The e-Diasporas Atlas* is a unique experiment in research on diasporas and publishing, a first in the restitution of scientific findings and their presentation.

Historically, the emergence of e-diasporas occurred in tandem with the diffusion of the Internet and the development of multiple online public services. At the end of the 1990s, a number of institutions joined forces with the new “e”-technologies (e-administration, e-democracy, e-education, e-healthcare, e-culture, e-tourism), which gave rise to the first presence on the Web of associations run by migrant populations. If the earliest websites were produced by IT professionals, we soon saw the diffusion of the Web in all of the diasporic communities and at all levels. The last ten years witnessed, in these communities, the use of both Webs 1.0 and 2.0 (blogs), as well as the widespread appropriation of the various social-networking platforms (Facebook, Twitter, LinkedIn, etc.).

These new communication and organization practices produce a vast, moving e-corpus whose exploration, analysis, and archiving have never before been attempted. The outcome of the efforts of more than 80 researchers worldwide, the e-Diaspora Atlas is the first of its kind, with some 8,000 migrant websites archived and observed in their interactions.

A migrant site is a website created or managed by migrants and/or that deals with them (at any rate, a site for which migration or diaporas is a defining theme). This could be a personal site or blog, the site of an association, a portal/forum, an
institutional site, or anything similar. Usage is not the criterion: a site often consulted by migrants (a media site, for example) is not necessarily a migrant site. What distinguishes ‘activity’ is first and foremost the production of content and practice of citation (hyperlinks). On the other hand, a migrant site need not necessarily be situated in a foreign country and may just as easily be in the country of origin. Migrant sites testify to a given e–diaspora’s occupation of the web.

What we call e-diaspora is a migrant collectivity as it organises itself and acts particularly on the web, and whose practices are those of a community whose interactions are ‘enhanced’ by digital exchange. An e–diaspora is a dispersed collectivity, a heterogeneous entity whose existence rest on an elaboration of a common direction, a direction not defined once and for all but which is constantly renegotiated throughout the evolution of the collective. An e-diaspora is a unstable collectivity because redrawn by every newcomer. She can retract or increase not by exclusion or inclusion but by membership or resignation.

It is both ‘online’ and ‘offline’, so what interests us is both the digital ‘translations’ of ‘physical’ actors/phomena (the online activities of associations for example) and the specifically (‘natively’) digital actors/phomena (e. g., a forum and its internal interactions), what are sometimes called pure players. The question of ‘rub–offs’—reciprocal influence between these two sorts of web entity—is of capital importance in analysis of an e–diaspora. It's thus clear that research carried out in the context of the e–Diasporas Atlas presupposes a knowledge of the diaspora in question and, based on exploration of the web, calls on new research in the field. It also implies a knowledge of the web and an appreciation of the singularity of the exchanges that take place there.

Furthermore, we prefer the term e–diaspora to that of ‘digital diaspora’ since the latter runs the risk of becoming a source of confusion given the increasingly frequent use of the notions of ‘digital native’ and ‘digital immigrant’ in a ‘generational’ sense (distinguishing those born before from those born during/after the digital era). The object of the e–Diasporas Atlas is not this ‘digital migrant’ but the connected migrant.
e-Diasporas Methodology

The digital methodological chain and the tools we developed for building the e-Diasporas Atlas which aims at mapping and analyzing the occupation of the web by diasporas. Such a chain is composed of four intertwined steps: 1) equipped web exploration and corpus building; 2) data enrichment (location, languages, text-mining); 3) network visualization-manipulation and graph interpretation; 4) collaborative sharing of (raw) data and findings.

Step 1: Web Exploration

In the making of a chapter of the Atlas, the first step is the building (and circumscription) of a corpus of websites. The researcher plays a crucial role in this process inasmuch as his knowledge of the fieldwork allows him to discriminate the relevant resources for a given diaspora.

The constitution of a corpus of websites is the method used to ‘capture’ an e-diaspora. A question of breakdown and selection that allows extraction of a diaspora web, it is also a task of definition in that a e-diaspora presents itself to a researcher only as a product of this ‘excision’ performed upon the web. Similarly, it is only because of such exploration/selection, this filtering/circumscription of a corpus, that what a migrant site actually is takes on meaning.

In order to complete this stage of collection, the researcher needs to be equipped. The identification of relevant websites is achieved semi-automatically thanks to a software called Navicrawler that allows to scan web grounds through a web-browser. Navicrawler is a Firefox add-on, designed and developed by Mathieu Jacomy. The interface is located on the left of the currently browsed page.

The essential functionality of the Navicrawler consists in scraping the out-links of the visited websites (listed and stored as “Next Sites”) For each website, the researcher can incorporate it to the corpus (it becomes “In Site”) or reject it (“Out Site”). He can also describe the websites by adding tags.

The logic of exploration induced by the Navicrawler is located at the crossroads of browsing and crawling. Unlike automatic crawling, it allows the researcher to perceive the context of links and thus to avoid a blackbox effect. At the end of this exploration stage, he is able to export his corpus as a graph in which the nodes
represent the websites and the edges stand for the links between them.

**Step 2: Data Enrichment (Digital Toolbox)**

As already mentioned, the social scientist plays a central role all along the process of corpus building and description/enrichment. However, he can be assisted in the content analysis by automatic tools. Our research team developed a digital toolbox that renders possible various “enrichment processes”, among which:

- Retrieving from a list of urls the information provided by the registrar (ICANN); about the registrant (owner of the domain name), especially his *geographical location*; about the server hosting the website, etc..
- Text-mining applied on the index of the corpus in order to retrieve *named entities*: persons, organizations, places, etc. (using Open Calais API).
- Recognition of the *languages* used in each website (and the distribution of languages in order to study multilinguism, an important issue in migration studies).

**Step 3: Network Visualization (Gephi)**

In order to visualize the exploration data, in other words to map the corpus previously built, we use a software of graph visualization named *Gephi*, a project initiated and hosted at its beginnings by our research team. It was developed by Matthieu Bastian, Matthieu Jacomy and Sébastian Heymann. This tool allows to spatialize and manipulate the corpus network. Two types of visualization are available:
- a spatialization based on the physical principle of attraction/repulsion (according to the presence or absence of a link between two nodes)
- a geographical spatialization that uses geocoded data: location of websites’ owner, websites’ addressees, servers, etc. (especially information retrieved during the data enrichment stage)
maps.e-diasporas.fr is a collaborative platform first developed and implemented by Mathieu Jacomy and the ICT Migrations team, for hosting the e-Diasporas Atlas. It is a tool for publishing and sharing research findings among scientific communities. It is composed of chapters (in our case, the various diasporas) and provides for each of them the following data:

- **Maps**: browsable graphs of the corpus, with different views according to the fields of classification.
- **Raw data**: the empirical data (texts, videos, images interviews, etc.) produced/retrieved and used during the research. The Digital Diasporas Atlas is part of the more general “digital humanities” project of providing access and diffusing not only the research results but also the research data.
- **Statistics**: they are automatically generated from both the classification and the graph structure and provide quantitative data about the relations between categories/actors. Statistics help strengthening the hypothesis formulated from the graph visualization.
findings

( more to read on e-diasporas.fr)

The e-Diasporas Atlas proposes at least two keys of interpretation (performance):

1. One is topological, centered on the analysis of the connectivity between the actors who bustle on the Web and

2. the other one is quanti-qualitative, and brings information stemming from the exploration of the contents of every site and from the confrontation with the fieldwork and the expertise of all researchers involved in this project.

The network is the very image of a diaspora. In its web-like structure can be read the spatial distribution of a geographically dispersed collective and its remote organization.

The e-Diasporas Atlas is above all a basic inventory of the networks on the Web; its scale is large and its topology complex, represented in the form of graphs. The comparison between geographical networks and networks observed on the Web seemed an obvious approach and gave rise to a few recurring observations, among which:

1. A large majority of e-diaspora sites are geolocated in North America and especially in the United States. This predominance is even more surprising when the presence on the Web strongly contrasts with the presence in the geographical areas, as was observed for Chinese, Palestinian and Nepalese diasporas (among others).

2. Wishing to understand the collective life of the e-diasporas, we sought to identify clusters of expatriates on the Web and the exceptions to these clusters, their centrality, their hierarchies, their relations and their assemblage. Although the place of origin appears as a reference on the Web, even if it does not act as a hub (cf. the Palestinian diaspora), it reinjects geography into the Web.
3. The auto-organization typical of Web networks facilitates the emergence of decentralized communities and acts as an ideal platform or vehicle for a transnational collective imaginary. This is the case, for instance, of the Huyghurs, the Hmong or the Nepalis.

4. Mapping the e-diasporas enables us to analyze, among other things, the relations entertained by various diaspora actors with their homeland: clearly labeled links with a strong state, as in the case of France, one-way to non-existent links, as in the case of Morocco, or Mexico for instance.

5. As for the links between actors within an e-diaspora, the e-Diasporas Atlas gives a glimpse, before going into further detail, of associations unaware of one another, of bloggers creating their own world, of activist groups or individuals seizing power on the Web, and sometimes, as in the case of the Arab springs, even managing to spark popular dissent and to impact on political events.

6. India, with its many and widely dispersed communities, was our laboratory of choice. Far from having exhausted the exploration of the Indian e-diaspora, a thorough search organized around regional and religious components resulted in nine sections, among which an effort to gain an overall picture.

The first thing we observe is the density and duration of the association sites, some of which go back more than 20 years. Next, a non-negligible visibility of the government sites, which seek to attract the most privileged and influential migrants.

Highly visible as well are two opposing activist movements: on the one hand, the nationalist religious hindutva movement, ideologically attached to the territory of India, institutionally installed in the United States and run by a conservative, male elite; on the other hand, a militant movement that rejects nationalism in favor of an all-South-Asian regional identity based on defense of the interests of all minorities marginalized because of caste, gender, sexual orientation, religion or ethnic belonging, whatever their geographical location.
The restitution of scientific findings and their presentation

The e-Diaspora Atlas is a hybrid product combining two readings:

The maps of the presence of the different e-diasporas on the Web at a time T make up a paper atlas that is in turn augmented by a smartphone application and a website.

The Atlas can be used in a number of ways:

- The paper version is composed of three composite maps, which introduce the reader to graphic representations of some thirty e-diasporas.

- An application – which can be downloaded free from the App Store – enables readers equipped with an iPhone to access articles and videos by the researchers concerned via the graph appearing with each diaspora. This additional pedagogical material expands on interpretation of the moments T frozen on the paper page.

- This application also provides a link to a third and even richer level of reading: a companion website provides a space allowing the user to appreciate the full dynamic evolution of the corpuses and visualize them in their entirety. This digital restitution makes it possible, in particular, to access personally the e-Diasporas research tools and platform, to manipulate the maps, to cross and compare the data and to move about within the graphs.


*BRETONS, CHINESE, EGYPTIANS, FRENCH EXPATS, FRENCH RE-PATS, TURKS, HMONG, INDIANS-ALL INDIANS-GENERAL, INDIANS-HINDUTVA, INDIANS-REAL ESTATE, ITALIANS, JEWISH DIASPORA, KERALA, LEBANESE, MACEDONIANS, MEXICANS, MAROCCANS, NEPALI,
UYGHUR, PALESTINIANS1, PALESTINAINS2, RUSSOPHONES, SIKHS, SOUTH-ASIAN LEFTIST MOBILIZATION, TAMILS, TUNISIANS, YOGOSPHERE, ZOROASTRIANS