

Espaces collaboratifs, métaphores graphiques et exploration visuelle.

La visualisation d'information au service de l'intelligence collective.

Guy Melançon, LIRMM UMR 5506

Nous vivons dans une société où l'information est abondante. En 2001, le projet « How much information » de l'université de Berkeley rapportait que l'on avait produit cette année-là un exabyte (1 million de terabytes) de données dont 99,997% sont disponibles sous forme électronique *seulement* (voir [Keim 2001]). En 2003, la quantité d'information rapportée à la population mondiale correspond à 800 Mo de nouvelles données produites par personne en un an [Peter and Varian 2003]. Cette abondance d'information apporte toutefois sont lots de questions à résoudre. Nombre de champs de recherche se penchent actuellement sur la construction de méthodes et d'outils capables d'exploiter cette richesse d'information. Les approches visuelles pour l'exploration et l'exploitation de ces informations connaissent un succès grandissant ; elles misent sur la capacité de l'œil humain à traiter efficacement des informations graphiques en y identifiant des motifs structuraux complexes – selon Ware [Ware 2000], 40% des activités de notre cortex sont dédiées au stimuli visuels. La visualisation d'information s'inspire d'idées ancrées dans des traditions d'origines diverses, dont la statistique graphique, la cartographie, le graphisme par ordinateur, l'interaction homme-machine, la psychologie cognitive, la sémiotique, le design graphique et l'art graphique [Tufte 1983; Tufte 1990; Tufte 1997], [Card, Mackinlay et al. 1999], [Herman, Marshall et al. 2000], [Spence 2001]. La visualisation porte ses fruits notamment lorsqu'elle vient en appui au travail de l'expert. C'est le cas en biologie et en bio-informatique [Schroeder, Gilbert et al. 2001], par exemple, et dans nombre d'autres domaines d'applications (voir [Johnson, Moorhead et al. 2006]). Lorsqu'on lance une requête sur Google Scholar avec les mots-clés « large information space », c'est sans surprise que l'on trouve parmi les liens les plus pertinents (au sens de Google) les références aux travaux actuels en visualisation d'information.

La complexité que l'on perçoit en première analyse lorsqu'on aborde des données massives est souvent inhérente aux données, à la complexité de chacun des attributs des données et à leurs agencements. Il suffit de penser par exemple aux données d'expression de la génomique ou de la protéomique, aux réseaux d'interactions de protéines très étudiés aujourd'hui à l'aide d'outils de visualisation. Or, une autre dimension vient aujourd'hui se glisser dans cette « complexité », *sociale* celle-là. La technologie s'ouvre à un type de navigation qui permet à l'utilisateur de suivre le cheminement des autres utilisateurs afin d'être guidée non pas par le contenu seulement *mais aussi par l'usage*. Familier de l'intelligence collective, il peut mettre à profit dans sa recherche d'information ce qu'il connaît ou croit connaître des autres utilisateurs appartenant à sa *communauté*. Il peut valider la pertinence des informations accédées par rapport à leur contexte social ou contexte d'usage ramené à sa propre expérience. Nous parlons de la *navigation sociale* telle que posée déjà en 1994 par Chalmers [Dourish and Chalmers 1994], où l'utilisateur a accès à la « trace » des activités des autres [Dieberger, Höök et al. 2001].

Ici encore la visualisation d'information a son rôle à jouer. Elle peut apporter un appui décisif pour la navigation en révélant visuellement la dimension sociale de l'environnement exploré. Les travaux qui se sont penchés sur la navigation sociale [Erickson and Kellogg 2000], [Riedl and St-Amant 2003] distinguent souvent les systèmes permettant à l'utilisateur de percevoir directement la présence et l'action des autres, et ceux où la trace sociale n'est accessible qu'au travers de l'historique des usages de l'espace d'information et des échanges antérieurs à sa connexion. Nombre de travaux se sont penchés sur le travail collaboratif et sur les artifices à mettre en œuvre pour faciliter l'interaction synchrone entre utilisateurs afin de donner « corps » aux échanges favorisant l'émergence de comportements « sociaux ».

La dimension spatiale, notamment aux travers des avatars, mais aussi de plongement dans des univers virtuels 2D ou 3D a le plus souvent été conçue de manière à apporter à l'utilisateur des indices visuels ou sonores rappelant l'environnement réel [Zhao and Stasko 2000], [Dieberger, Höök et al. 2001]. Cependant, deux autres dimensions sont présentes lorsqu'il s'agit de naviguer un espace d'information : la dimension sémantique et la dimension sociale [Dourish and Chalmers 1994]. La visualisation d'information, parfois moins attaché à plonger l'information dans la géométrie euclidienne du monde réel, tente de concevoir des vues dominées par la dimension sémantique – le défi peut se résumer à trouver une correspondance entre « proximité » sémantique et proximité sur l'écran ou « proximité dans l'interaction ».

Ainsi, on peut imaginer enrichir un corpus de documents (textes, audio et/ou vidéo) d'informations concernant leurs usages *et leurs usagers*.

Visualisation et FOAD

On pourrait imaginer par exemple enrichir une plate-forme de formation FOAD de cartographies permettant à l'utilisateur, apprenant (étudiant) ou enseignant, de visualiser les liens entre les documents : proximité de contenus, mais proximité d'usage parce que des utilisateurs ayant consulté le premier document ont aussi consulté le second (« à la Amazon ») ; liens entre documents et intervention sur les forums, par exemple : parce que des utilisateurs ont échangé sur un exercice apparaissant dans le document, permettant ainsi à l'utilisateur de visualiser que cet échange est en rapport son objet d'étude et de consulter le fil de discussion sans avoir à rechercher les forums ; etc. Pourquoi ne pas permettre à l'utilisateur enseignant de superviser le travail d'un étudiant en lisant la cartographie augmentée d'indices décrivant son parcours personnel ? Plus généralement, si les concepteurs d'un cours en décrivent la cartographie dès sa conception en indiquant les liens entre objets pédagogiques, on imagine l'enseignant pouvant aller « piocher » parmi les objets pédagogiques ceux qu'il mettra au programme du semestre en visualisant leurs interdépendances. Du coup, la cartographie des objet pédagogiques sélectionnés déduite de la cartographie de tout le cours et pourrait être accessible aux étudiants afin qu'ils décident d'une stratégie pour effectuer l'ensemble des activités du cours¹.

Des systèmes comme NusEye pour la navigation de contenus syndiqués [Jarrett and Dennis 2005] ou celui de Wang et Chen [Wang and Chen 2004] propose des navigations de réseaux sociaux associant acteurs et contenus. Ces systèmes souffrent cependant de ne pouvoir traiter des informations massives impliquant plusieurs centaines d'utilisateurs, voire plus, et de grands ensembles documentaires ou espaces d'indexation. Typiquement, les espaces considérés sont des réseaux qui explicitent les liens entre utilisateurs, entre utilisateurs et documents ou entre documents. Lorsque l'on considère des ensembles de grandes tailles, utilisateurs et documents, et qu'il est de plus nécessaire de prendre en compte différents attributs ou types de relations, les réseaux deviennent engorgés et se prêtent mal aux techniques de dessin de graphes traditionnellement utilisées en analyse des réseaux sociaux [Scott 2000] ; [Battista, Eades et al. 1998], [Kaufmann and Wagner 2001]. Cette masse d'information éprouve aussi les techniques reposant sur des indices de centralités permettant d'identifier les acteurs clés et les « communautés » présentes dans le réseau.

Plus récemment, des approches reprenant les travaux sur le clustering de graphes se sont penchées sur l'identification de communautés d'acteurs dans les réseaux sociaux. Ces techniques tirent profit de propriétés structurelles des réseaux dits « petits-mondes » (small world) appelés ainsi par Watts [Watts 1999; Watts 2003] en référence aux travaux originaux de Milgram [Milgram 1967]. Les travaux de Auber *et al.* [Auber, Chiricota et al. 2003], ceux de Newman et Girvan [Girvan and Newman 2002; Newman and Girvan 2004] ou de Boutin *et al.* [Boutin and Hascoët 2003; Boutin and Hascoët 2004; Boutin, Thièvre et al. 2005; Boutin, Thièvre et al. 2006] proposent des méthodes permettant d'identifier la structure de communautés d'un réseau. L'adaptation de la méthode de Auber *et al.* au réseau du trafic aérien mondial offre des résultats prometteurs [Amiel, Melançon et al. 2005]. Ces méthodes s'inspirent des idées présentes dans les travaux de Granovetter [Granovetter 1973] et décident du rôle d'un lien dans le réseau : lien « fort » apportant cohésion à l'intérieur d'une communauté, ou plutôt lien « faible » s'ouvrant sur l'extérieur.

Le bénéfice de ce découpage tient à ce qu'il permet ensuite de naviguer dans le réseau en quête d'information. Les techniques de navigation conjuguant zoom et/ou déformations sémantiques (voir par exemple [Herman, Marshall et al. 2000]) donnent ainsi à la fois une vue détaillée de quelques acteurs et documents tout en maintenant à vue la globalité de l'espace. Des indices calculés en fonction de la région visitée permettent à l'utilisateur de repérer des régions d'intérêt. L'application Vizster, bien qu'elle se limite à visualiser les acteurs d'un réseau, est une tentative allant dans cette direction [Heer and Boyd 2005] qui se compare à certains produits que l'on trouve dans le monde commercial.

Les métaphores visuelles sont un vecteur très performant de la pensée analogique et très présent dans les modes de communication en société (la publicité, en particulier, leur laisse une place de

¹ C'est déjà l'usage : les auteurs d'ouvrages didactiques proposent souvent en début d'ouvrage une sélection de chapitres pour un cours d'un semestre, et le présente parfois sous forme graphique.

choix) [Mollerup 1999], [Schroeder 2002], [Phillips and McQuarrie 2004]. On peut miser sur la visualisation comme agent actif de l'intelligence collective. Cela exige de repenser la métaphore visuelle en y associant l'interaction. La visualisation d'information doit encore travailler au développement de techniques qui s'inséreront dans les pratiques des usagers pour faciliter la recherche d'information et la navigation dans les grandes masses de données. Les efforts de recherche sur l'évaluation des systèmes de visualisation vont dans ce sens [Fekete and Plaisant 2004], [Wijk 2005]. Il y a fort à faire afin de mieux cerner les métaphores visuelles et les modes de navigation adaptés en fonction des données visualisées ou des tâches effectuées [Shneiderman 1996], [Chi 2000]. Après une décennie de résultats concernant les aspects plus techniques de la discipline, l'étude des usages est mis au premier plan des travaux de cette communauté de recherche [Johnson, Moorhead et al. 2006].

Références

[Amiel, Melançon *et al.* 2005] Amiel, M., G. Melançon and C. Rozenblat (2005). "Réseaux multi-niveaux : l'exemple des échanges aériens mondiaux." M@ppemonde **79**(3-2005).

[Auber, Chiricota *et al.* 2003] Auber, D., Y. Chiricota, F. Jourdan and G. Melançon (2003). Multiscale navigation of Small World Networks. IEEE Symposium on Information Visualisation, Seattle, GA, USA, IEEE Computer Science Press, pp. 75-81.

[Battista, Eades *et al.* 1998] Battista, G. d., P. Eades, R. Tamassia and I. G. Tollis (1998). Graph Drawing: Algorithms for the Visualisation of Graphs, Prentice Hall.

[Boutin and Hascoët 2003] Boutin, F. and M. Hascoët (2003). Focus-Based Clustering for Multi-Scale Visualization. Proceedings of the Seventh International Conference on Information Visualization, IEEE Computer Society: 53.

[Boutin and Hascoët 2004] Boutin, F. and M. Hascoët (2004). Focus dependent multi-level graph clustering. Proceedings of the working conference on Advanced visual interfaces. Gallipoli, Italy, ACM Press: 167-170.

[Boutin, Thièvre *et al.* 2005] Boutin, F., J. Thièvre and M. Hascoët (2005). Multilevel Compound Tree Construction Visualization and Interaction. Human-Computer Interaction Interact'05, pp.

[Boutin, Thièvre *et al.* 2006] Boutin, F., J. Thièvre and M. Hascoët (2006). Focus-based filtering + clustering technique for power-law networks with small world phenomenon. SPIE Electronic Imaging / Visualization and Data Analysis, San Jose, California, SPIE IS&T, pp.

[Card, Mackinlay *et al.* 1999] Card, S. K., J. D. Mackinlay and B. Shneiderman (1999). Readings in Information Visualization. San Francisco, Morgan Kaufmann Publishers.

[Chi 2000] Chi, E. H. (2000). A Taxonomy of Visualization Techniques Using the Data State Reference Model. IEEE Symposium on Information Visualization, IEEE Computer Society, pp. 69.

[Dieberger, Höök *et al.* 2001] Dieberger, A., K. Höök, M. Svensson and P. Lönnqvist (2001). Social navigation research agenda. CHI '01 extended abstracts on Human factors in computing systems. Seattle, Washington, ACM Press: 107-108.

[Dourish and Chalmers 1994] Dourish, P. and M. Chalmers (1994). Running Out of Space: Models of Information Navigation. Ancillary Proceedings of BCS HCI'94, Glasgow, U.K., pp.

[Erickson and Kellogg 2000] Erickson, T. and W. A. Kellogg (2000). "Social translucence: an approach to designing systems that support social processes." ACM Transactions on Computer-Human Interaction **7**(1): 59-83.

[Fekete and Plaisant 2004] Fekete, J.-D. and C. Plaisant (2004). Les leçons tirées des deux compétitions de visualisation d'information. IHM2004, International Conference Proceedings Series, Namur, Belgium, ACM Press, pp. 7-12.

[Girvan and Newman 2002] Girvan, M. and M. E. J. Newman (2002). "Community structure in social and biological networks." Proceedings of the National Academy Science USA **99**: 7821-7826.

[Granovetter 1973] Granovetter, M. S. (1973). "The Strength of Weak Ties." The American Journal of Sociology **78**(6): 1360-1380.

[Heer and Boyd 2005] Heer, J. and D. Boyd (2005). Vizster: Visualizing Online Social Networks. IEEE Symposium on Information Visualization (INFOVIS'05), pp. 5.

- [Herman, Marshall *et al.* 2000] Herman, I., M. S. Marshall and G. Melançon (2000). "Graph Visualisation and Navigation in Information Visualisation: A Survey." IEEE Transactions on Visualization and Computer Graphics **6**(1): 24-43.
- [Jarrett and Dennis 2005] Jarrett, A. C. and B. M. Dennis (2005). NusEye: designing for social navigation in syndicated content. Proceedings of the 2005 conference on Diversity in computing. Albuquerque, New Mexico, USA, ACM Press.
- [Johnson, Moorhead *et al.* 2006] Johnson, C., R. Moorhead, T. Munzner, H. Pfister, P. Rheingans and T. S. Yoo, Eds. (2006). NIH-NSF Visualization Research Challenges Report, IEEE Computer Society.
- [Kaufmann and Wagner 2001] Kaufmann, M. and D. Wagner, Eds. (2001). Drawing Graphs, Methods and Models. Lecture Notes in Computer Science, Springer.
- [Keim 2001] Keim, D. A. (2001). "Visual exploration of large data sets." Communications of the ACM **44**(8): 38-44.
- [Milgram 1967] Milgram, S. (1967). "The small world problem." Psychology today **1**(61).
- [Mollerup 1999] Mollerup, P. (1999). Marks of Excellence, Phaidon Press.
- [Newman and Girvan 2004] Newman, M. E. J. and M. Girvan (2004). "Finding and evaluating community structure in networks." Physics Reviews E **69**(026113).
- [Peter and Varian 2003] Peter, L. and H. R. Varian. (2003). "How Much Information." Retrieved March 2006, from <http://www.sims.berkeley.edu/how-much-info-2003>.
- [Phillips and McQuarrie 2004] Phillips, B. J. and E. F. McQuarrie (2004). "Beyond Visual Metaphor: A New Typology of Visual Rhetoric in Advertising." Marketing Theory **4**(1-2): 113-136.
- [Riedl and St-Amant 2003] Riedl, M. O. and R. St-Amant (2003). Social navigation: modeling, simulation, and experimentation. Proceedings of the second international joint conference on Autonomous agents and multiagent systems. Melbourne, Australia, ACM Press: 361-368.
- [Schroeder 2002] Schroeder, J. (2002). Visual Consumption, Routledge.
- [Schroeder, Gilbert *et al.* 2001] Schroeder, M., D. Gilbert, J. van Helden and P. Noy (2001). "Approaches to visualisation in bioinformatics: from dendrograms to Space Explorer." Information Sciences **139**(1-2): 19-57.
- [Scott 2000] Scott, J. P. (2000). Social Network Analysis: A Handbook, SAGE Publications.
- [Shneiderman 1996] Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization. IEEE Conference on Visual Languages (VL'96), Boulder, CO, IEEE CS Press, pp. 336-343.
- [Spence 2001] Spence, R. (2001). Information Visualization. Harlow, England, ACM Press/Addison-Wesley.
- [Tufte 1983] Tufte, E. R. (1983). The Visual Display of Quantitative Information. Cheshire, CT, USA, Graphics Press.
- [Tufte 1990] Tufte, E. R. (1990). Envisioning Information. Cheshire, CT, USA, Graphics Press (8th printing, June 2001).
- [Tufte 1997] Tufte, E. R. (1997). Visual Explanations. Cheshire, CT, USA, Graphics Press.
- [Wang and Chen 2004] Wang, J.-C. and C.-L. Chen (2004). "An Automated Tool for Managing Interactions in Virtual Communities-Using Social Network Analysis Approach." Journal of Organizational Computing and Electronic Commerce **14**(1): 1-26.
- [Ware 2000] Ware, C. (2000). Information Visualization: Perception for Design. Orlando, FL, Morgan Kaufmann Publishers.
- [Watts 1999] Watts, D. J. (1999). Small Worlds, Princeton University Press.
- [Watts 2003] Watts, D. J. (2003). Six Degrees: The Science of a Connected Age, W. W. Norton & Company.
- [Wijk 2005] Wijk, J. J. v. (2005). The Value of Visualization. IEEE Visualization, IEEE Computer Society, pp. 79-86.

[Zhao and Stasko 2000] Zhao, Q. A. and J. T. Stasko (2000). What's happening?: the community awareness application. CHI '00 extended abstracts on Human factors in computing systems. The Hague, The Netherlands, ACM Press.