

PROGRAMME DE RECHERCHE ¹

RESEARCH PROGRAM ¹

Nom / *Last name* : Venturini

Prénom / *First name* : Tommaso

Equipe(s)-projet(s) d'affectation souhaitée(s) / *Project(s)-team(s) assignment wishes* : DANTE

Intitulé du programme de recherche / *Title of research program* :

Coevolution of structures and flows in complex networks

Two parallel strains of research emerged in the social and in the computational sciences that, despite their different disciplinary roots, present a striking affinity. Reading from the presentation of DANTE:

A large amount of work has... highlighted the role of the network topology on the dynamics of the spreading. However, the dynamics of the networks (i.e. topology changes), and in the networks (i.e. spreading processes), are still generally studied separately. There is therefore an important need developing tools and methods for the joint analysis of both dynamics.

This argument could be repeated with the very same words for the social sciences. A great deal of reflection has been dedicated to the evolution of macro-structures (i.e. the history of institutions, norms and social fields) and to the dynamics of micro-interactions (i.e. the interpersonal contacts and exchanges). Very little to the interference between micro and macro processes.

Modelling collective life, **we generally assume that structural transformation and individual interactions** (or 'micromotives' and 'macrobehaviors' to use Shelling famous notions) **occur at different time scales** – as a railway renovating at a speed that does not disrupt the circulation of individual trains, or as ants excavating at a speed that does not disrupt the architecture of the nest. While acknowledging the oversimplification of such separation, we maintained it because of the difficulty to obtain information on both dynamics at the same time. In a very convenient way, the separation between micro and macro reflected the divide (customary in social sciences) between qualitative and quantitative research methods.

In the last few years, however, the access to collective information has dramatically changed. Reading again from the presentation of DANTE:

Large datasets describing such networks are nowadays more "accessible" due to the emergence of online activities and new techniques of data collection. These advantages provide us an un-precedent avalanche of large data sets, recording the digital footprints of millions of entities (e.g. individuals, computers, documents, stocks, memes etc.) and their temporal interactions.

The advent of digital traceability affects both social and computational sciences and urges their collaboration. It opens the possibility to study simultaneously the dynamics *of* network and the dynamics *in* networks, but it also raises a multitude of conceptual and technical questions that can only be solved by an interdisciplinary alliance.

Exactly one year ago, in March 2015, I called for such a collaboration in a paper for the Journal of Artificial Societies and Social Simulation (Venturini, T., Jensen, P., & Latour, B. Fill in the Gap: A New Alliance for Social and Natural Sciences. *JASSS*, 18(2):11. jasss.soc.surrey.ac.uk/18/2/11.html). The fact, since its publication, the paper has never left the rank of the most viewed articles of the journal (jasss.soc.surrey.ac.uk/admin/top20.html) hints at the need of such interdisciplinary alliance.

Identifying a need, however, is much easier than solving it. And this is way I wish to join (for a few years at least) the team of DANTE. In my previous researches (cf. the above forms), I have worked on the cartography of hyperlinks; on scientometrics landscape; on open legislative data; on international climate negotiations; on online debates. In all these experiences, I painfully felt the **lack of the computational tools to handle simultaneously the local and global dynamics** of large scale networks.

This lack, which I intend to address with the help of INRIA researchers, exists at several levels:

- In the **collection** of traces, where it is crucial to harvest in the largest and 'rawest' way, in order to impose the least possible constraints to the later data analysis.
- In the **storage** of data, where innovative memory structures are necessary to assure (for large dataset) the conservation of temporal meta-data *and* the efficacy of data retrieval.
- In the **analysis**, where intelligent algorithms are necessary to handle collective transformations that not only occur at different temporal and spatial scales, but can also speed up or slow down, extend or retract.
- In the **modelling** of collective dynamics, where a new generation of models are demanded to address *in a fully combined* way, the movement through structures and the transformation of structures.

- In the **visualization** of information, to develop interfaces allowing a seamless navigation between different temporal and spatial aggregations.
- In the **communication** of findings, to provide interpret and ‘narrate’ results in ways that maintains the continuity between macro-structure and micro-interaction.

To INRIA, I will bring the knowledge (and the datasets) I have collected on various collective phenomena and, more generally, **the viewpoint of a social scientist well-acquainted to the milieu of computational research**.

[1] The main goal of the DANTE team is to lay solid foundations to the characterisation of dynamic networks, and to the field of dynamic processes occurring on large scale dynamic networks. [2] In order to develop tools of practical relevance in real-world settings, we propose to ground our methodological studies on real data sets (ibid.).

Because of my interest in the digital traces of public debates, my research would benefit from the techniques developed in the first point. Because of my expertise in social methodology, I can offer help on the second. **My research agenda on digital methods is, I believe, perfectly symmetrical to DANTE’s**. This is why this research project is deliberately unspecific: the point is not imposing a case study to the INRIA, but to offer an interdisciplinary collaboration that can be mutually beneficial.

An example (purposely far from my previous research) can however illustrate the potential such collaboration. A few years ago, some of DANTE’s researchers employed from signal processing to study temporal and spatial patterns in the use of public bicycles in Lyon (cf. Borgnat, P. et al. 2011. Shared Bicycles in a City: A Signal Processing and Data Analysis Perspective. *Advances in Complex Systems*, 14(3):415-438). While idea is fascinating, the investigation could have been pushed further with more support from the social sciences. Analysing the movement of Lyon’s bicycles, the researchers assumed the urban organisation as stable. As the data on the *Vélo’v* accumulate year after year, however, it become possible to investigate, how the urban topology affects and is affected by the introduction of a new system of public mobility (for instance enlarging or displacing the touristic areas of Lyon).

And if the *Vélo’v* example sounds a bit trivial, think of how, in a similar way the appearance of new type of financial transactions – ‘high frequency trading’ – has influenced the very structure of the Internet (with trading companies moving their server-farms closer to Internet Exchange Points and even building new optic and microwave backbones to gain milliseconds).

In both examples, the coevolution of flows and topology is lost if structures and interactions are studies separately. The understanding of phenomena of such kind demands a profound rethinking of the way in which we study network, both socially and computationally. My double background in social and data sciences places me, I believe, in ideal position to contribute to such rethinking, especially in collaboration with INRIA.

CONTRIBUTIONS MAJEURES

MAJOR CONTRIBUTIONS

Nom / *Last name* : Venturini

Prénom / *First name* : Tommaso

Fiche 1 : Controversy Mapping

1. Description de la contribution / *Description of the contribution*

Controversy mapping (CM) is an original pedagogical and research methods germinated some twenty years ago in the French school of Science and Technology Studies (STS). The objective of this approach is double:

1. **Methodologically**, it investigates sociotechnical controversies (e.g. the disputes about GMOs, nuclear power, nanotechnologies...) to reveal how science and technology participate to the construction of collective life. Sociotechnical debates are the moment in which the role of scientific ideas and technical objects are explicitly discussed reveal the networks of internal components and external supporters that constitute them.
2. **Politically**, CM intends to facilitate public discussions about science and technology, by listing the actors of these debates; deploying their position and viewpoints; describing their alliances and oppositions; and investigating their balance of power. By this study, controversy mapping hope to make sociotechnical debates more readable to the public opinion and to encourage engagement.

2. Contribution personnelle de la candidate/du candidat / *Personal contribution of the applicant*

- In my **Ph.D. thesis**, I carried out such an investigation using the analysis of the legal disputes about seed patenting to reveal a deeper and larger discontent with agricultural modernization.
- In the **MACOSPOL** project, I collaborated with Bruno Latour to consolidate the method of CM, by organizing a vast discussion with all the researchers and teachers employing it.
- With the project **EMAPS** (www.emapsproject.com) and **MEDEA** (<http://medea.medialab.sciences-po.fr>), for both of which I was the Principal Investigator, I tried to demonstrate how CM could be usefully employed to analyse public debate about the adaptation to climate change (see also the contribution 4)
- By mounting the project **FORCCAST** (<http://forccast.hypotheses.org>) and coordinating its 1st axis for two years, I contributed to the development of CM as teaching tool at university and high school level.
- More recently I have been writing a **'fieldguide' on CM**. This text constitutes both my HDR thesis and a book currently under review of MIT Press ('Inside Technology' Collection).

3. Originalité et difficulté / *Originality and difficulty*

Despite its solid tradition and upward trajectory, CM is still a highly experimental approach which cannot count on the consolidation of traditional research methods. Its use both in the classroom and in research is hindered by the difficulty to combine its two sources of inspiration: the conceptual heritage of French STS (and Actor-Network Theory in particular) and the data exploration techniques of Digital Methods. My main contribution to CM has precisely been to combine these two heterogeneous strands assuring their theoretical and practical compatibility.

4. Validation et impact / *Validation and impact*

1. **As a teaching method**, CM is currently taught to some 1500 students in various European and American universities (Paris, Copenhagen, Milan, Oxford, Manchester, Amsterdam, Liège, Oslo, Padova, Trento, São Paulo, Rio de Janeiro, Buenos Aires...). More than Since 2012, the pedagogical development of controversy mapping is coordinated through the IDEFI project FORCCAST (<http://forccast.hypotheses.org>).
2. **As a research method**, CM has supported several national and international research projects:
 - a. MACOSPOL (mapping controversies on science for politics) 2007-09
 - b. MEDEA (mapping environmental debate on adaptation) 2011-14
 - c. EMAPS (electronic maps to assist public science) 2011-14

5. Diffusion / *Dissemination*

Academic publications

Munk, A. K., Meunier, A., & **Venturini, T.** (2016). Data Sprints: A Collaborative Format in Digital Controversy Mapping. In D. Ribes & J. Vertesi (Eds.), *DigitalSTS: A Handbook and Fieldguide* (forthcoming).

Venturini, T., Ricci, D., Mauri, M., Kimbell, L., & Meunier, A. (2015). Designing Controversies and Their Publics. *Design Issues*, 31(3), 74–87. http://doi.org/10.1162/DESI_a_00340

Borra, E., Weltevrede, E., Ciuccarelli, P., Kaltenbrunner, A., Laniado, D., Magni, G., ... **Venturini, T.** (2015). Societal Controversies in Wikipedia Articles. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15* (pp. 193–196). <http://doi.org/10.1145/2702123.2702436>

Borra, E., Weltevrede, E., Ciuccarelli, P., Kaltenbrunner, A., Laniado, D., Magni, G., ... **Venturini, T.** (2014). Contropedia - the analysis and visualization of controversies in Wikipedia articles. In *OpenSym 2014 Proceedings*.

- Venturini, T. (2010). Diving in magma: how to explore controversies with actor-network theory. *Public Understanding of Science*, 19(3), 258–273. <http://doi.org/10.1177/0963662509102694>
- Venturini, T. (2012). Building on faults: how to represent controversies with digital methods. *Public Understanding of Science*, 21(7), 796 – 812. <http://doi.org/10.1177/0963662510387558>
- Venturini, T. (2012). Designing Controversies. In *Design and Displacement, 4S-EASST Conference*.
- Venturini, T. (2008). Piccola introduzione alla cartografia delle controversie. *Etnografia e Ricerca Qualitativa*, 3, 369–394.

Digital publications

The website <http://controverses.sciences-po.fr/archiveindex/> (that I conceived and helped to develop) contains the best examples of controversy atlases realised by the students of some of the universities that employ this method (included the works of my own students at Sciences Po and in Milan).

Fiche 2 : Visual Network Analysis

1. Description de la contribution / *Description of the contribution*

Once limited to the domain of mathematics, networks have recently invaded most disciplines in social and natural sciences. To a large extent, this success is to be credited to the increasing availability of point-and-click software for network analysis. **Taming the complexity of graph mathematics through user-friendly interfaces**, tools such as Pajeck, Ucinet, Guess and Gephi opened network computation to a growing number of scholars. Gephi (gephi.org), in particular, has played a major role in the popularization of networks, thanks to its capacity to represent networks graphically. Translating structural properties, into visual features, Gephi materialized the basics of networks analysis (i.e. clusters, paths, bridges, authorities...) making them accessible to scholars with little mathematical experience.

2. Contribution personnelle de la candidate/du candidat / *Personal contribution of the applicant*

In collaboration with Mathieu Jacomy (the inventor of Gephi), I have worked to develop a visual approach to network analysis. Such approach exploits the virtues of the force-directed algorithms, which spatialize networks simulating a system of physical forces in which nodes repulse each other while edges pulls them together. In the state of equilibrium, the geometric distance between nodes is *roughly* proportional to network distance (see the works of Andreas Noack). Force-directed algorithms can therefore translate visually some of the classic indicators of graph mathematics (communities of nodes tend to cluster together, nodes with high-centrality tend to end up in the centre of the network, nodes with high betweenness tend to be between clusters, etc.)

Drawing on this observation, **I developed a series of techniques to read networks visually**, identifying and describing their structural properties in a language accessible to larger audience of scholars (particularly social scientists who may be reluctant to tackle graph mathematics). My contribution consisted in making explicit an approach that is often employed implicitly, pinpointing its premises, advantages and limitations.

3. Originalité et difficulté / *Originality and difficulty*

I said above that in spatialized networks geometric distance is *roughly* proportional to network distance. Though this has been repeatedly observed, no actual mathematical proof of this proportionality exists. Different force-directed algorithms will spatialize networks in a slightly (and sometime considerably) different way. Even the same algorithm, launched twice on the same network, may produce different results (as the system of forces can stabilise at different local equilibria). This variability generally does not conceal the main clusters and structural holes, but they may have important consequence on the position of individual nodes and sub-clusters (especially when the connectivity is distributed in a homogenous way).

One of my current lines of research addresses this problem, searching help from the mathematical community to find way to **conceptualise and measure the ‘quality’ of network spatialisation**. This will allow to know which algorithm produce the best results on which network and to compute the margin of error in the position of nodes.

4. Validation et impact / *Validation and impact*

While the explicit reflection on visual network analysis is relatively new, its tenets have been implicitly implemented in the Gephi. From the onset and even more now that its development is managed at the Sciences Po médialab, this piece of software has been inspired to the approach of visual network analysis and its success validates the interest of such strategy. Compared to other tools (Pajeck for instance), Gephi is less powerful in the computation of graph metrics, but it offers the most advanced set of tools to spatialize and visualise networks.

5. Diffusion / *Dissemination*

Academic publications

- Venturini, T., Munk, A., & Jacomy, M. (2016). Actor-Network VS Network Analysis VS Digital Networks Are We Talking About the Same Networks? In D. Ribes & J. Vertesi (Eds.), *DigitalSTS: A Handbook and Fieldguide* (forthcoming).
- Venturini, T., Bounegru, L., Jacomy, M., & Gray, J. (2016). How to Tell Stories with Networks: Exploring the Narrative Affordances of Graphs with the Iliad. In *Datafied Society*. Amsterdam: University Press (forthcoming).

- Venturini, T., Jacomy, M., Baneyx, A., & Girard, P. (2016). Hors champs: la multipositionnalité par l'analyse des réseaux. *Rezeaux* (forthcoming).
- Jensen, P., Morini, M., Marton, K., Venturini, T., Vespignani, A., Jacomy, M., ... Fleury, E. (2016). Detecting global bridges in networks. *Network Science*.
- Venturini, T., Jacomy, M., & Carvalho Pereira, D. (2015). *Visual Network Analysis*. Paris. Sciences Po médialab working papers.
- Severo, M., & Venturini, T. (2015). Intangible cultural heritage webs: Comparing national networks with digital methods. *New Media & Society*, 1–20. <http://doi.org/10.1177/1461444814567981>
- Venturini, T. (2015). Review of “An Aesthesis of Networks: Conjunctive Experience in Art and Technology, by Anna Munster.” *Digital Studies / Le Champs Numerique*.
- Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS One*, 9(6), e98679. <http://doi.org/10.1371/journal.pone.0098679>

Digital publications

Besides the academic publications listed above, the techniques of visual network analysis have also been implemented in an online tool developed by the Sciences Po médialab and called **Manylines** (tools.medialab.sciences-po.fr/manylines) [self-assessment: A-3; SO-4; SM-2; EM-3; SDL-4; OC-da-2/cd-0/ms-0/tpm-3]. Manylines allows users to upload their networks, refine their visual rendering (in terms of node position, size and colour) and then to creating stories or presentations aligning different dynamic screenshot of the network.

Fiche 3 : Sciences Po médialab quali-quantitative tools

1. Description de la contribution / *Description of the contribution*

One of the main occupation of the médialab (the laboratory that I founded and coordinated for six years) has been to develop a series of digital tools for social research. This effort has not been limited to the technical development, but also involved a deep reflection on the theory of methods. In different ways, all our tools aim at **overcoming the divide, customary in the social sciences, between qualitative and quantitative methods**.

Accustomed to paying a high price for every fragment of information, social scientists traditionally faced a trade-off between gathering superficial data on large populations (with quantitative methods such as surveys and questionnaires) or a rich information on few actors (with qualitative methods - interviews, ethnographic observations). Thanks to digital traceability, it is now possible to obtain rich information on large populations at a reasonable price. The mission of the médialab has been to explore the new digital datasets and develop tools to navigate from macro patterns to tiny details without overlooking all that lies in-between.

Here is a short description of five of the main tools developed at the médialab (a fifth one, Manylines, is described just above in the fiche 2 on “visual network analysis”)

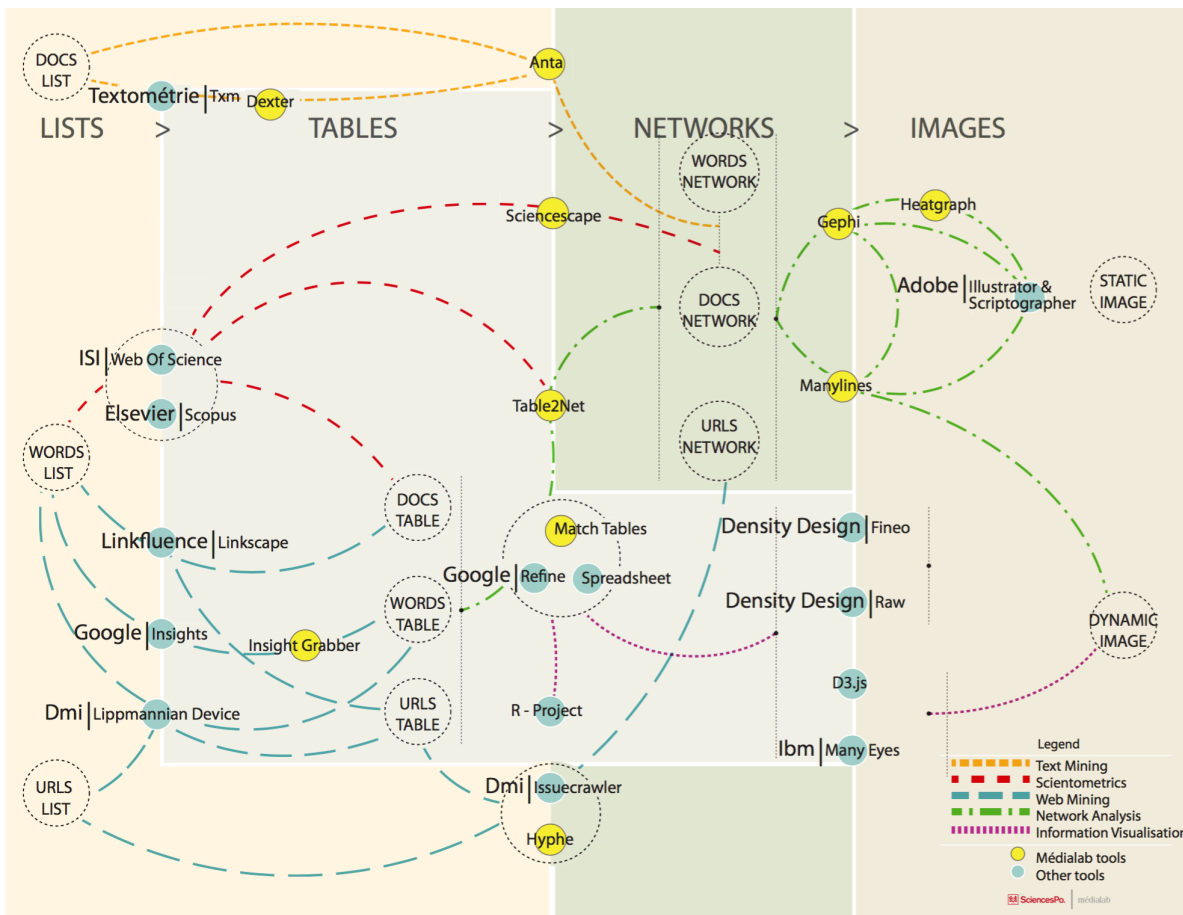
Gephi (<https://gephi.github.io/>) [self-assessment: A-5; SO-3; SM-4; EM-4; SDL-4; OC-da-2/cd-0/ms-0/tpm-3]. Gephi is the reference software for visualizing and exploring all kinds of networks and complex systems, dynamic and hierarchical graphs. Runs on Windows, Linux and Mac OS X. Gephi is open-source and free.

Table2Net (<http://tools.medialab.sciences-po.fr/table2net/>) [self-assessment: A-3; SO-3; SM-2; EM-2; SDL-; OC-da-2/cd-0/ms-0/tpm-3]. Table2Net is a simple but powerful javascript tool that allows to extract networks from data in a tabular form. The tool allows to chose which columns contain information on the node, which information of the edges and to export different types of graph (mono-partite, bi-partite, oriented, weighted, dynamic...).

Hyphe (<http://tools.medialab.sciences-po.fr/#hyphe>) [self-assessment: A-4; SO-4; SM-3; EM-4; SDL-4; OC-da-2/cd-0/ms-0/tpm-3]. Hyphe aims at providing a tool to crawl data from the web to generate networks between what we call WebEntities, which can be singles pages as well as a website or a combination of such.

Sciencescape (<http://tools.medialab.sciences-po.fr/sciencescape/>) [self-assessment: A-3; SO-3; SM-2; EM-3; SDL-4; OC-da-2/cd-0/ms-0/tpm-3]. Sciencescape allow users to analyse bibliographical information extracted from Scopus or ISI Web of Science. It produced temporal and network diagrams of keywords, authors, journals and other scientometrics information.

ANTA (<http://github.com/medialab/ANTA>) [self-assessment: A-2; SO-3; SM-2; EM-2; SDL-4; OC-da-2/cd-0/ms-0/tpm-3]. ANTA (or actor-network text analyser) is a text analysis tool, allowing to collect and curate corpora of documents, to extract and select the noun-phrases contained in them and to generate bi-partite networks of documents and noun-phrases.



2. Contribution personnelle de la candidate/du candidat / *Personal contribution of the applicant*

My contribution in the development of the médialab tools has been double.

- On the one hand, my role has been to make sure that all the tools that we developed stick to the **quali-quantitative approach** described above. This entailed making explicit the tenets of this approach and working with developers to make sure that this conceptual inspiration is not lost because of technical constraints.
- On the other hand, I have served as the **point of contact with the scholars of Sciences Po**, making sure that tools remained close to their research needs. This involved making sure that the interfaces were ‘friendly’ enough, but also making sure that the methodological implications of the tools were clear and coherent.

3. Originalité et difficulté / *Originality and difficulty*

As the image presented above illustrates, the tools developed at the médialab constitute not an accidental ensemble, but **an integrated ‘ecosystem’** in which the output of one tool become the input of another one. The colours of the lines in the image hint to the fact that tools are not supposed to be used in isolation, but are part of different ‘methodological lines’. Assuring this global **coherence and interoperability** has been my responsibility for the six years I have spent at the médialab.

4. Validation et impact / *Validation and impact*

The following table presents the visibility of the different websites maintained by the médialab and provides an overview of their impact

Site web	Sessions in 2015	Unique users in 2015	Viewed pages in 2015
Climaps (EMAPS)	5 230	3 166	13 103
Hyphe tool	2 542	680	53 201
Hyphe presentation	3 641	2 009	45 949
Manylines	1 355	979	1 997
MEDEA	263	134	1 139
Medialab Tools	88 393	61 187	148 703
Médialab SITE	31 296	22 857	85 203
SigmaJS (lib graphes)	202 626	148 561	269 242
Modes of Existence	34 510	19 587	96 007
La Fabrique de la Loi	16 021	13 481	30 950

5. Diffusion / *Dissemination*

Academic publications

- Jacomy, M., Girard, P., Ooghe-Tabanou, B., & Venturini, T. (2016). Hyphe, a Curation-Oriented Approach to Web Crawling for the Social Sciences. In *ICWSM*.
- Venturini, T., Cardon, D., & Cointet, J.-P. (2015). Méthodes digitales: Approches quali/quantitative des données numériques - Curation and Presentation of the Special Issue. *Réseaux*, 188, 9–21. <http://doi.org/10.3917/res.188.0009>
- Venturini, T., Jensen, P., & Latour, B. (2015). Fill in the Gap: A New Alliance for Social and Natural Sciences. *Journal of Artificial Societies and Social Simulation*, 18(2), 11. Retrieved from <http://jasss.soc.surrey.ac.uk/18/2/11.html>
- Latour, B., Jensen, P., Venturini, T., Grauwin, S., & Boullier, D. (2012). “The whole is always smaller than its parts”: a digital test of Gabriel Tarde’s monads. *The British Journal of Sociology*, 63(4), 590–615. <http://doi.org/10.1111/j.1468-4446.2012.01428.x>
- Venturini, T. (2012). Great expectations: méthodes quali-quantitative et analyse des réseaux sociaux. In J.-P. Fourmentraux (Ed.), *L'Ère Post-Media. Humanités digitales et Cultures numériques* (Vol. 104, pp. 39–51). Paris: Hermann.
- Venturini, T., & Guido, D. (2012). Once Upon a Text : an ANT Tale in Text Analysis. *Sociologica*, 3. <http://doi.org/10.2383/72700>
- Venturini, T. (2012). Médialab de Sciences Po : cartographier le web pour les sciences sociales. Retrieved from <http://www.ina-sup.com/node/2842>
- Venturini, T., & Latour, B. (2010). The Social Fabric: Digital Traces and Quali-quantitative Methods. In *Proceedings of Future En Seine 2009* (pp. 87–101). Paris: Editions Future en Seine.

Digital publications

The website tools.medialab.sciences-po.fr gathers and presents all the tools published by the Sciences Po médialab. The code source and the technical documentation of all these tools (and of others still under development) is available on Github at github.com/medialab where the médialab tools are released under an open source licence.

Fiche 4 : Environmental Debates Mapping

1. Description de la contribution / *Description of the contribution*

As it should be clear from the three contributions described above, my research focuses on harnessing digital data and computational advances and to turn them in tools and methods for social research. To make sure that these tools and methods deliver the innovation they promise, I have myself be among their first alpha-users and I have used them extensively for the analysis of public debates on the environments.

Beside the political reasons that encourage me to privilege this domain, **environmental debates are interesting because they challenge the traditional categories of social sciences**. The need to take into consideration a variety of elements (who would have imagined that CO2 would become one of the main actors of international diplomacy? Who would have imagined that biodiversity would become as precious as fossil fuels?) promotes a radical renewal of the study of collective phenomena and demands new analytical tools.

2. Contribution personnelle de la candidate/du candidat / *Personal contribution of the applicant*

I have already mentioned the projects EMAPS (www.climaps.eu), MEDEA (medea.medialab.sciences-po.fr) and ClimateS (www.climatenegotiations.org). Through these projects, I have tried to demonstrate that controversy mapping, visual network analysis and, more generally, digital quali-quantitative methods could be fruitfully applied to investigate debates on climate change. These three projects (very different in size and partners involved) have been entirely under my responsibility. **I conceived their project design, wrote their funding proposals, gathered their research consortium, assured their leadership, managed their budget and reporting.**

3. Originalité et difficulté / *Originality and difficulty*

Investigating climate debates through digital methods has proved to be challenging beyond expectation as it required to apply a set of experimental methods to a relatively unsettled object of study. This explains the highly tentative natures of the results obtained. My projects are far from having exhausted the research questions that they raised and this represents a strong push to continue my investigations in this areas.

The double originality of this line of research (in methods and in object) also explains why, in the projects above, we have experimented an **innovative research format that we named “data-sprint”**. Explicitly inspired to open source events, data-sprints gather in the same physical location scholars coming from different disciplines (sociology, data sciences, design) and the representatives of the potential users (in the case of my projects, diplomats, activists, practitioners, journalists...). Unlike hackathons and barcamps, however, data-sprints involve a more extensive preparation to assure the achievement of actionable results.

4. Validation et impact / *Validation and impact*

It is (too) often said that social data is the wealth of the digital revolution: “data is the new oil”. Maybe, but if so it is certainly an ‘unconventional oil’ – that can only be exploited through a hard work of extraction and cleaning. My projects were design to **demonstrate the interest of digital methods, but also the conditions of their use**

beyond 'big data' hype. In particular, our experience has shown that the exploitation of digital traces demands:

1. A patient work of collecting and processing digital traces and turn them into in research data.
2. An exchange with the potential users to select the relevant data and analytical techniques.
3. A close cooperation between scholars from different disciplines: media studies (to identify the most interesting traces and recognize their bias); data sciences (to manage large amounts of data without losing quality analysis); information design (to make the data accessible to a wider audience).

EMAPS highlighted these three elements and integrated them in a research protocol (the 'data-sprint') that is now widely used at the médialab and has begun to spread to other projects and research initiatives.

5. Diffusion / Dissemination

Academic publications

- Venturini, T., Munk, A., & Meunier, A.** (2016). *Data-Sprint: a Public Approach to Digital Research*. (C. Lury, P. Clough, M. Michael, R. Fensham, S. Lammes, A. Last, & E. Uprichard, Eds.) *Interdisciplinary Research Methods (forthcoming)*.
- Venturini, T., Baya Laffite, N., Cointet, J.-P., Gray, I., Zabban, V., & De Pryck, K.** (2014). Three maps and three misunderstandings: A digital mapping of climate diplomacy. *Big Data & Society, 1*(2). <http://doi.org/10.1177/2053951714543804>
- Venturini, T., Meunier, A., Munk, A. K., Borra, E. K., Rieder, B., Mauri, M., ... Laniado, D.** (2014). Climaps by Emaps in 2 Pages (A Summary for Policy Makers and Busy People in general). *Social Science Research Network*, (ID 2532946). Retrieved from <http://papers.ssrn.com/abstract=2532946>
- Venturini, T., Gemenne, F., & Severo, M.** (2013). Des Migrants et des Mots. Une analyse numérique des débats médiatiques sur les migrations et l'environnement. *Cultures & Conflits, 88*(4).
- Venturini, T.** (2007). Verba Volant, Scripta Manent: The Discontinuity Effect of Explicit Media. *American Behavioral Scientist, 50*(7), 879–896. <http://doi.org/10.1177/0002764206298313>
- Venturini, T.** (2007). Les trous noirs de la Révolution Verte. *Décroissance & Technique, 3*.
- Venturini, T.** (2005). Communities VS Networks, from implicit to explicit technical knowledge. In *Conference on Communities and Technologies*.

Digital publications

[Climaps.eu](http://climaps.eu) is an online atlas providing data, visualizations and commentaries about climate adaptation debate. It contains 33 issue-maps. Each of the maps focuses on one issue in the adaptation debate and provides:

- an interactive visualization;
- a discussion of the map and the findings that it discloses;
- a description of the protocol through which the map has been created;
- the raw and the cleaned data on which the map is based and the code employed to treat them.

Climaps.eu also contains 5 issue-stories guiding the users in the combined reading of several maps.

The atlas is addressed to climate experts (negotiators, NGOs and companies concerned by global warming, journalists...) and to citizens willing to engage with the issues of climate adaptation. It employs advanced digital methods to deploy the complexity of the issues related to climate adaptation and information design to make this complexity legible.

[Medea.medialab.sciences-po.fr](http://medea.medialab.sciences-po.fr) aims to investigate scientific and diplomatic discussions on climate change through two datasets specifically built for this project.

1. The network of IPCC experts, which was visualized to reveal:
 - the geographic diversity of the IPCC (particularly the participation of Southern World countries);
 - the special role played by the IPCC authors affiliated with French institutions;
 - the 'bridge authors' who, participating to multiple working groups, assure the IPCC integration.
2. The discussion in the UN Framework Convention on Climate Change as reported by Earth Negotiation Bulletin (www.iisd.ca/vol12/), which was visualized to reveal:
 - the comparison between the discussions on mitigation and on adaptation within the UNFCCC;
 - the evolution of themes and actors of climate negotiations;
 - the matrix of the relative commitment of different countries on different tables negotiations.