Proceedings of the IATUL Conferences   2014 IATUL Proceedings

Jun 3rd, 12:00 AM

Metrics, Altmetrics, Data Visualization

Tuija Sonkkila
Aalto University, tuija.sonkkila@aalto.fi

Metrics, altmetrics, data visualization

Tuija Sonkkila, Data curator, Aalto University Library
tuija.sonkkila@aalto.fi

ABSTRACT

With examples from past, present and future in the making, we look how altmetrics visualizations are positioned into different visualization categories.

INTRODUCTION

By visualization we refer here to data visualization where information has been *abstracted in some schematic form, including attributes or variables for the units of information*[1]. In his extensive and recommended iTunes U course on Data Literacy and Data Visualization, Braumoeller categorizes visualization into spatial, time series, relational, and graphical narratives.

But first things first, as Yau (2013) reminds:

> *without context, data is useless, and any visualization you create with it will also be useless.*

The wider context of altmetrics is still open. As of writing this, the first draft of NISO Alternative Assessment Metrics (Altmetrics) Project[2] is not yet released. Among other non-traditional research outputs, this project explores potential assessment criteria for visualizations.

However, we do need to know the context where data was born: source, time and place, if applicable. For a good introduction to diverse issues on data provenance in altmetrics, see Chamberlain (2013).

BADGES

A state-of-the-art style of presenting altmetrics is to show, either in raw numbers or related to some reference set, how many views, saves, mentions etc. our science outputs have gathered. Here, the JavaScript badges by the Altmetric company have become a de-facto standard. A recent novelty is a scoreless variant, an enhancement that has been on the wish list of their clientele (Adie 2014).

The donut-shaped visualization is like an Executive Summary of altmetrics, concise and carefully branded. In Braumoeller's categorization, it is a combination of a relational graph and a graphical narrative. From the landing page we can continue looking at details of various sources. Among them is also some spatial visualization, the map of tweeters.

Until early 2014, ImpactStory, the US based nonprofit organisation, offered a selection of altmetrics badges too. After the open API was depreciated, badges are available only within the context of ImpactStory profiles[3]. The narrative aspect is even more prominent here. Visually, the badges are

---

[1] http://en.wikipedia.org/wiki/Data_visualization
[2] http://www.niso.org/topics/tl/altmetrics_initiative/
[3] See e.g. https://impactstory.org/DanielGraziotin

[Type text]

designed to avoid chartjunk[4]. With only two colors, they succeed to deliver information both about the actors, and about the magnitude of interest.

REPORTING

PLoS ALM Metrics page[5] is a combination of altmetrics, arranged around a skeleton of different categories of public attention: Viewed, Cited, Saved and Discussed. The naming follows ImpactStory's badge labels.

The only visual element on the Metrics page is a time series, a stacked barchart. It shows a monthly cumulative view statistics, drilling down the total scores presented on top of the page. Hovering above the chart brings up a tooltip that contains a table with detailed statistics of views in different formats. As a bonus, you can compare figures to a number of reference sets, visualized as a green line on the same chart. In other words, this is field-normalized altmetrics in action.

Although we can argue if the barchart in its present form adds or blurs our understanding of the information, the new PLoS ALM Reports[6] is a step forward in this respect. Query results are visualized on the spot with bubble charts, treemaps and geo charts by the Google Chart API. Here we have now spatial, time series and relational visualizations all on the same page. Repositories and CRIS platforms are probably moving to this direction too. Ready-made charts are a big time-saver in a field where data-savvy artists cum coders can produce breath-taking results, but with a price tag.

COMMUNITIES OF PRACTICE

In November 2012, PLoS and ImpactStory co-sponsored an Altmetrics Hackathon. In there, Juan Alperin et al coded ALM Viz[7]. When you need to plug a concise time series visualization on a page with limited space, this is a handy tool.

When and by whom did the article got attention? In his doctoral thesis on citing, Nelhans suggests case studies in scientific practice, something that would be a healthy exercise in altmetrics, too. With visually modest yet informative directed network graphs by HistCite[8], Nelhans shows how different citation patterns form a citation topology (p. 186-194). Here, networks are read as top-down charts of ancestors where the vertical dimension is not used. If we mimicked these networks with altmetrics data, we could think of applying the vertical axis for swimlines that represent different types of sources.

The R statistical software is famous of its active user base of scientists. rOpenSci[9] is a collection of packages for accessing open repositories. To work with altmetrics specifically, there are interfaces for metrics offered by Altmetric, rAltmetric[10], and by PLoS ALM, alm[11]. The power of these solutions is their integration to the versatile graphic libraries of R.

---

[4] http://en.wikipedia.org/wiki/Chartjunk

[5] See e.g. http://www.plosone.org/article/metrics/info%3Adoi%2F10.1371%2Fjournal.pone.0064489

[6] http://almreports.plos.org

[7] https://github.com/jalperin/almviz

[8] http://en.wikipedia.org/wiki/Histcite

[9] http://ropensci.org/

[10] http://ropensci.github.com/rAltmetric/

[11] https://github.com/ropensci/alm

[Type text]

ReaderMeter by Mario Taraborelli was the first example of a visual narrative or infographic[12] of author-level and article-level metrics. A scientific business card or profile page, in other words. Launched in 2010, it was a mashup built on top of readership data from the reference management and document sharing tool Mendeley.

Mendeley and its long-time rival, Zotero, are strong workhorses. Therefore, it seems a logical move from them to expand focus to become academic relaying services. If you are already using the software as your main workbench, you might want to pull other relevant data to the same platform, and push outcomes to repositories, something that Zotero is planning to do (Funding... 2014).

TOWARDS SCALING

Rendered with modern interactive web technics, visualizations of altmetrics raw counts can be visually pleasing[13] but not much more. They are relational visualizations but still, a table of values is only a table of values. Like the ImpactStory Blog (Priem and Piwovar) puts it:

> *The low-hanging altmetrics fruit – thing like simply counting tweets – are increasingly plucked.*

A big step forward would be, if we could compare values along certain scales. Like the Scholarly Kitchen blog has pointed out (Davis 2013), visualizations like the Better Life Index by OECD[14] leave the freedom of weighting and ranking to the user.

From the same OECD data, Jeremy Boy has made an interactive radar chart[15]. Here, you can drag objects closer, like under a magnifying glass, and even compare metrics to others by putting them as layers on top of each other. In Braumoeller's categorization, this is a prime example of relational visualization. However, everything is predefined, and all you can do is play with the visual elements. Still, this type of chart could act as a showcase of different altmetrics sources.

FROM A DIFFERENT ANGLE

In his keynote at the European Communication Summit 2013, Stefaner reminds us that at the time of the first space travels, expectations ran high to see the Moon from close. But instead, eyes were turned to Earth. For the first time, we were able to see our planet from above as a coherent organism.

Perhaps altmetrics will become an eye-opener too. As known[16], altmetrics can also be used for discovery. To quote Kraker (2014), the potential is huge but so far, little used:

> *Similarly to citations, they [altmetrics] can create pathways through science. After all, a citation is nothing else but a link to another paper.*

---

[12] http://en.wikipedia.org/wiki/Infographic

[13] See e.g. Aalto University PLoS altmetrics (D3.js) http://m.lib.aalto.fi/dev/d3/plosalm.html

[14] http://www.oecd.org/statistics/datalab/bli.htm

[15] http://peopleviz.gforge.inria.fr/trunk/oecd_BLI_radar_chart/

[16] See e.g. https://twitter.com/ernestopriego/status/293726140860874753

[Type text]

As part of of his doctoral thesis, Kraker published an application for visualizing scientific domains. Demonstration data consist of articles from Mendeley (Kraker 2013)**.** His hypothesis is that the more often two papers are read together, the closer they are subject-wise.

A physical metaphor can be useful. Stefaner explains how he applied it while building an interactive web application of risk assessment for World Economic Forum (2012). Positioned on same-centric circles around the risk we like to examine, the closer the circle, the closer related it is to the one in the middle.

From the perspective of user interface design and visualizations, when uncertainty or serendipity is a virtue rather than a problem, we need solutions that favour exploring. One example is the search intent prototype made in the Finnish project Revolution of Knowledge Work (Ruotsalo). Here, the radar-shaped universe of evolving search results is related to the same visual thinking as in the works by Kraker and Stefaner above.

THE FUTURE IS HERE

The variety of altmetrics sources is already big, and it is steadily growing in size.

Traditional modes of publishing are changing, which increases the number of digital artefacts. The core function of a journal is no longer a package to articles rather than a brand or a carrier of trust. A similar disruption will happen with books. As soon as scientific communities are able to cite datasets and other volatile outputs, these will become objects of linking too.

There will also be new ways to "perform science" online. From today's perspective, it seems plausible that more and more scientists, at least in some fields, will begin to show not only what they have accomplished but also what they do - right now. There are two interlinked, powerful forces behind this trend.

First, transparency and reproducibility. Today, with computational, web-based platforms such as IPython Notebook[17], added with version control, you can build a forensic track, if you like, from your research. At the same time, embedded computers with e.g. the new Wolfram Language[18], are expanding the base of core computing; we will be able to link to the very place and time when some measurement is taken. And not just link but operate with it, including visualize. An interesting question in this context is, at which point do we need to take a step back in order to see where we stand. How close to the birth of science is it sensible to go with concepts like linking and reproducing, without loosing the man in the middle, the scientist?

Second, evidence-based approach to academic life. In the world of competing resources, HE and scientists need to actively demonstrate their value, whether they like it or not. About the same topic in academic libraries specifically, see Kelly (2011).

The Feltron Annual Reports by Nicholas Felton are legendary[19]. Back in 2005, his first report on himself was considered a foolish endeavor albeit a graphical masterpiece as such. Since then, the Quantified Self movement and My Data have become mainstream. Against this background, this year's April's Fool blog

---

[17] http://ipython.org/notebook.html
[18] Stephen Wolfram's Introduction to the Wolfram Language: https://www.youtube.com/watch?v=_P9HqHVPeik
[19] http://feltron.com/

[Type text]

posting by ImpactStory about their new Total Impact Score was, perhaps inadvertently, a little more than a joke.[20]

CONCLUSION

In the latter part of his keynote, Stefaner gives examples on how some of his visualizations include a transformation from digital to analog; to data sculptures or, like in the visualizations made at the Art of Data Cuisine at Pixelache 2012 in Helsinki, to data food[21].

Despite their tongue-in-cheek character, these kind of experiments carry a strong human message: we are playful social creatures. Who knows, maybe the winner of the Aalto University Campus 2015 architectural competition, which *reveals its functions manifesting courage, ambition, internationality and creativity to be easily observed by the passers-by*[22], will one day carry visual elements on its façades, driven from altmetrics data, or from science itself[23].

---

[20] http://blog.impactstory.org/announcing-a-better-way-to-measure-your-value-the-total-impact-score-2/

[21] http://www.pixelache.ac/open-data-cooking-workshop/

[22] http://campus2015.aalto.fi/en/finalists/vare/

[23] Thanks for the idea to Timo Kiravuo who in the Aalto University Yammer group brainstormed a mechanically-produced fog cloud hovering above the University main building, signaling current energy consumption.

[Type text]

REFERENCES

Adie, E. 2014. Kick out the jam: scoreless donuts now available. http://www.altmetric.com/blog/kick-out-the-jam-scoreless-donuts-now-available/

Braumoeller, B. F. Data Literacy and Data Visualization https://itunes.apple.com/us/course/data-literacy-data-visualization/id693097601

Chamberlain, S. 2013. Consuming Article-Level Metrics: Observations And Lessons From Comparing Aggregator Provider Data, *Information Standards Quarterly*, Summer 2013, Vol 25, Issue 2, pp. 4-13 http://www.niso.org/apps/group_public/download.php/11303/FE_Chamberlain_Consuming_ALMs_isq_v25no2.pdf

Davis, P. 2013. Visualizing Article Performance - Altmetrics Searches for Appropriate Display. *Scholarly Kitchen* Sep 30, 2013 http://scholarlykitchen.sspnet.org/2013/09/30/visualizing-article-performance-altmetrics-searches-for-appropriate-display/

Funding for Altmetrics Research and Expanded API. *Zotero Blog*. April 21st, 2014. https://www.zotero.org/blog/funding-for-altmetrics-research-and-expanded-api/

Kelly, B. 2011.The Need for an Evidence-based Approach to Demonstrating Value. *UK Web Focus* http://ukwebfocus.wordpress.com/2011/12/28/the-need-for-an-evidence-based-approach-to-demonstrating-value/

Kraker, P. 2013. Head Start, Mendeley Labs http://labs.mendeley.com/headstart/

Kraker, P. 2014. Open and transparent altmetrics for discovery. http://science20.wordpress.com/2014/01/13/open-and-transparent-altmetrics-for-discovery/

Nelhans, G. 2013. Citerings praktiker. Det vetenskapliga publicerandet som teori, metod och forskingspolitik. Göteborgs Universitet. *Filosofi, lingvistik och vetenskapsteor*i. http://gup.ub.gu.se/publication/184221 (Summary in English)

PLOS ALM Hackathon. 2012. http://article-level-metrics.plos.org/alm-workshop-2012/hackathon/

Priem, J., Piwovar, H. 2014. Top 5 altmetrics trends to watch in 2014. *ImpactStory blog*. http://blog.impactstory.org/top-5-altmetrics-trends-to-watch-2014/

Ruotsalo, J. et al. 2013. Directing exploratory search with interactive intent modeling. In *Proceedings of CIKM 2013, the ACM International Conference of Information and Knowledge Management.* http://research.ics.aalto.fi/mi/online-papers/Ruotsalo13cikm.pdf

Stefaner, M. 2012. Project for World Economic Forum http://moritz.stefaner.eu/projects/world-economic-forum/

Stefaner, M. 2013. Keynote at European Communication Summit 2013 http://vimeo.com/71798105

[Type text]

Taraborelli, D. 2010. ReaderMeter: Crowdsourcing research impact
http://nitens.org/taraborelli/ReaderMeter

Yau, N. 2013. Data Points: Visualization That Means Something. Somerset, NJ, USA: Wiley.

[Type text]