Ubiquity Symposium

Big Data

Big Data and the Attention Economy

by Bernardo A. Huberman

Editor's Introduction

While attention has always been prized above money, few people have had the means to attract it to themselves. But the new digital economy has provided everyone with a loudspeaker; thus efforts at getting noticed have rapidly escalated in global society. The attention economy focuses on the mechanisms that mediate the allocation of this scarce entity. Social networks and big data play a role in determining what is noticed and acted upon.

Ubiquity Symposium

Big Data

Big Data and the Attention Economy

by Bernardo A. Huberman

Not surprisingly, instant access to information across geographic and institutional boundaries has made its value plummet in economic terms. Search engines, myriad apps, social networks, cloud computing, and billions of websites offer instant news, entertainment, human connections, videos, and the answer to most questions at the speed of light. Welcome to the digital economy, a massless entity whose unstoppable spreading creates much of the growth we see in the world today.

We value what is scarce, not what is plentiful, and the precious entity that we all seek, and is hard to manufacture at will, is attention. Such focus is always finite, ephemeral and claimed by many at the same time. This explains the intense efforts made at obtaining it through targeted advertising, shocking news, catchy titles, pop-ups, the artistic presentation of ideas online, videos, and most disheartening, spam.

The attention economy, which focuses on how this scarce entity is allocated among content and people, now plays a prominent role in every aspect of life. It affects what consumers think and buy, the outcome of elections, political discourse, group think, scientific and technological advances, and how decision makers decide among available choices [1–4].

While attention has always been prized above money [5], few people have had the means to attract it to themselves. But the new digital economy has provided everyone with a loudspeaker, and so efforts at getting noticed are rapidly escalating at a global scale. An interesting example of this arms race is provided by academic life. Attention in the scholarly world is often the main currency, for we publish to get the attention of others, we cite the work of colleagues so that they receive attention, and we cherish the prominence of great work if only because of the attention it gathers [2, 3]. And while attention to topics and scholars has been around since the establishment of learned societies and academic disciplines, the digital

medium has given it such a level of prominence that it threatens to distort the research agenda and how it is valued by society as a whole.

Nowadays it is not enough to publish one's research results. They need to be broadcasted to several audiences, with clever titles and illustrations that will hopefully elevate them above the many other related papers so as to catch the attention of the media and hopefully lead to increased visibility and funding. In many cases that attention also translates into hiring decisions, awards, and promotions.

As millions of people use the web for their social, informational, and consumer needs, they propagate information through vast social networks that extend far beyond the traditional channels determined by geography and personal relations. At the same time, content providers vie for the limited attention of people by resorting to strategies aimed at maximizing the amount of attention devoted to their content [5]. These strategies range from the personalization of content, impactful videos, and click-bait headlines, to the dynamic rearrangement of items in a given page to suit the needs of the user. In all these cases the ultimate goal is the same: to draw the attention of users to specific content before they drift away to something else.

The role social networks play in the dissemination of ideas, purchases, and reputations is also accelerating due to the ease with which vast amounts of data can be gathered and analyzed on a scale that was impossible using traditional methods. A simple example of how that can work is shown in the figure below [6]. It graphically depicts the results of an analysis of a network of recommendations responsible for the purchase of books. By studying the networks that grew up around each book—who bought and recommended it, and who responded to the recommendation—one notices that social networks take on different characteristics depending on the type of books that are recommended.

In the figure, red dots and lines indicate people who purchased a product while blue dots and lines represent people who received a recommendation. The network around a medical book (small graph in the upper left-hand corner) is a rather scattered one where recommendations, on average, don't travel very far. The network surrounding a Japanese graphic novel, on the other hand, which occupies the central part of the picture, shows a thick flow of information among densely connected groups of people.

Figure 1. A recommendation network for books sold by Amazon. Red dots and lines indicate people who purchased a product while blue dots and lines represent people who received a recommendation.

A similar methodology can be used for any other kind of information linking people, their ideas and their consumption profiles. Several years ago, for example, Adamic and Glance [7] studied the linking patterns and discussion topics of political bloggers during an election in order to measure the degree of interaction between liberal and conservative blogs and to uncover any differences in the structure of the two communities. Among their many finding, a significant one was the difference in the behavior of liberal and conservative blogs, with conservative blogs linking to each other more frequently and in a denser pattern than the liberal ones. Similarly, intramural email patterns inside an organization can provide an automatic way of identifying communities of practice within an organization [8, 9].

Another important aspect of social networks is how they mediate the allocation of attention to given topics, products, or ideas while ignoring others. To understand this, consider as an

example how a news story spreads among a group of people. When it first comes out, the story catches the attention of a few, who may further pass it on to others in their social network if they find it interesting enough. If many people start to pay attention to this story, its exposure in the media will continue to increase. In other words, a positive reinforcement effect sets in such that the more popular the story becomes, the faster it spreads.

Another factor that determines success in the allocation of attention is novelty, for few people would keep attending to content they viewed already or is stale in terms of its validity. That is why most news aggregators, and sites ranging from Twitter and Facebook to arXiv, order the links of a given page by their novelty, so as to guarantee a high degree of attention. This strong interplay between novelty and collective attention is universally manifested in a rather swift initial growth of the number of people attending to a new item and its eventual slowdown as interest fades among the population [10].

Another factor that affects the allocation of attention stems from the finite number of items that a user can attend to in a given time interval. This psychological constraint is compounded by a strong empirical regularity observed in web browsing that goes under the name of the law of surfing [11,12]. This law states the probability of a user accessing a number of web pages in a single session markedly decays with the number of pages, thereby constraining the amount of information that ever gets explored in a single surfing session. Because of this, a typical user seldom visits pages beyond the first one displaying search results, and consequently a page ranked near the bottom by a search engine is unlikely to be viewed by many users. This behavior tends to reinforce the leading position of those top items and further increase their popularity, which in turn penalizes new content that is not well known yet. Thus it is easy for an item to get locked in a top ranking and hard for other bottom items to surface, even though the latter can often be more valuable.

And yet in spite of all these obstacles, we manage to remain up to date with our work and even discover every once in a while, interesting new facts and ideas that are relevant to our intellectual activities. We often do that through a social network of like-minded people, colleagues, and friends that quickly propagate novel ideas and their opinions about them. These networks, sometimes called informal colleges or communities of practice, were identified a long time ago as important channels for the dissemination and validation of new results in a given discipline [9,14]. With the advent of the web, the scope and swiftness of these networks has increased by several orders of magnitude.

And yet this is not the complete story, for there are other drivers of social attention that can be as effective as novelty in eliciting it. Popularity is one such factor, which often leads us to read

and examine ideas if only because others do. Or style, as is the case when visual elements make an idea or presentation initially compelling because of its elegance or esthetic value. Richard Lanham has eloquently described the significant role that the arts and letters play in this attention economy by serving as creators of attention structures, mostly through fields like rhetoric and the visual arts [15].

In the digital economy, social attention and its swift allocation through vast social networks plays a central role in the dissemination of ideas, news, and ideologies that can at times be both extremely influential and also harmful. The speed of this development not only is disrupting the traditional channels of information transmission such as scientific journals, television, and printed media, but also leading to the expression consumer preferences and political outcomes that have large effects on the functioning of the global society.

References

- [1] Goldhaber, M. H. Attention economics and the net. *First Monday*. 1998.
- [2] Franck, G. Science communication, a vanity fair. *Science* 286, 53–55. (1999).
- [3] Klammer, A. and Van Dalen, H. P. Attention and the art of scientific publishing. *Journal of Economic Methodology* 9 (2002), 289–315.
- [4] Parolo, P. D. B., Pana, R.K. , Ghosh, R., Huberman, B. A., Kaski, K., and Fortunato, S. Attention decay in science. [Journal of Informetrics](#) 9 (2015), 734-745.
- [5] Huberman, B. A., Onculer, A., and Loch, C. Status as a valued resource. *Social Psychology Quarterly* 67, 1 (2004), 103–114.
- [6] Lefkovich, J., Adamic, L., and Huberman, B. A. The dynamics of viral marketing. *ACM Transactions on the Web* 1, 5 (2007).
- [7] Adamic, L. and Glance, N. The political blogosphere and the 2004 US election: divided they blog. In *Proceedings of the Third International Workshop on Link Discovery*. 2005.
- [8] Tyler, J. R., Wilkinson, D. M., and Huberman, B. A. Email as spectroscopy: Automated discovery of community structure within organizations. In *Proceedings of the International Conference on Communities and Technologies*. Kluwer Academic Publishers, 2003.
- [9] Crozier, M. *The Bureaucratic Phenomenon*. University of Chicago Press, Chicago, 1964.

- [10] Wu, F. and Huberman, B. A. Novelty and collective attention. In *Proceedings of National Academy of Sciences*. 2007.
- [11] Huberman, B. A., Pirolli, P. L. T., Pitkow, J. E., and Lukose, R. M. Strong regularities in World Wide Web surfing. *Science* 280, 5360, (1998), 95–97.
- [12] Huberman, B. A. *The Laws of the Web: patterns in the ecology of information*. MIT Press, Cambridge, 2002.
- [13] Wasserman, S. and K. Faust. *Social Network Analysis*. Cambridge University Press, Cambridge, 1994.
- [14] Lanham, R. *The Economics of Attention: Style and substance in the age of information*, Chicago University Press, Chicago, 2006.
- [15] Huberman, B. A. and Adamic, L. Information dynamics in the networked world. In *Complex Networks*, E. Ben-Naim, H. Frauenfelder and Z. Toroczkai eds. Springer, 2004 371-395.

Biography

Bernardo A. Huberman is a Fellow and Vice President of Core Innovation at CableLabs.

DOI: 10.1145/3158337