, una tal quale **ritrosia a confidare al dominio del pubblico dati, fatti e notizie.***

[Cesare Battisti, Il Trentino, 1898]

[http://it.wikisource.org/wiki/Pagina:](http://it.wikisource.org/wiki/Pagina:Il_Trentino.djvu/10)

[Il_Trentino.djvu/10](http://it.wikisource.org/wiki/Pagina:Il_Trentino.djvu/10)

Resistenza! II

Se pubbliche, queste informazioni, permettono di conoscere:

- ▶ come viene amministrata la cosa pubblica
- ▶ come vengono utilizzate le risorse pubbliche economiche, culturali, ambientali e umane

La trasparenza è alla base della *civic accountability*: partecipazione attiva dei cittadini alla valutazione della efficienza dell'amministrazione pubblica.

- ▶ l'ente che possiede il dato è geloso della propria conoscenza
- ▶ ci fanno i conti in tasca...
- ▶ ma ci sono norme di legge!
- ▶ allora facciamo operazioni di facciata...

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Webstacles

- ▶ *portmanteau* di “web obstacles”
- ▶ barriere artificiali al raggiungimento di un dato
- ▶ formati “strani” (proprietary e/o maldocumentati)
- ▶ CAPTCHA
- ▶ limitazioni “QoS” (Quality of Service, in senso deteriore)
- ▶ “graficalizzazione” (testo convertito in bitmap)

Es. ARPA Lombardia (nel seguito)

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Scraping/extraction

- ▶ *scraping*: letteralmente “grattare”
- ▶ browser e conoscenza di HTTP/HTML
- ▶ a volte basta leggere l'URL
(es. `http://weather.uwyo.edu/cgi-bin/uamap?REGION=naconf&OUTPUT=gif&TYPE=obs&TYPE=an&LEVEL=500&date=2020-04-02&hour=0`)
- ▶ programmi wget/curl-like
- ▶ browser scripting (greasemonkey, iMacros ecc.)
- ▶ headless browsing (`http://phantomjs.org`)
- ▶ librerie di scraping (Scrapy, HtmlUnit)
- ▶ voglia di sperimentare e un po' di pazienza

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Son tutti... [1]



Enrique Balbontin

http://it.wikipedia.org/wiki/Enrique_Balbontin

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Son tutti “open” con il sourcedata degli altri



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Dati puntuali giorno per giorno (e settimanali)



N.B. da allora il sito è stato “rinnovato” più volte

Dall'opendata alla physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi della scienza

Data physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Dai! Fallo!

Qualità dell'Aria
Regione Lombardia

SITO DE QUALITÀ DELL'ARIA di ARPA Lombardia
Nome: Aria - Richiesta dati

Archivio dei dati storici

Dati storici e real-time

Provincia:
Città:
Centralina:
Inquinante:
Sensori:

Sezione:

Da: Al:

E-mail:

Richiedi

Zona del territorio / Provincia

Città:
Centralina:

VAI

PM10 la media delle concentrazioni negli agglomerati urbani

5 dati storici

- L'evoluzione negli anni
- Il trend del PM10
- Il confronto con i limiti della normativa
- Download archivio storico

Oltre la rete

- Impianto per la misura dell'ozono in Italia ed in Europa
- PM10 a confronto
- Il sistema "Qualità dell'Aria della Regione"

Meteo

- L'importanza delle previsioni meteorologiche
- Piano di inquinamento preventivo per i prossimi giorni

Documentazione

- Relazioni annuali provinciali
- Campagne mesh mobile
- La valutazione modellistica

~150 staz
X
~5 item
X
~30 anni
=====
>20k query (mail)

Dall'opendata alla physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi della scienza

Data physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Soluzione?

- ▶ PASSATO: nessuna!
Salvo avere soldi (captcha human solving)
- ▶ FUTURO (che comincia dal 2005!):
scraping + scripting = scrapy
bash and filters (I'm an old dog, sorry), oggi si userebbe
'scrapy' o altro

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

ScrApting?!? I

Con:

```
wget
```

```
-q -O $DATA/$(date +%Y%m%d%H%M)_aria$num.html  
http://ita.arpalombardia.it/ITA/qaria/lista$num.asp  
2>&1
```

Nel *crontab*:

```
0 1 * * * /root/aria/aria.sh
```

Si ottiene il *raw data*:

```
201402110100_aria01.html.gz  
201402110100_aria02.html.gz  
201402110100_aria03.html.gz  
201402110100_aria04.html.gz  
201402110100_aria05.html.gz  
201402110100_aria06.html.gz  
201402110100_aria07.html.gz  
201402110100_aria08.html.gz  
201402110100_aria09.html.gz  
201402110100_aria10.html.gz
```

[Dall'opendata alla
physicalization](#)

Andrea Trentini

[Perché](#)

[Trasparenza](#)

[Opendata](#)

[Resistenza!](#)

[Scraping](#)

Aria

Car sharing

Bike sharing

[Chi raccoglie ma
non concede...](#)

(a latere) WiFi analysis

[OD websites](#)

[Accademia e crisi
della scienza](#)

[Data
physicalization](#)

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

[References](#)

ScrApting?!? II

Il *raw data* contiene (sorgente):

```
...class="linktesto"><B>Cormano</B><br><font
color="#666666"></font></a></td> <td width="1"
bgcolor="#003366"></td> <td width="54" align="right"
class="testobox">nd</td> <td width="18" valign="middle"
align="center"></td> ... <td width="1"
bgcolor="#003366"></td> ... <td width="1"
bgcolor="#003366"></td> <td width="54" align="right"
class="testobox">86</td> <td width="18" valign="middle"
align="center"></td>...
```

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

ScrApting?!? III

Col filtro:

```
... | sed 's/<br>/ /g' | vilistextum -a - - |
    tr -d '\177-\377' | sed 's/</(min.di)/g' |
    tee tmp.txt | tr -s "\n" " " |
    sed -f sed.script |
    grep -ivf ../righePost.txt |
    tr -s " " "; " | sort > $FILE.dat
```

Si ottiene il *CSV data*:

Milano_Abbiategrasso.csv

Milano_Juvara.csv

Milano_Liguria.csv

Milano_Marche.csv

Milano_Messina.csv

...

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Phisiradio

References

ScrApting?!? IV

Che contiene i campi:

Data,S02,PM10,PM2.5,N02,C0,03,Benzene

2005/12/5, , ,143,3.2, ,6.8,

2005/12/6, , ,196,4.6, ,9.0,

2005/12/7, , ,135,4.5, ,8.2,

2005/12/8, , ,165,5.4, ,10.6,

2005/12/9, , ,190,4.5, ,8.0,

2005/12/10, , ,150,4.7, ,7.1,

2005/12/11, , ,135,3.8, ,6.3,

2005/12/12, , ,155,3.5, ,6.2,

2005/12/13, , ,103,3.0, ,5.4,

...

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Conclusioni

- ▶ verifica storica dei dati! (gira dal 2005)
- ▶ grafici generati in automatico (gnuplot) con i dati aggiornati
- ▶ risultati ottenuti (anche politici):
 - ▶ effetti di $\text{AreaC} = 0!$
 - ▶ in corso di studio effetto AreaB e COVID sul'aria (controintuitivo)
 - ▶ spingere ARPA a pubblicare (oltre che a cercare di "scappare", vedi seguito)
 - ▶ ARPA pubblica tutto a fine 2017 (un pochino di merito me lo prendo)

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Gli scARPAti di casa...

- ▶ durante una mia presentazione a IODD14 sul lavoro di *scraping*...
N.B. contatto con Crespi (opendata Lombardia), ha in seguito fatto da ponte con ARPA
- ▶ presenti dirigenti ARPA...
- ▶ sconvolti nello scoprire che c'era qualcuno che in effetti quei dati li usava e li immagazzinava!

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

ScARPA scARPA galantuomo! (Manzoni)

Notare il buco nei dati... coincidenze? Io non credo!

<insert your favourite Adam Kadmon meme>

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

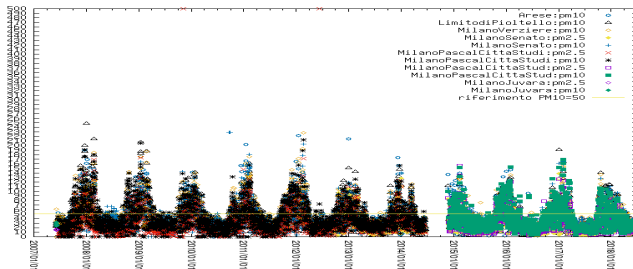
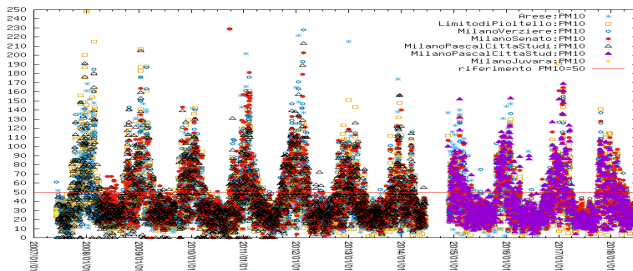
Barra prestiti

Lampada

Dekatron

Physiradio

References



Carsharing I

Trova gli URL...

- ▶ <http://www.sharengo.it/core/publiccars>
- ▶ http://www.car2go.com/api/v2.1/vehicles?oauth_consumer_key=car2goweb&format=json&loc=milano
- ▶ http://enjoy.eni.com/ajax/retrieve_vehicles
(a tratti da err 200)

Attiva un *crontab* (ogni minuto parte lo *scraper*)

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Carsharing II

Ottieni il dato (JSON):

```
[{"car_name":"Fiat 500",  
"fuel_level":43,"secondary_fuel_level":"","  
"lat":45.46435,"lon":9.188686,  
"address":"Via,1 00001 Roma",  
"virtual_rental_type_id":2,"virtual_rental_id":583904,  
"car_category_type_id":1,"car_category_id":8,  
"onClick_disabled":false,"carModelData":[100,141]},  
{"car_name":"Fiat 500",  
"fuel_level":53,"secondary_fuel_level":"","  
"lat":45.520947,"lon":9.227421,  
"address":"Viale Monza, 316, 20128 Milano MI",  
"virtual_rental_type_id":2,"virtual_rental_id":681541,  
"car_category_type_id":1,"car_category_id":8,  
"onClick_disabled":false,"carModelData":[100,141]},...
```

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

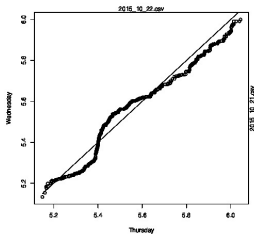
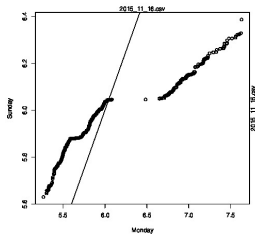
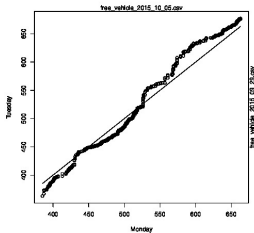
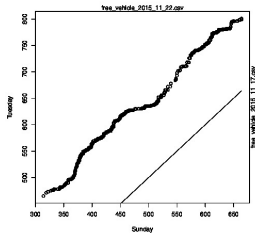
Physiradio

References

Elabora e plotta... I

Dall'opendata alla
physicalization

Andrea Trentini



Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

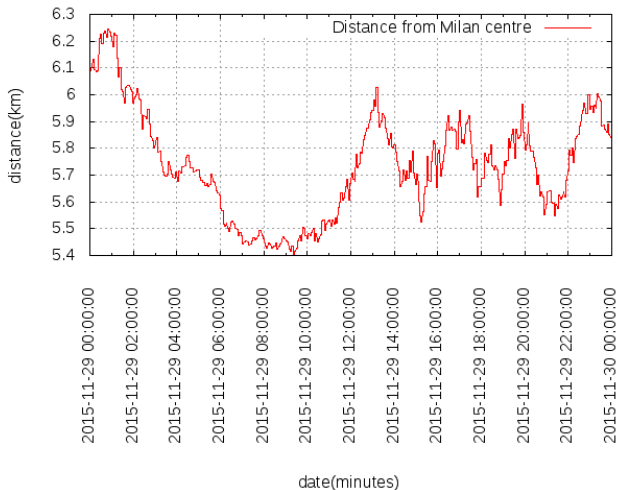
Lampada

Dekatron

Physiradio

References

Elabora e plotta... II



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

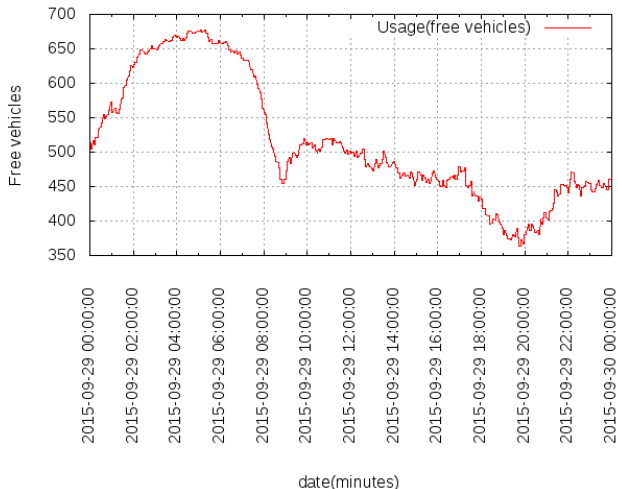
Lampada

Dekatron

Physiradio

References

Elabora e plotta... III



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

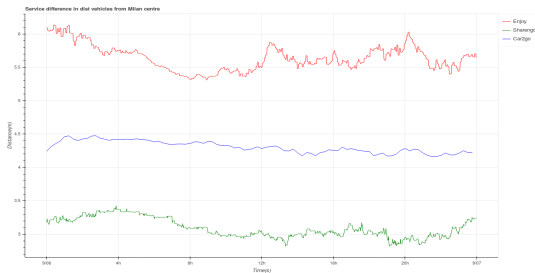
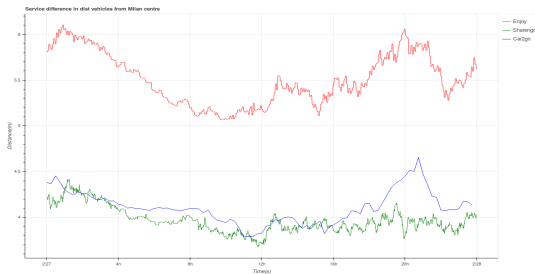
Lampada

Dekatron

Physiradio

References

Evoluzione storica delle distanze dal centro



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

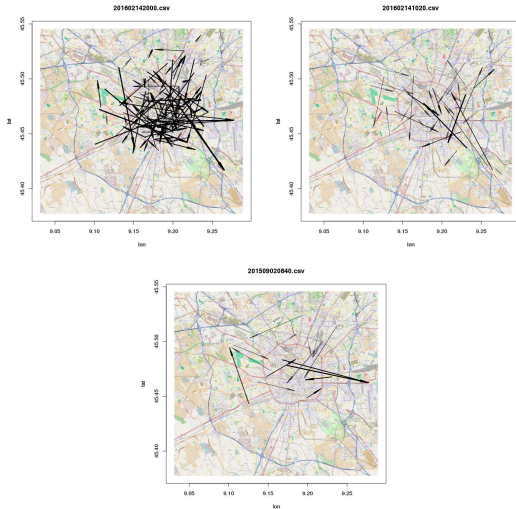
Lampada

Dekatron

Physiradio

References

Tratte



Video: <http://www.youtube.com/watch?v=60QK7Go2C1U>

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Per ora solo data scraping, pulizia dei dati e pubblicazione su zenodo... (partita tesi da poco)

Solito 'wget':

```
wget -q http://www.bikemi.com/it/mappa-stazioni.aspx
```

CSV risultanti:

```
date (YYMMDDHHMM):LAT:LON:location name:
```

```
  bikes:electrical:free slots (0 = station full)
```

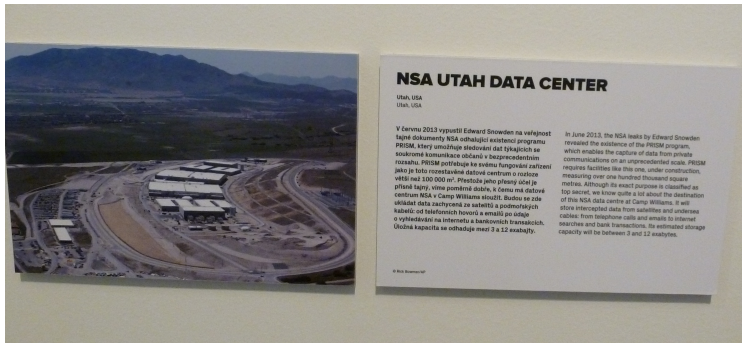
```
1512210720:45.464683238626:9.18879747390747:  
1-Duomo:0:0:23
```

```
1512210720:45.4663612119451:9.19765949249268:  
2-San Babila:0:2:22
```

...

Chi raccoglie ma non concede...

NSA (Noi Sappiamo Abbastanza)



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

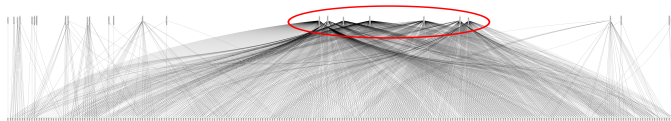
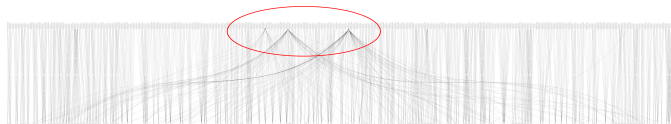
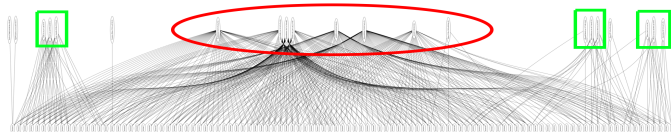
Lampada

Dekatron

Physiradio

References

WiFi analysis



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Siti web 'opendata' (finalmente!)

- ▶ Comuni (es. <http://dati.comune.milano.it>),
Regioni (es. <http://dati.lombardia.it>),
Nazioni (es. <http://dati.gov.it>)
- ▶ enti vari (es. <http://www.wikidata.org>)
- ▶ ufficiali (esercizio: trovare i siti dei progetti
"soldipubblici", "italiasicura", "opencantieri",
"opendemanio")
- ▶ crowdsourced

E meta-siti:

- ▶ <http://www.sciamlab.com/opendatahub/dataset>
- ▶ <http://opendatabarometer.org>
- ▶ <http://index.okfn.org>
- ▶ <http://opendatamonitor.eu>
- ▶ <http://www.opendata500.com>

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

API e miele...

- ▶ c'è anche il lato buono
- ▶ Application Programming Interface
- ▶ REST
 - `http://en.wikipedia.org/wiki/Representational_state_transfer`
 - `http://www.restapitutorial.com`
- ▶ XML/JSON
- ▶ es. governo inglese: `http://data.gov.uk/data/api`
(ma anche ormai quasi tutti i siti opendata)
- ▶ CKAN (CMS orientato all'opendata)
- ▶ "APIficatori":
 - ▶ `http://apify.com`
 - ▶ `http://www.scraperaapi.com`
 - ▶ `http://rapidapi.com/collection/top-scraperaapi-tools`

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

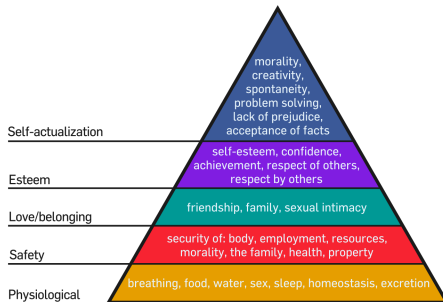
Dekatron

Physiradio

References

Che ce ne facciamo?

- ▶ *civic accountability* (già detta)
- ▶ servizi “terze parti”!!! (anche *profit*)
- ▶ “cultura” personale e utilità sociale (Maslow)



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Servizi terze parti... I

- ▶ Mapnificent (pub transport)
- ▶ Commutefrom (travel)
- ▶ Mapumental (real estate)
- ▶ Illustreets (qualità vita)
- ▶ Quolimi
- ▶ Are you safe? Freeda.it
- ▶ waqi.info
- ▶ openaq.org
- ▶ incidenti stradali vs. autovelox (BG)
- ▶ visionzerosf.org

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Servizi terze parti... II

- ▶ Bikedistrict
- ▶ CityMapper (app)
- ▶ Trains Timetable Italy (app italiana)
- ▶ Flood alerts
- ▶ Don't eat at (risultati ispezioni igieniche!)
- ▶ Noise watch
- ▶ Smart city dashboard (es. London)
- ▶ Public toilet map (crowd, stile tripadvisor)
- ▶ Fontanelle Milano (un dataset che sembrava inutile...)
- ▶ Openpolis!!! e Spaghetti Opendata

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Accademia e crisi della scienza

- ▶ riproducibilità
- ▶ verificabilità
- ▶ falsificabilità
- ▶ approccio critico
- ▶ trasparenza dei metodi/processi
- ▶ <http://retractionwatch.com>
- ▶ ...

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

F.A.I.R. data

- ▶ Findable
- ▶ Accessible
- ▶ Interoperable
- ▶ Reusable

Importante il meccanismo del DOI (<http://www.doi.org>): un URI assegnato stabilmente ad un dataset (cfr. <http://zenodo.org>) per poterlo citare in un *paper*.

http://en.wikipedia.org/wiki/FAIR_data

[http:](http://www.force11.org/group/fairgroup/fairprinciples)

[//www.force11.org/group/fairgroup/fairprinciples](http://www.force11.org/group/fairgroup/fairprinciples)

<http://www.nature.com/articles/sdata201618>

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Il problema dei dati

- ▶ Power...
any citizen, **with enough knowledge**, can retrieve data (that can be better trusted if coming from third party sources) from public servers and study the effect of laws such as a change of tobacco taxation on number of smokers, the banning of some types of vehicles on air quality [11], the introduction of specific legislation to lower the number of unoccupied people, etc...
- ▶ ...needs knowledge
data analysis is not for the masses [9]: the average citizen either does not know the existence of open data, or he/she is not able to extract information from a dataset.

Serve “democratizzazione”

- ▶ “DIY, hacking, and craft” are powerful democratizing tools [10]
- ▶ IoT (Internet of Things) devices: “the potential to democratize the access, use, and appropriation of data”, since “most of the data is ‘black box’ in nature: users often do not know how to access or interpret data”. These devices, “**blended** into homes”, can be used to **engage non-technical users**. [7]
- ▶ The democratization aspect is also proposed in [12] where ‘Domestic Widgets’ are used to successfully support household creativity and co-creation of data representations.
- ▶ E, *of course*, **visualizzazione!**

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Data physicalization

<http://dataphys.org> e in particolare

<http://dataphys.org/list>

- ▶ “recente” (in realtà parte da 5500 B.C., ma oggi informatica!)
- ▶ rendere **tangibile** il dato
- ▶ dinamicamente?
- ▶ interattivamente?

Installazioni artistiche e/o interattive:

- ▶ fontana clock
<http://www.youtube.com/watch?v=gusJeslMbLc>
- ▶ Marco Brianza <http://www.marcobrianza.it>
- ▶ BLOOP
<http://www.instagram.com/blopfestival>

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

3D surface dinamica

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

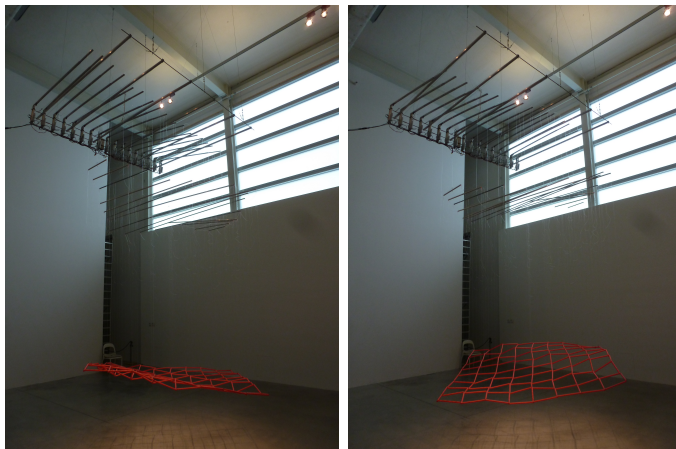
Barra prestiti

Lampada

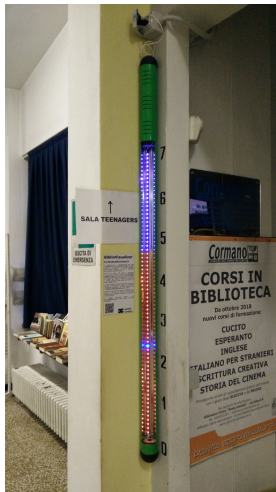
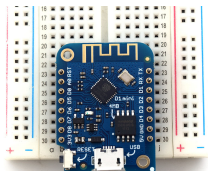
Dekatron

Physiradio

References



Barra prestiti (HackLabCormano)



http:

[//gitlab.com/HackLabCormano/BiblioVisualizer](https://gitlab.com/HackLabCormano/BiblioVisualizer)

[CC BY-SA 4.0] 2020 © Andrea Trentini - Dipartimento di Informatica - Università degli Studi di Milano

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Data gathering

“Cruscotto” CSBNO <http://webopac.csbno.net>:



(dati di tutto il CSBNO, cumulativi, aggiornati ogni ora)

Fonte dati:

frassino.comperio.it/kpi/clavis-csbno.json

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Cosa visualizzare?

```
{"dbName":"clavis-csbno", "utenti":"334518",  
"prestiti":"159497", "prestiti_oggi":"1520",  
"utentiplus":"7157", "utenti_attivi":"88484"}
```

- ▶ Ci vuole un dato un po' variabile...
- ▶ Di cui calcolare/misurare un range di valori
- ▶ Scelto "prestiti_oggi", variabile da 0 (a inizio giornata) a circa 5000 (a fine giornata) con punte di 7000 (cfr. <http://atrent.it/CSBNO/CSBNO.log>)

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Reazioni e follow-up

- ▶ Sindaco ha presentato la barra durante la festa annuale della biblioteca
- ▶ “Grimaldello” verso contapersone, collaborazione con biblio
- ▶ Esperimenti di connessione fra device: mostra ingressi e uscite lanciando un “LED sweep”
- ▶ CSBNO inizialmente entusiasta, aveva promesso liberazione dati... ora forse torna alla carica

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

KLamp - grafico a torta?



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

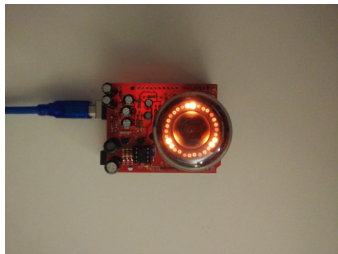
Dekatron

Physiradio

References

Dekatron - altra torta?

O *gauge*, o effetto dinamico in tempo reale...



<http://en.wikipedia.org/wiki/Dekatron>

<https://www.youtube.com/watch?v=iF8CnG7ee-A>

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Clockatron - NTP synced



<http://hackaday.io/project/168875-clockatron>

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

Physiradio - musicalizzazione

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References



(paper allegato)

Blending in



Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

References I



E. Balbontin and A. Macchiavello.

Son tutti finocchi col culo degli altri. Corso avanzato di lingua savonese.

Oscar bestsellers. Mondadori, 2005.



Tim Berners-Lee.

Government linked data.

<http://www.w3.org/DesignIssues/GovData.html>, 2009.



Tim Berners-Lee.

Linked data.

<http://www.w3.org/DesignIssues/LinkedData.html>,
June 2009.

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti


Lampada


Dekatron


Physiradio

References

References II

 Christian Bizer, Tom Heath, and Tim Berners-Lee.
Linked data-the story so far.
International journal on semantic web and information systems, 5(3):1–22, 2009.

 Tim Davies and Practical Participation.
Open data, democracy and public sector reform: A look at open government data use from data.gov.uk.
Practical Participation, 2010.

 HM Government.
Open data white paper - unleashing the potential.
Technical report, 2012.

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

References III

 Steven Houben, Connie Golsteijn, Sarah Gallacher, Rose Johnson, Saskia Bakker, Nicolai Marquardt, Licia Capra, and Yvonne Rogers.

Physikit: Data engagement through physical ambient visualizations in the home.

In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pages 1608–1619. ACM, 2016.

 Open Knowledge Foundation.

History of the Open Definition, 2014.

 Aare Puussaar, Ian G Johnson, Kyle Montague, Philip James, and Peter Wright.

Making open data work for civic advocacy.

Proceedings of the ACM on Human-Computer Interaction, 2(CSCW):143, 2018.

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References

References IV



Joshua G Tanenbaum, Amanda M Williams, Audrey Desjardins, and Karen Tanenbaum.

Democratizing technology: pleasure, utility and expressiveness in diy and maker practice.

In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2603–2612. ACM, 2013.

non era tra quelli scaricati da dataphys, ma viene citato spesso.



Andrea Trentini.

Lombardy EPA *Obtorto Collo* data and anti-pollution policies fallacies.

Journal of e-Learning and Knowledge Society, 10(2), 2014.



David Verweij, David Kirk, Kay Rogage, and Abigail Durrant.

Domestic widgets: Leveraging household creativity in co-creating data physicalisations.

2019.

Dall'opendata alla
physicalization

Andrea Trentini

Perché

Trasparenza

Opendata

Resistenza!

Scraping

Aria

Car sharing

Bike sharing

Chi raccoglie ma
non concede...

(a latere) WiFi analysis

OD websites

Accademia e crisi
della scienza

Data
physicalization

Installazione

Barra prestiti

Lampada

Dekatron

Physiradio

References