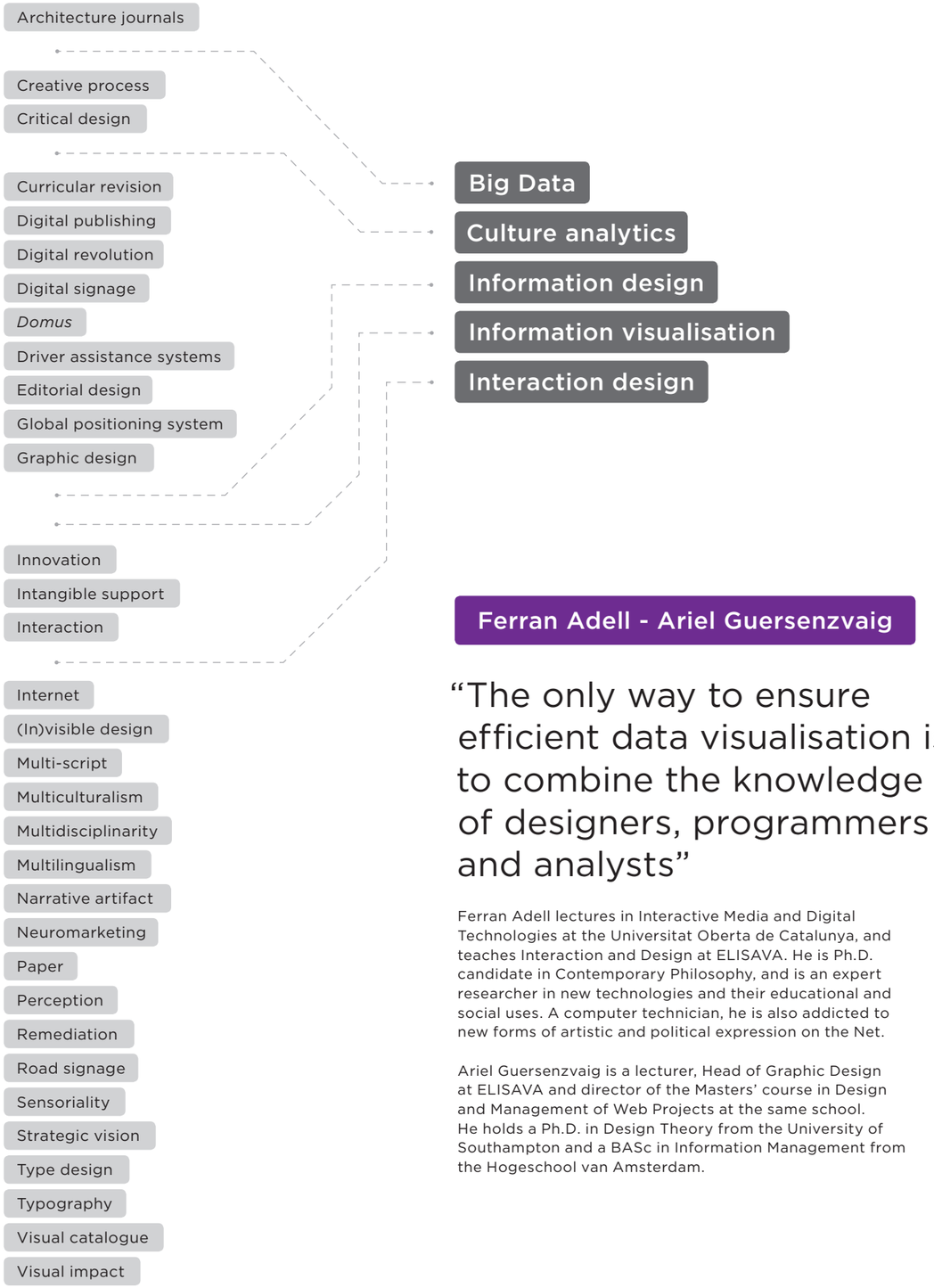


Ferran Adell - Ariel Guersenzvaig

"The only way to ensure efficient data visualisation is to combine the knowledge of designers, programmers and analysts"

Ferran Adell lectures in Interactive Media and Digital Technologies at the Universitat Oberta de Catalunya, and teaches Interaction and Design at ELISAVA. He is Ph.D. candidate in Contemporary Philosophy, and is an expert researcher in new technologies and their educational and social uses. A computer technician, he is also addicted to new forms of artistic and political expression on the Net.

Ariel Guersenzvaig is a lecturer, Head of Graphic Design at ELISAVA and director of the Masters' course in Design and Management of Web Projects at the same school. He holds a Ph.D. in Design Theory from the University of Southampton and a BASc in Information Management from the Hogeschool van Amsterdam.

Big Data and New Ways of Visualising Information

In the current context we must come up with new ways of organising and displaying information for users, whether they be scientists working with the huge results of a particle accelerator, or ordinary people looking up their local climate data obtained from numerous environmental sensors.

The term Big Data refers to large sets of digital data that is captured, stored, searched, manipulated and visualised by means of computer systems. The increase in number of sources and the exponential growth of implicit and explicit data capture by smartphones, tablets and a host of new digital sensors have produced a new scenario as regards data treatment.

The digital revolution has transformed our ways of relating to our environment and of constructing knowledge. We find ourselves before a typical paradigm shift—crisis, revolution and establishment of a new paradigm. Kuhn's classical diagram[1] is of no use when we speak of technology and are not judging one specific example. Technological society is dynamic and changing, and unsparingly rejects the systemic theories we have grown accustomed to when it comes to constructing knowledge. Variables have taken centre stage in today's society; in the case of technology, they have so reduced the validity of reflection that more than results the academic world suggests possible paths.

One of the most relevant changes directly brought about by new technologies is our increasing ability to generate, store and manage data. The constant upgrading of computer capacity and of storage devices, the effective participation of users as prosumers and the growing volume of active and passive data-compiling sensors have turned upside down the equations designed to define the growth of the corpus of knowledge.

In 2008 the volume of data increased by 487 billion gigabytes (GB) and it is estimated that in 2012 this amount multiplied by five; no growth reduction is

1. T.S. Kuhn, *The Structure of Scientific Revolution*. University of Chicago Press, Chicago, 1962.

foreseen, therefore the challenge of preserving and transmitting knowledge is but in its initial stages. We have reached one zettabyte (ZB) of digital storage, the equivalent of over a thousand billion former floppy discs.[2]

The Challenges

Belonging as we unquestionably do to a generation that has produced more information than those that have come before us, philosopher David Casacuberta asks himself in his blog whether we could perhaps also be the first generation unable to transmit this information. The fact is that even in the bosom of technological disciplines, as in most techno-social phenomena, ethical and philosophical reflection is unable to progress at the same speed as new technologies and their implications, 'And this is the great paradox of our culture: we have by large the highest record of cultural actions, yet in formats that will not survive the passage of time. What will actually remain of Barack Obama's online campaigns? What has become of the information lovingly stored in the vintage pages of Geocities? Are any copies still left of Wordstar? Can they still be executed? How much information is lost forever when an organisation decides to fill a container with unusable floppy discs and throw them away?'[3]

Managing and transmitting acquired knowledge is surely one of the greatest challenges that mankind as a social species is facing, and the only way of meeting this challenge is to devote time and resources to advancing the management and classification of data and developing new ways of visualising information. The huge potential of Big Data processing systems as a basis for cultural analysis seems to be one of the possible answers for transforming data into conclusions that can be 'archived'. This means that while it is almost impossible for us to transmit the totality of the Facebook messages that made up Barack Obama's online political campaign, we can transform this data into infographics, analyses or compilations and capture their useful essence, which would otherwise be lost. This is the only way we can guarantee a continuity in the history of knowledge and allow future generations to continue to study the past and so be able to learn from it.

We cannot make backups of the connections between things; we cannot provide cloud computing for the interaction between society and trending topics on Twitter; we cannot archive the interconnected nature of data, which is why the cultural analysis that derives from it is even more important.

Data today is the basis of knowledge; without it there can be no technology (or medicine, or alternative fuel sources, for instance). And without the analysis of data, no serious in-depth social, cultural or economic study can be made. Without deep reflection on its uses, its limitations and related ethical rights, data is only data. The terms *data* and *information* are often mistakenly used, even in scholarly articles, despite being clearly defined. Data is not useful *per se*—

2 C. Sáez, *¿Dónde están mis datos?* [online], Grup Godó, Barcelona. [Accessed: 8 March 2013]. Available at: http://www.lavanguardia.com/estilos-de-vida/20120719/54326650737/donde-estan-mis-datos.html

3 D. Casacuberta, *Entender nuestra cultura, net art y el Big Data* [online]. [Accessed: 8 April 2013]. Available at: http://elbigdata.net/entender-nuestra-cultura-net-art-y-el-big-data/

it must be related to something in order to suggest meanings, knowledge, conclusions. It doesn't constitute information, nor is it able to communicate meaning. Only the organised set of processed data becomes information, which is when it takes on the utility it lacked as independent data.

We must therefore face up to important tasks in information management, i.e., not only storage and recovery but also the process of data transformation and information visualisation, in scientific and academic contexts and in other social environments. Data is the element on which we have learnt to construct scientific knowledge, and for some decades now, social knowledge as well. What isn't so obvious is that we are now ready to make the most of data when it is measured in thousands of millions of bytes.

▲ DOMO. *Every Minute of the Day*, 2012. (Source: http://img.xatakaciencia.com/2012/07/informacion-internet-minuto.png).

Big Data Culture

In a context in which technological progress is far too easily separated from academic reflection we need to take a deep breath and consider which path will lead us to integrate Big Data as a key element of contemporary research, and how we shall present the results of this research to the public.

Contemporary society is governed by great economic powers. In the case of Big Data, the commercial race for data was begun by large corporations to secure themselves a sizeable portion of what promises to be the sweetest cake of the next few years. So, those companies, organisations and governments that are spending money researching and developing the potential of Big Data will be the ones that in the not too distant future will reap the fruits of the rich harvest of data. However, while Big Data and its implications are the subject of numerous postgraduate and masters degrees in the best American universities and in renowned research centres such as MIT, it is still quite difficult in Spain to draw the attention of academia and promote new initiatives in the field.

As well as triggering new and promising occupations for scientists and data analysts,[4] Big Data is a cultural phenomenon that affects us all. In a world in which we have reached the zettabyte of digital information, both our relationship with data processing and our way of showing data must needs change. We are no longer able to interpret information without computer mediation; we need software for processing it but also for shaping a volume of information that escapes our cognitive ability. And this is the truly important process that has a

4 T. H. Davenport and D. J. Patil, *Data Scientist: The Sexiest Job of the 21st Century* [online]. Harvard: *Harvard Business Review* [Accessed: 8 April 2013]. Available at: http://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/

▲ *A Picture Is Worth a Thousand Words*. (Source: http://www.blindfiveyearold.com/the-future-of-search-and-seo)

direct bearing on today's designers: the patterns with which we present information to the public must necessarily change when we are contending with a totally different type of data.

In an interesting article, Bill Franks stated 'Our brains are meant to see in pictures. Grids and columns of data, while ubiquitous, make it very difficult to see trends or patterns. Additionally, a lot of the new data sources available today, such as genetic data or social network data, don't lend themselves to traditional spreadsheets and graphs. These data types require a different way of displaying them to allow us to see the underlying patterns and stories in the data.'[5]

Furthermore, if the need to change the way in which we present data was already obvious when we were managing much smaller amounts of information, today it is simply indisputable. The example of the map of the United States suggested by Franks in his article is a clear and straightforward exercise that shows us the importance of information visualisation and its benefits. Imagine a map of the regions of Spain or the provinces of Catalonia: if we show the map to an audience with a minimum knowledge of their local geography it should be able to easily identify some of the places on the map and some of their main characteristics—whether or not it borders on another region, whether or not it has a coast, what its geographical features are, etc. On the other hand, if we show the same audience a list of data in the form of tables that present the same information as the map, including geographical references, geolocation coordinates and a large number of numerical data, its ability to receive and process this information will be much smaller, giving rise to name mistakes and misunderstandings.

Is a picture worth a thousand words? Well, although broadly speaking we cannot absolutely confirm the popular saying, when we are dealing with huge amounts of data, *it's a must!* Visualisation converts databases into accessible and valuable information, thereby enabling us to reach certain conclusions that would otherwise be impossible. But we are not only referring to the treatment

5 B. Franks, *The value of a good data visual: immediacy* [online]. Harvard: *Harvard Business Review* [Accessed: 8 April 2013]. Available at: http://blogs.hbr.org/cs/2013/03/the_value_of_a_good_visual_imm.html

of geographical data—just think how hard it is to express in numerical notation the relationships established between members of different social networks. The friend of a friend of a friend, who in turn is the friend of another friend who wasn't a friend of the first friend … and so on, to infinity and beyond! What can be shown in infographics as an entity relation diagram (ERD) becomes a communicative fantasy if all we are working with are numerical data, tables and bar graphs.

"We need software for processing it but also for shaping a volume of information that escapes our cognitive ability"

Data visualisation enables us to establish relationships, patterns and stories in a much more practical and intuitive way than that offered by databases. So, as Bill Frank asserted, the popular maxim applicable around the world that 'a picture is worth a thousand words' is true.

The introduction of modern technologies in the service of new ways of visualising information also entails other advantages, such as interaction with users. Visualisation is no longer static, and we can manipulate, redefine and condition a complex graph with a simple click of the mouse. As a result, in order to focus on a specific spot and enlarge a certain area in a given city, we no longer need a set of maps of different scales.

Interactivity opens up new fields for cultural analysis starting from visual maps, and generates new relational spaces between data, users and mediators of information. This is applicable to the field of trade—we can now show information to customers according to presumptive approaches and gradually adapt it to their needs, questions or concerns.

'When the chart is up on the screen and a client asks a question that requires a different view of the data, it is easy to drill into that view on the fly. No more sending an email later in the day with another chart.'[6] Yet where there is interaction with users there must also be a concern with usability. It is not technical efficiency or visual success that should guide visualisation projects, but user interest, ease of use and adaptation to visual intuition. Here, once again, UX (User Experience) designers must contribute their knowledge in order to promote user experience and ensure they meet their aims.

6 Ibidem.

John Sviokla mentions the three main benefits of modern visualisation methods.[7] In the first place, good visualisations are efficient because they enable users to analyse large amounts of data very quickly. They also help analysts study the nature of problems in depth and discover new positions, aspects or approaches for solving them. Last but not least, a good visualisation of data helps create shared points of view of specific situations and unite experts around necessary actions.

The new methods of information treatment guarantee correct interpretation of data and of its use in the service of knowledge. Visualisation allows us to confirm or refute social or political theories, to show something we already know (telling the facts), and to open up new paths of exploration transforming large

7 J. Sviokla, *Swimming in Data? Three Benefits of Visualization* [online]. [Accessed: 8 April 2013]. Available at: http://blogs.hbr.org/sviokla/2009/12/swimming_in_data_three_benefit.html

data sets (that are very difficult to manage and analyse) into models that enable us to predict the behaviour of a system and learn to manage it more efficiently.

'I believe that we will naturally migrate toward superior visualizations to cope with this information ocean. Since the days of the cave paintings, graphic depiction has always been an integral part of how people think, communicate, and make sense of the world. In the modern world, new information systems are at the heart of all management processes and organizational activities.'[8]

8 Ibidem.

An interesting example of the application of information visualisation for the development of science is the project entitled *Planet Hunters*, created by the Madrid design studio Vizzuality for NASA in collaboration with physicists at the University of Oxford. Planethunters.org is a platform open to astronomers, professional and amateur alike, the ultimate objective of which is to detect the existence of planets outside our solar system. Planet Hunters tries to visualise the huge amount of data provided by the space probe Kepler, in orbit around the Sun. Although NASA uses extremely powerful computers to analyse such data, the project strives to obtain a visually accessible presentation of the information so that it may be intelligible to the human mind, taking advantage of our great talent for recognising visual patterns. The data visualised are the brightness measurements of 150,000 stars made every thirty minutes. Platform users attempt to detect changes in brightness, for these may indicate the existence of a planet. Human beings are able to see things, such as small planets the size of the Earth, that the algorithms of the most powerful computers are unable to detect.

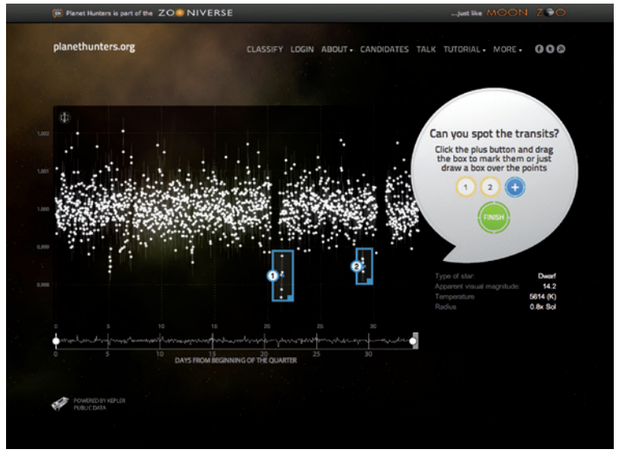
Data Scientists and Data Designers

An analysis of Big Data and its implications for information visualisation, however, wouldn't be complete if it didn't explain the tools and knowledge required in the application of these new methods. Amanda Cox, graphic editor of *The New York Times*, says that the qualities needed for its development are now very difficult to find in a single professional. Programming, statistics and graphic design come together in work that must be collaborative, the sum of different visions and abilities.[9] Behind all good visualisation projects there must be good design, effective codes and correct data management and analysis.

This is the only way we shall arrive at a safe harbour, establishing a marked informational hierarchy and promoting academic study of how we perceive information, how we manage it and what our intuition tells us about it. The computer code is thus placed in the service of a new cultural objective. In order to transform data into information we must prioritise this informational objective over and above dark and scarcely functional statistics.

While Big Data is currently the focus of all debates and it is difficult to establish which initiatives will truly accomplish significant progress, some projects have already made significant contributions. Some time ago Manovich, one of

9 S. Berinatto, *The Power of Visualization's "Aha!" Moments* [online]. [Accessed: 8 April 2013]. Available at: http://blogs.hbr.org/hbr/hbreditors/2013/03/power_of_visualizations_aha_moment.html



▲ Vizzuality. Interface of the *Planet Hunters project*, 2011. (Source: www.vizzuality.com)

the most important new media theoreticians, realised the need for software in the service of cultural analysis and started a huge project on what he himself termed *Cultural Analytics*,[10] the structuring, recovery and visualisation of information with a view to the efficiency of the said cultural analysis.

'Cultural Analytics is the use of computational and visualization methods for the analysis of massive cultural data sets and flows. […] We can create more comprehensive and inclusive understanding of human cultural evolution and dynamics using all digitized and born-digital cultural artifacts in any media from all of human history. We can map changing cultural patterns around the world in *real time* using all available web data. Reaching these goals will require not only solving many technical problems but also developing new methodologies and concepts for data-driven cultural analysis.'[11]

Manovich understood that in order to develop the project he needed to program new computer tools that would enable him to digitise and process huge amounts of information in order to obtain really useful visual maps. As mentioned, the only way to ensure efficient data visualisation is to combine the knowledge of designers, programmers and analysts in pursuit of a common objective. Therefore, the profile we should associate with the challenges of information can no longer be defined bearing in mind a single discipline. We must overcome our fear of digital technology through design and our fear of design through digital technology—this is the only way we shall find a common space for reflection and progress that will allow us to deal with the problems derived from Big Data with the gravity they deserve, over and above the advertising boom and focusing on their true meaning.

10 L. Manovich, *Software Studies: Cultural Analytics* [online]. [Accessed: 8 April 2013]. Available at: http://lab.softwarestudies.com/2008/09/cultural-analytics.html#1

11 Ibidem.