

# Cartographier le Web a-t-il un sens ?

Par **Alexandre MONNIN**

Philosophe, chercheur Inria (équipe Wimmics),  
membre du réseau d'experts de la mission Etalab

En conférence le 15 décembre

**Longtemps, le Web et les technologies numériques ont été assimilés à un espace qualitativement différent de l'espace géophysique arpenté par les géographes. Le « cyberspace », comme son nom l'indique, n'en demeurerait pas moins une forme d'espace. En ce sens, la spatialité endossait le rôle d'opérateur susceptible de suturer l'écart entre « réel » et « virtuel ». Pour autant, cette réflexion laissait une question essentielle en suspens : à quelle « espèce d'espace » revenait-il d'unifier des pans en apparence aussi hétérogènes de la réalité ?**

La réponse à cette question exige d'en passer par la médiation de la cartographie qui, seule, permet d'échafauder un espace de représentations commensurable où les cartes du Web s'insèrent au sein d'un corpus très riche. Bien sûr, toute carte est moins la représentation fidèle d'un espace qu'une opération de mise en visibilité d'une étendue censée répondre à certaines contraintes cognitives, dans le but de faciliter des tâches aussi diverses que la navigation ou la planification. La carte n'est donc en aucun cas le décalque fidèle (avant réduction) du territoire. Cartographier le Web présente néanmoins un défi spécifique. Les cartes du Web associent fréquemment deux types de représentations spatiales : une carte du monde à plat (dont on sait qu'elle procède d'un compromis plus visuel que sémantique, destiné à assurer le passage d'une représentation sphérique du globe à son équivalent sous forme plane) et une carte qui épouse les contours d'un graphe. Notons que les espaces mis en jeu reposent sur des « axiomatiques »<sup>1</sup> très différentes : contiguïté pour les surfaces, connexité s'agissant du graphe. Bien sûr, une réalité aussi complexe que le Web pourrait, sans contredit, articuler ces deux dimensions. Toutefois, cette hétérogénéité oblige à s'interroger sur ce qui est ainsi représenté : Internet et ses réseaux de câbles, de routeurs, de serveur, ou le Web, qui constitue une couche applicative d'Internet (bien que, de ce point de vue, ses frontières soient floues par ailleurs). Le premier présente une architecture de réseaux territorialisés, d'où l'articulation de deux modalités de représentation. Seulement, à l'instar de l'économie du numérique qui repose sur les contenus du Web, ce sont également ces derniers qui demandent à être cartographiés. Or, le Web, à bien des égards, s'apparente à un espace foncièrement *déterritorialisé*.

Examinons quelques-unes de ses caractéristiques. On nomme « architecture du Web » la couche de standards jugés les plus fondamentaux. Parmi eux, le plus important n'est autre que

le système de nommage à base d'URI (*Uniform Resource Identifiers*, généralement connues sous le nom d'URL, un acronyme abandonné dans les standards depuis 1998), le protocole HTTP (*HyperText Transfer Protocol*) et les langages tels que HTML ou RDF. L'évocation de ces quelques éléments, stabilisés au fil d'un processus de standardisation entamé à la suite de la naissance du W3C (*World Wide Web Consortium*) en 1994, est marquée par une mise à distance des éléments spatiaux au cœur de ces standards. Ainsi, à la différence du protocole Internet TCP/IP, protocole dit de *transport*, le protocole HTTP est un protocole de *transfert*. Des éléments concrets, des « données », sont *transportés* par le réseau Internet, mais, on va le voir, ce sont avant tout des *représentations* qui sont transférées sur le Web, représentations qui sont référées à ce qu'elles représentent en un sens tout à fait spécifique. Plus généralement, sur le Web, les éléments techniques présents sur un serveur (bases de données, fichiers, etc. - pour les cas les plus simples) constituent moins la *cible* que le *support* des interactions. À cela s'ajoute le fait d'appréhender le système de nommage du Web non comme un ensemble d'adresses (URL) mais d'identifiants (URI) qui désignent des objets en tous genres, « accessibles » comme « inaccessibles », articulant ainsi une *géographie de l'identité* de ces objets. Du point de vue de l'architecture du Web, il ne s'agit pas de télécharger des fichiers dont on suivrait l'évolution à la manière des systèmes de *versioning*, tels que Github. Outre que les services ou les applications (aux résultats constamment changeants) sont identifiables et adressables sur le Web, et ce depuis l'origine, un tel système de suivi des versions fait défaut. En réalité, l'échelle pertinente pour comprendre ce qui est identifié sur le Web n'est pas celle du fichier ou du document, mais de la « ressource ». Ce concept, dont le sens n'a véritablement été élucidé qu'avec le travail de Roy Fielding sur le style d'architecture REST (pour *REpresentation State Transfer*) désigne « la sémantique de ce que [l'on] entend identifier ». Autrement dit, non telle ou telle itération de la page d'accueil du journal *Le Monde* par exemple, ni tel ou tel fichier photo représentant Tim Berners-Lee, mais « la page d'accueil du *Monde* »

<sup>1</sup> Michel Lussault.

ou « la photo de Tim Berners-Lee » (voire Tim Berners-Lee lui-même !). Le contenu de la ressource ou la manière de l'appréhender sont ainsi laissés entièrement libres, aussi n'est-il pas rare de croiser des ressources s'apparentant à des déictiques, telles que « mon compte Gmail ». Il appartient en effet à l'organisme qui les identifie de publier les contenus en adéquation avec les ressources. Cette adéquation se marque dans le temps (la page d'accueil du *Monde* varie constamment) et ponctuellement (en fonction des modalités selon lesquelles les représentations des ressources sont générées : langages, formats techniques, paramètres d'accessibilité, etc.). Les ressources qualifient leurs représentations, qui sont elles-mêmes générées pour les satisfaire de manière régulière, au double sens normatif et temporel du terme. En s'engageant à donner accès à des contenus pertinents, les organismes amenés à gérer des noms de domaines créent une structure d'attente dont on ne sait, *a priori*, de quelle manière elle sera satisfaite (personne ne peut prévoir à quoi ressemblera précisément la page d'accueil du *Monde* dans trois jours). Fielding compare ainsi les ressources, comme principe d'attente, de génération et de coordination des représentations transférées, à des « ombres » ou des « concepts » : inaccessibles en tant que tels, bien que manipulés par l'intermédiaire de leurs représentations.

Deux images sont ainsi mises à distance : celle de la grande bibliothèque et celle de l'hypertexte. En multipliant le nombre de représentations satisfaisant une même ressource, le Web rompt avec la centralité reconnue au document comme point fixe des environnements documentaires. D'ailleurs, c'est bien dans le but de faire place à la fixité du document dans l'univers bibliothéconomique que furent imaginés les URN (*Uniform Resource Names*), ces identifiants de contenus documentaires stables, destinés à fournir le complément pérenne des URL. C'est dire que le Web n'a jamais été une réalité statique. Loin de s'imposer d'elle-même, il a fallu faire sa place à l'idée de grande bibliothèque. Autre idée battue en brèche, celle selon laquelle le Web serait un hypertexte. Sans rentrer dans les détails, notons simplement que l'hypertexte suppose de relier des pages (ou « lexies ») au moyen de liens. Mais les liens, sur le Web, ne sont pas des réalités premières mais dérivées. Leur existence dépend de pointeurs (les URI) qui assurent (éventuellement) une fonction d'accès. Or, dans le cas contraire, avec l'erreur 404 par exemple, le système ne s'effondre pas. On comprend dès lors qu'il n'est nullement besoin de se référer à un index général cataloguant les différentes « pages » du Web avant

de publier un contenu. Il n'existe en effet aucun organisme gérant le moindre index de ce type dont on tirerait ensuite de quoi établir une carte du Web. L'octroi d'une permission n'est pas nécessaire : il suffit de gérer un nom de domaine et des moyens techniques adéquats pour publier les contenus que l'on souhaite. L'indirection entre URI et représentations (la ressource s'interposant entre les deux) procède de ce fait d'un choix fondamental en terme de design qui assure au Web son caractère décentralisé.

Qui plus est, les contenus générés sur le Web sont très largement dépendants des interactions (requêtes à partir d'une URI). Comme l'indiquait Sébastien Heymann, l'un des créateurs du logiciel *Gephi* de représentation de graphes : « les liens hypertextes parcourus sont des réalisations (...) c'est-à-dire des occurrences d'un possible prédéfini »<sup>2</sup>. On le comprend, le Web n'est pas indexable en tant que tel, sans médiation. Les robots de *Google* eux-mêmes interagissent à partir des URI et le résultat de ces interactions est capturé et canonisé dans l'index du moteur : voici comment naissent les « pages » Web. L'opération de cartographie comme d'indexation opère en réalité une substitution d'une réalité à une autre. Il s'agit donc moins de représenter fidèlement son objet que d'agencer des réalités hétérogènes, aboutissant à simplifier en vue de faciliter certains usages (fortement orientés, on le comprendra aisément). Paradoxalement, c'est en ce sens que la cartographie du Web apparaît bien comme une « authentique » cartographie : il n'y a pas lieu, en effet, de nier l'épaisseur des opérations nécessaires à sa réalisation. On ne dira plus : « la carte n'est pas le territoire », mais « il n'y a pas de territoire avant sa cartographie ». Ce qui ne revient pas à dire qu'il n'y a rien. En revanche, ce qu'il y a, « sur le Web », ce sont avant tout des réseaux d'interactions multiples que l'on pourrait représenter de bien des manières, sans consentir de privilège particulier aux cartes.

Notons cependant pour finir que le Web, du fait de ce principe d'indirection, qui place la ressource au cœur de son architecture, présente néanmoins un point de similitude fondamental avec l'espace. Michel Lussault place la distance au cœur de la spatialité. Distance qu'il appréhende avant tout comme un écart que les hommes et les sociétés s'efforcent de combler. À l'échelle du Web, la distance possède une charge plus positive : elle est l'attente sise au cœur de son

<sup>2</sup> *Du Web à l'idée du Web : conception d'outils pour les sciences humaines*, S. Heymann, 2008.

architecture, le ressort d'où il tire son caractère décentralisé. L'accès, nous dit Lussault, euphémise la distance et la séparation. Sans doute. La leçon, au cœur du Web, est cependant plus fine encore. Car si l'accès, ou plutôt l'actualisation d'une ressource pensée comme virtualité, ne la rend pas présente, les interactions quotidiennes avec les « objets physiques » ne nous y donnent pas davantage accès. À moins de s'en tenir à une intelligence de l'espace séparée du temps, aujourd'hui obsolète, les objets occupent un espace-temps dont une coupe temporelle n'est, précisément, qu'une coupe. Qu'il s'agisse de serrer la main de quelqu'un ou d'accéder à une représentation de la page d'accueil du *Monde*, dans les deux cas, les objets ne se livrent pas *in toto*. La place qu'occupe ponctuellement un objet ne le contient pas intégralement. Pour penser l'objet, encore faut-il penser une trajectoire virtuelle mais aussi son extension : à savoir tout ce qui lui permet d'être tel qu'il est. Lussault le note lui-même : le lieu est hors de soi (il « ek-siste »), à l'instar des images ou des paroles qui s'y rapportent : « une part de ce qui compose [le lieu] existe hors de lui »<sup>3</sup>. Il en va de même avec les objets, telle est la leçon que nous livre l'architecture du Web. On ne peut tout simplement parler d'objet sans poser une forme de distance permettant d'articuler ce qui s'actualise et se présente et, de ce fait, en vient à combler, partiellement, cette distance. Cela même qui motive d'assembler, d'articuler et de qualifier ces différents éléments d'une manière « régulière » constitue l'objet. Autrement dit, comprendre le Web exige de dépasser ce que le philosophe anglais Alfred North Whitehead nommait le principe de localisation simple, à savoir la thèse selon laquelle « Dire qu'un élément matériel a une localisation simple signifie que – en exprimant ses relations spatio-temporelles – il est approprié d'affirmer qu'il est là où il se trouve, en une région définie de l'espace, et pendant une durée définie, en dehors de toute référence essentielle aux relations de cet élément matériel à d'autres régions de l'espace et à d'autres durées<sup>4</sup> ».

<sup>3</sup> *L'homme spatial*, Michel Lussault, éd. Seuil, 2007, p. 99.

<sup>4</sup> *La Science et le monde moderne*, Alfred North Whitehead, éd. du Rocher, 1994.

