

« Pardonnez cette platitude » : de l'intérêt des ethnographies de laboratoire pour l'étude des processus algorithmiques¹

Florian Jaton²

Intelligence artificielle ! Le terme a beau pointer vers des agencements actuels et problématiques, il peut aisément sonner creux aux oreilles de celles et ceux qui, comme moi, se réclament des *Science & Technology Studies (STS)*³. « Intelligence » ? Une notion pour le moins dangereuse qu'il conviendrait, a minima, d'employer au pluriel tant les enquêtes d'Edwin Hutchins, Reviel Netz, Lucy Suchman et bien d'autres, ont démontré qu'il n'y avait que des processus cognitifs équipés, collectifs, affectifs et distribués⁴. « Artificiel » ? Les travaux empiriques puis théoriques entrepris par Michel Callon, Bruno Latour, John Law, Michael Lynch, Susan Leigh Star et bien d'autres, sur la formation des énoncés scientifiques et des assemblages techniques n'ont cessé de remettre en cause l'opposition entre « naturel » et « artificiel », « réel » et « construit », « humain » et « machine ». On comprendra donc une certaine réticence à aujourd'hui traiter d'un objet estampillé « Intelligence Artificielle » : que de malentendus en perspective ! Pour celles et ceux qui s'inscrivent dans la lignée des auteurs précités, l'objet « IA » est au mieux un serpent de mer, au pire une manifestation douloureuse de l'absence de portée sociale et politique des travaux de sensibilité STS⁵.

Du fait peut-être, en partie, de l'extrême équivocité du terme IA, davantage d'auteurs se réclamant des STS se sont saisis d'un objet connexe mais tout autant d'actualité : les méthodes informatiques de calcul, souvent appelées « algorithmes ». Depuis quelques années, en réaction notamment à la multiplication des travaux, pour beaucoup issus des *Media* et des *Surveillance Studies*, traitant des effets sociaux des algorithmes, plusieurs auteurs de sensibilité STS ont proposé de « rentrer dans les algorithmes » afin de mieux cerner leurs fonctionnements⁶. L'objectif scientifique et politique de cette ligne de recherche est clair : au lieu de se limiter à l'étude des effets des algorithmes, il est maintenant temps d'enquêter *sur les causes de ces effets* pour mieux les changer.

Bien que salutaires, ces études sociales des algorithmes – qui croisent parfois des problématiques liées aux dispositifs dits « d'intelligence artificielle » – se heurtent néanmoins à une limite qui a trait selon moi à leurs matériaux d'enquête. En effet, en basant leurs discussions sur le fonctionnement d'algorithmes déjà constitués – en prenant notamment appui sur des articles scientifiques, des rapports et des manuels –, ces études peinent à considérer les hésitations, tests et autres dispositifs pratiques qui ont

¹ Zilsel 5, Février 2019. Antépénultième version d'un article prenant par au numéro spécial *Frictions* intitulé « Les STS et l'IA : une rencontre manquée ? » dirigé par Olessia Kirtchik.

² École polytechnique fédérale de Lausanne, Digital Humanities Institute, florian.jaton@gmail.com.

³ Peter Debaz et Sheila Jasanoff ont à mon avis raison de dire que ce qui relie – lâchement – les chercheuses et chercheurs de cette communauté de recherche hétérogène labellisée *Science & Technology Studies* est la conviction que « depuis les travaux de Ludwick Fleck, Thomas Khun, et David Bloor, la science n'est pas le froid royaume de l'empirisme logique et des principes premiers, que nous préexistons nos connaissances du monde (tout comme le monde préexiste nos connaissances), que la matérialité est centrale à la formation et à la mise à l'épreuve des vérités scientifiques, et que les sciences et les dynamiques des pratiques scientifiques et techniques sont un terrain fertile à l'analyse sociale, politique et éthique » (ma traduction). Peter Dear et Sheila Jasanoff, « Dismantling Boundaries in Science and Technology Studies », *Isis*, vol. 101, No. 4, 2010, p. 761. Sur l'hétérogénéité des STS, voir Dominique Vinck, « Les science studies : de la marginalité thématique à la refondation de la discipline », *SociologieS* [En ligne], Dossiers, Sociétés en mouvement, sociologie en changement, mis en ligne le 07 mars 2016, consulté le 2 juin 2016. URL : <http://sociologies.revues.org/5248>

⁴ Edwin Hutchins, *Cognition in the Wild*, Cambridge (Mass.), The MIT Press, 1995 ; Reviel Netz, *The Shaping of Deduction in Greek Mathematics : A Study in Cognitive History*, Cambridge (UK), Cambridge University Press, 2003 ; Lucy Suchman, *Human-Machine Reconfigurations : Plans and Situated Actions*, Cambridge (UK) ; Cambridge University Press, 2006. Voir aussi le récent et important travail de Simon Penny sur les liens entre computation et cognition : Simon Penny, *Making Sense : Cognition, Computing, Art, and Embodiment*, Cambridge (Mass.), The MIT Press, 2017.

⁵ Sur les raisons potentielles de l'absence de portée politique et sociale des travaux STS, voir William Thomas, « Les études de controverses et le récit du renouveau méthodologique en STS », *Zilsel*, No 2, 2017, pp. 131-147.

⁶ Notamment Dominique Cardon, *À quoi rêvent les algorithmes. Nos vies à l'heure des Big Data*, Paris, Seuil, 2015 ; Daniel Neyland, « Bearing Account-able Witness to the Ethical Algorithmic System », *Science, Technology & Human Values*, Vol 41, No 1, 2016, pp. 50-76 ; Adrian MacKenzie, *Machine Learners : An Archeology of a Data Practice*, Cambridge (Mass.), The MIT Press, 2017. Je reviendrai sur ces textes dans la prochaine partie de l'article.

précédemment participé à la constitution de ces fonctionnements. S'ensuit la mise sous silence problématique d'une myriade d'objets, d'habitudes et d'équipements qui fondent pourtant la véridiction et la fonctionnalité des processus algorithmiques considérés, qu'il s'agisse de méthodes d'apprentissage automatique (*machine learning*) prenant part à des systèmes dit « d'intelligence artificielle » ou non. L'argument, somme toute assez banal, que je souhaite développer dans cette note consiste ainsi à souligner l'intérêt des études ethnographiques de laboratoire qui, elles, n'ont pas d'autres choix que de considérer le fonctionnement des algorithmes comme des *résultats incertains*. En s'attachant à rendre compte des processus pratiques, hétérogènes et instables qui participent à la constitution progressive des algorithmes, de leurs balbutiements initiaux au sein de laboratoires – académiques ou industriels – à leur solidification graduelle au sein de documents spécifiques (programmes informatiques, brevets, articles scientifiques, rapports techniques, etc.), ces enquêtes de terrain – encore largement à faire – permettraient de préciser l'écologie des univers algorithmiques avec lesquels nous interagissons, fournissant ainsi davantage de prises pour mieux les apprécier ou les contester. Pour étayer mon propos, je prendrai l'exemple d'une étude menée au sein d'un laboratoire d'informatique qui a permis de décrire la centralité des *ground truths*, ces bases de données référentielles qui servent à la fois d'assises matérielles et d'éta-lons de mesure à bon nombre de méthodes informatiques de calcul⁷.

De l'analyse des effets des algorithmes à celle de leurs fonctionnements

En Amérique du nord et en Europe occidentale, les années 1990 furent particulièrement riches en discours à tendance utopique quant au déploiement des technologies informatiques⁸. La période était en effet propice : entre la distribution croissante de micro-ordinateurs dans les sphères professionnelles et privées, les débuts prometteurs de la technologie Internet, les annonces quant à la mise en place d'« autoroutes de l'information » ou encore l'arrivée imminente d'un nouveau millénaire, tout était aligné pour que se développe une vision prospective – et bien souvent promotionnelle⁹ – des « nouvelles technologies de l'information ». Face à ces discours mettant en équivalence « progrès technologique » et « avancées sociales »¹⁰, de nombreux chercheurs et chercheuses en sociologie et anthropologie s'attachèrent à mesurer les ressorts concrets, sociaux et politiques des processus d'informatisation. Pour ce qui est de la recherche francophone, l'analyse minutieuse de Lafrance et Brouillard sur la mise en place des infrastructures télécom au Canada, ou encore l'analyse comparative des politiques des autoroutes de l'information aux Etats-Unis, en Europe et au Japon faite par Vedel sont assez illustratives des approches alors privilégiées : il s'agissait de décrire les dynamiques des grands projets d'informatisation afin d'être en mesure de distinguer le sensationnel de l'ordinaire, l'illusoire du légitime¹¹. Aux Etats-Unis, comme suggéré par Poon et d'autres¹², la loi de 1993 stipulant que 5% des fonds du *Human Genome Research Project* devaient être investis dans la recherche sur l'implication sociale, légale et éthique de la recherche en génomique contribua à positionner les praticiens des sciences sociales comme gardiens de la justice et des libertés fondamentales face au déploiement des nouvelles technologies. La seconde édition de l'anthologie *Computerization and Controversy* publiée par Kling en 1996 – qui traite aussi bien des problématiques professionnelles, domestiques, fictionnelles, affectives et sécuritaires en lien avec les

⁷ Florian Jatton, « We Get the Algorithms of Our Ground Truths : Designing Referential Databases in Digital Image Processing », *Social Studies of Science*, Vol 47, No 6, 2017, pp. 811-840.

⁸ Deux exemples, certes arbitraires mais illustratifs, de l'utopisme technologique des années 1990 pourraient être Ross Weiland, « 2001 : A Meetings Odyssey », *Successful Meetings*, Vol 42, No 13, 1993, pp. 34-39 ; mais aussi Thomas A. Stewart, « Boom Time on the New Frontier », *Fortune*, Vol 128, No 7, 1993 : 153-161.

⁹ Michael Kranz, « Cashing in on Tomorrow », *Time*, 15 juillet, 1996.

¹⁰ Sur les ressorts de la rhétorique de la « révolution informatique » du début des années 1990, voir Rob Kling et Suzanne Iacono, « Making a Computer Revolution », *Journal of Computing and Society*, Vol 1, No 1, 1990, pp. 43-58. Pour une généalogie du terme « société de l'information », voir Armand Mattellart, « L'âge de l'information : genèse d'une appellation non contrôlée », *Réseaux*, Vol 18, No 101, 2000, pp. 19-52.

¹¹ Jean-Paul Lafrance et Pierre Brouillard, « À la conquête électronique de la toison d'or ou la mise en place des contenus et des services sur les autoroutes de l'information », *Réseaux*, No 78, 1996, pp. 53-83 ; Thierry Vedel, « Les Politiques des autoroutes de l'information dans les pays industrialisés : une analyse comparative », *Réseaux*, No 78, pp. 11-28. La revue *Réseaux* a fortement contribué à mettre en avant l'étude sociale des technologies informatiques dans le monde francophone durant les années 1990. Voir également le double numéro 72-73 (1995) consacré à l'économie des télécommunications.

¹² Martha Poon, « Corporate Capitalism and the Growing Power of Big Data : Review Essay », *Science, Technology, & Human Values*, Vol 41, No 6, 2016, pp. 1-21 ; Jenny Reardon, « On the Emergence of Science and Justice », *Science, Technology, & Human Values*, Vol 38, No 2, 2013, pp. 176-200 ; Hub Zwart et Annemiek Nelis, « What is ELSA Genomics ? Science & Society Series on Convergence Research », *EMBO Reports*, Vol 10, No 6, 2009, pp. 540-544.

processus d'informatisation – donne en ce sens un bel aperçu des thématiques qui traversaient l'étude des conséquences sociales et politiques des technologies informatiques durant cette époque charnière¹³. Cette ligne de recherche se réclamant d'un « réalisme social »¹⁴ vis-à-vis des technologies informatiques se poursuit durant les années 2000 avec toutefois un progressif rétrécissement thématique, notamment en Amérique de nord¹⁵, qui participa à l'éclosion de l'objet « algorithme » dans les sciences sociales. Des efforts ciblés et conséquents furent d'abord investis dans la critique du discours sur l'*empowerment* et l'accessibilité des informations mis en avant par les promoteurs des technologies du Web. Hoffman et Novak ont par exemple montré que l'accessibilité et l'usage des technologies du Web aux Etats-Unis étaient grandement fonction de différences raciales¹⁶. Lawrence et Giles ont quant à eux montré que, contrairement aux discours promotionnels faisant état d'une mise en accès quasi illimitée, les moteurs de recherche disponibles à la fin des années 1990 ne parvenaient à indexer qu'une petite fraction du Web¹⁷. Dans le même registre, Introna et Nissenbaum se sont attachés à souligner l'influence souterraine – et potentiellement néfaste – des heuristiques utilisés pour la classification des URLs par ces mêmes moteurs de recherche¹⁸. La période post-11 septembre 2001 qui suivit fit la part belle aux critiques des biais inscrits dans les programmes et *algorithmes* – le terme apparaît à ce moment-là¹⁹ – de surveillance et de détection précoce. Dans son étude des implications sociales des technologies d'extraction de données (*data mining*), Gandy prévient par exemple qu'elles sont la porte ouverte à de la discrimination rationnelle participant au renforcement des habitudes corrélatives entre statut social et appartenance à un groupe²⁰. Dans une perspective empruntant à l'économie politique, Zureik traite des technologies et algorithmes biométriques ainsi que de leurs propensions à banaliser le profilage social et les politiques publiques de catégorisation et d'exclusion de groupes nationaux²¹. Un autre exemple de cette ligne de recherche en *Surveillance Studies* pourrait être le travail d'Introna et Wood : leur analyse des algorithmes de reconnaissance faciale souligne les biais potentiels de ces dispositifs alors souvent présentés comme impartiaux²². La congruence et les échanges entre ces deux univers thématiques – « étude des nouveaux médias » et « surveillance » – aboutit, à l'orée des années 2010, à de nombreux travaux sur les discriminations et les mises en invisibilité induites par le recours à des dispositifs algorithmiques²³.

¹³ Rob Kling (ed), *Computerization and Controversy : Value Conflicts and Social Choices*, Londres, Academy Press Limited, 1996 (1991).

¹⁴ « Social realism » est en effet l'expression utilisée par Kling pour caractériser le genre propre à l'analyse sociale des effets des technologies numériques dans : Rob Kling, « Hopes and Horrors : Technological Utopianism and Anti-Utopianism in Narratives of Computerization », in Rob Kling (ed), *Computerization and Controversy : op. cit.*

¹⁵ En France, le processus semble avoir été inverse : au lieu de rétrécir, les thématiques de recherche liées aux technologies numériques se sont élargies à l'étude des pratiques de quantification, notamment au sein du *Centre de Sociologie de l'Innovation* via les travaux de Claude Rosental, Michel Callon et Fabian Muniesa. Voir entre autres Rosental, *La trame de l'évidence : Sociologie de la démonstration en logique*, Paris, PUF, 2003 ; Michel Callon et Fabian Muniesa, « Economic Markets as Calculative Collective devices », *Organization Studies*, Vol 26, No 8, 2005, pp. 1229-1250.

¹⁶ Donna L. Hoffman et Thomas P. Novak, « Bridging the Racial Divide on the Internet », *Science*, Vol 580, No 5362, 1998, pp. 390–391.

¹⁷ Steve Lawrence et C. Lee Giles, « Accessibility of Information on the Web », *Nature*, Vol 400, Juillet, 1999, pp. 107-109.

¹⁸ Lucas D. Introna et Helen Nissenbaum, « Shaping the Web : Why the Politics of Search Engines Matter », *The Information Society*, Vol 16, No 3, 2000, pp. 169-185.

¹⁹ Alors qu'il lui était préféré les termes « software », « code » ou encore « software-algorithm », le terme unique « algorithm » commence à être de plus en plus commun dans la littérature anglo-saxonne des *Surveillance Studies* à partir des années 2000. Il serait intéressant de mieux connaître les canaux par lesquels le terme « algorithm » a fini par s'imposer dans la littérature critique anglo-saxonne.

²⁰ Oscar H. Gandy Jr., « Data Mining and Surveillance in the Post-9.11 Environment », in Kristie Ball et Frank Webster (eds), *The Intensification of Surveillance Crime, Terrorism and Warfare in the Information Age*, London, Pluto Press, 2002, pp. 26-41.

²¹ Elia Zureik et Karen Hindle, « Governance, Security and Technology : The Case of Biometrics », *Studies in Political Economy*, Vol 73, Été, 2004, pp. 113-137.

²² Lucas Introna et David Wood, « Picturing Algorithmic Surveillance : The Politics of Facial Recognition Systems », *Surveillance & Society*, Vol 2, No2-3, 2004, pp. 177-198.

²³ Sur les discriminations induites par les algorithmes, voir par exemple Felicitas Kraemer, Kees Overveld et Martin Peterson, « Is There an Ethics of Algorithms ? », *Ethics and Information Technology*, Vol 13, No 3, 2010, pp. 251-260. ; mais aussi Tarleton Gillespie, « The Relevance of Algorithms », in Tarleton Gillespie, Pablo Boczkowski et Kirsten Foot (eds), *Media Technologies*, Cambridge (Mass.), The MIT Press, 2013, pp. 167-194. Sur l'invisibilisation, voir par exemple Tania Bucher,

Il ne faut pas sous-estimer l'importance de ces études critiques des effets sociaux des algorithmes. Ces travaux ont en effet opéré, à leur niveau, comme contre-feu à la rhétorique marchande et séductrice des promoteurs des nouvelles technologies informatiques (j'y reviendrai plus bas). Pour autant, leur multiplication et, d'une certaine manière, leur redondance ont néanmoins eu – et continuent d'avoir – un effet néfaste sur l'étude des processus algorithmiques. C'est cet effet subreptice que Ziewitz a récemment bien su déceler avec sa notion de « drame algorithmique » (*algorithmic drama*)²⁴.

L'argument de Ziewitz, que je partage, opère comme suit. Alors que les enquêtes des années 1980-90, aussi bien dans le monde anglo-saxon que francophone, prenaient soin de mettre en avant les éléments périphériques mais cruciaux (standards, habitudes, politiques publiques, équipements) permettant aux méthodes informatiques de calcul – que l'on peut nommer « algorithme », si on y tient – de fonctionner, l'empilement des enquêtes plus tardives sur les effets des algorithmes a eu comme conséquence de mettre de côté ces éléments périphériques. Petit à petit, alors que les travaux critiques sur l'agentivité (*agency*) des algorithmes se multipliaient, ces objets ont progressivement commencé à apparaître comme autonomes : les réseaux dont ils ont pourtant besoin pour se déployer et se maintenir ont été lentement ignorés. Ce glissement progressif constitue le premier acte du « drame algorithmique » : de dispositifs ancrés empiriquement et délicatement irrigués, les algorithmes sont progressivement devenus de *puissantes entités flottantes*.

Ziewitz affirme en outre qu'une fois détachés des fragiles réseaux leur permettant de se déployer, les algorithmes sont également devenus de plus en plus mystérieux. En effet, de quoi ces puissantes entités pouvaient-elles être faites ? Comme les connexions entre les algorithmes et ce qui les soutient ont eu tendance à ne plus être considérées dans la littérature sociologique, surtout anglo-saxonne, ces objets ne pouvaient qu'être faits d'ingrédients « théoriques », « immatériels », ou « abstraits » faisant parfois référence aux mathématiques²⁵, à du code informatique²⁶, ou une combinaison des deux²⁷. N'ayant plus de prises sur le contenu de ces ensembles, la notion de « complexité » fut appelée en renfort²⁸ : peu importe ce à quoi pouvaient bien faire référence les mathématiques ou le code qui constituent les algorithmes, ceux-ci se devaient d'être hautement complexes puisqu'ils sont abstraits *et* puissants. Comment quelque chose pourrait-il être, dans le même mouvement, évanescence *et* agissant ? C'est ce genre de question induite – bien souvent en creux – par la multiplication des travaux sur les effets des algorithmes qui introduisit subrepticement le deuxième acte du « drame algorithmique » : les algorithmes sont progressivement devenus des entités *insondables*²⁹. Et Ziewitz de rajouter : « l'opacité des opérations tend à être considérée comme un nouveau signe de leur influence et leur pouvoir »³⁰. Le drame algorithmique est ainsi circulaire : les algorithmes sont puissants *parce qu'ils sont insondables, parce qu'ils sont puissants, parce qu'ils sont insondables...*

C'est contre ce cercle vicieux que sont intervenus, plus ou moins explicitement, plusieurs auteurs de sensibilité STS. Dans son ouvrage *À quoi rêvent les algorithmes*, Cardon propose par exemple d'« ouvrir la boîte noire » et d'analyser les ressorts de ces dispositifs³¹. C'est là pour lui un projet politique

« Want to Be on the Top ? Algorithmic Power and Threat of Invisibility on Facebook », *New Media & Society*, Vol 14, No 7, 2012, pp. 1164-1180. ; également Engin Bozdag, « Bias in Algorithmic Filtering and Personalization », *Ethics and Information Technology*, Vol 15, No 3, 2013, pp. 209-227. Une monographie très consultée fut aussi Christophe Steiner, *Automate This : How Algorithms Came to Rule our World*, New York, Penguin, 2012.

²⁴ Malte Ziewitz, « Governing Algorithms : Myth, Mess, and Methods », *Science, Technology, & Human Values*, Vol 41, No 1, 2016, pp. 3-16.

²⁵ Robin K. Hill, « What an Algorithm Is », *Philosophy and Technology*, Vol 29, No 1, 2016, pp. 35-59.

²⁶ Lawrence Lessig, « Code Is Law : On Liberty in Cyberspace », *Harvard Magazine*, Janvier, 2000.

²⁷ Brent D. Mittlestadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter et Luciano Floridi, « The Ethics of Algorithms : Mapping the Debate », *Big Data & Society*, Vol 3, No 2, 2016, pp. 1-21.

²⁸ À noter que « complexité » n'est pas ici considérée dans ses acceptations mathématiques et informatiques. Sur les potentiels liens entre théories de la complexité, *machine learning* et sciences sociales, voir Dominique Boullier et El Mahdi El Mhadi, « Des modèles aux pratiques : le Machine Learning à l'épreuve de la complexité », à paraître dans *Revue d'Anthropologie des Connaissances*.

²⁹ On retrouve cet argument dans Matteo Pasquinelli, *Anomaly Detection : The Mathematization of the Abnormal in the Metadata Society*, Berling, Transmediale, 2015.

³⁰ Malte Ziewitz, 2016, op. cit., p. 6.

³¹ Dominique Cardon, 2015, op. cit. Dominique Cardon fait selon moi figure de précurseur de l'analyse sociale des algorithmes en STS, du moins en ce qui concerne les technologies du Web. Dès 2013, il publia par exemple une importante recension de l'architecture calculatoire de l'algorithme PageRank de Google : Dominique Cardon, « Dans l'esprit de PageRank : Une enquête sur l'algorithme de Google », *Réseaux*, No 177, 2013, pp. 63-95. Voir également le récent et important travail sur la

important : « Une radiographie critique des algorithmes est un enjeu démocratique aussi essentiel qu'inaperçu »³². Et au travers de plusieurs chapitres fort bien amenés, où il en vient par exemple à proposer quatre familles d'algorithmes en fonction de leurs rapports aux données numériques, les fonctionnements internes de ces dispositifs apparaissent, sinon plus précis, du moins plus saisissables. Dans un article publié dans une des revues références des STS – *Science, Technology, & Human Values* – Neyland annonce à peu de chose près le même programme : « Au lieu de considérer l'agentivité et le pouvoir comme inscrits dans les algorithmes, comme le suggère l'éloquent drame algorithmique, cette recherche STS peut être utilisée pour soulever des questions quant aux ensembles de relations participant à la mise en existence des algorithmes »³³. En s'inspirant de l'ethnométhodologie et son concept de « accountability », Neyland se propose ensuite d'analyser comment un système algorithmique participe à la redéfinition d'espaces spécifiques, en l'occurrence un aéroport et une station de train. En se basant sur des échanges d'e-mails et des rapports de séances, il rend compte de la façon dont le cahier des charges de cet algorithme redéfinit le travail de celles et ceux qui seront amenés à interagir avec ce système. Plus récemment, Mackenzie analyse le fonctionnement interne d'algorithmes d'apprentissage automatique (*machine learning*) en se basant sur des manuels, des publications scientifiques et des cours à destination d'étudiants en informatique³⁴. Il tente ainsi une relecture de plusieurs techniques d'apprentissage automatique à l'aune du concept deleuzio-foucauldien de « diagramme ». S'ensuit une riche succession d'analyses socio-philosophiques des techniques du gradient stochastique, des réseaux de neurones, des machines à vecteurs de support ou encore des méthodes de partitionnement récursif.

Ces travaux d'inspiration STS sur les algorithmes sont à bien des égards salutaires. Ils font en effet d'une pierre deux coups : en mettant en avant certains réseaux de relations qui font tenir les algorithmes, ces travaux relativisent leur pouvoir – car ils dépendent de nombreux autres éléments – ainsi que leur aspect insondable – car ils en passent bel et bien par des lieux et des situations assignables. Pour autant, ces travaux se heurtent à une limite qui, selon moi, contraint leurs ambitions politiques. C'est cette limite que je voudrais maintenant tenter d'explorer.

Une invisibilisation négative

Nous l'avons mentionné, ces récentes études sociales des algorithmes souhaitent rendre visible les composantes des algorithmes qui participent à la production d'effets souvent contestés. Mais pourquoi ces auteurs tentent-ils cet exercice ? N'y aurait-il d'autres choses à faire ? Peut-être, certainement. Mais en ce moment précis, pour ce sujet précis des méthodes informatiques de calcul, la mise en visibilité de composantes est, du moins en théorie, favorable à l'expression potentielle de revendications. C'est cet argument partagé plus ou moins explicitement par ces trois auteurs qu'il nous faut à présent considérer. Comme le suggèrent Star et Strauss³⁵, il y a des processus qui, peut-être, méritent provisoirement de rester invisibles sous peine de perdre ce qui donne sens à leurs résultats. Les tours de magie sont un exemple parmi d'autres : rendre visible – par exemple en produisant un compte-rendu sociologique – une partie des pratiques nécessaires à la maîtrise relative d'un tour de magie pourrait affecter négativement le tour en question³⁶. La surprise et l'émerveillement du résultat seraient comme bridés par les opérations terre à terre et terriblement incertaines qui l'ont rendu possible. En termes spinozistes, on pourrait dire que cette mise en visibilité tendrait à diminuer l'expression d'une puissance d'agir, transporter, émouvoir³⁷. En suivant la distinction proposée par Star et Strauss, l'invisibilité des ressorts du

résurgence des réseaux de neurones en science informatique : Dominique Cardon, Jean-Philippe Cointet et Antoine Mazières, « La revanche des neurones. L'invention des machines inductives et la controverse de l'intelligence artificielles », *Réseaux*, No 211, 2018, pp. 173-220.

³² Dominique Cardon, 2015, op. cit., p. 13.

³³ Daniel Neyland, 2016, op. cit., p. 52. Ma traduction.

³⁴ Adrian Mackenzie, 2017, op cit.

³⁵ Susan L. Star et Anselm Strauss, « Layers of Silence, Arenas of Voice : The Ecology of Visible and Invisible Work ». *Computer Supported Cooperative Work*, Vol 8, No 1-2, 1999, pp. 9-30.

³⁶ L'anthropologue Graham Jones a publié une monographie fort intéressante sur une communauté de magiciens et de *performers* de la région parisienne. Cela invalide-t-il cet argument quant à la relation entre « visibilité » et « invisibilité » ? Je n'en suis pas sûr. Car Jones porte un regard anthropologique sur une communauté de praticiens, et non pas un regard sociologique sur les ressorts des tours de magie de cette communauté (même si « communauté » et « tours de magie » y apparaissent comme liés). Pour se faire une idée sur la question, voir Graham Jones, *Trade of the Tricks : Inside the Magician's Craft*, Berkeley, University of California Press, 2011.

³⁷ Important de noter que cette réflexion ne s'applique pas à la sociologie des attachements telle que développée par Antoine Hennion qui, précisément, tente de documenter les médiations constitutives d'affections. Voir notamment Antoine Hennion,

tour de magie peut être ainsi considérée, provisoirement, comme *positive* : elle aide les produits de certaines pratiques « magiques » à être, par manque d'un meilleur terme, *adéquats*. L'absence d'un compte rendu faisant état de ses incertitudes et de sa fragilité aide le tour de magie à devenir tour, tout comme elle aide le public à devenir public du tour. En somme, il y aurait là un appel – peu contesté – à l'espérance de maîtrise³⁸.

Mais à d'autres moments, des processus très différents voient parfois l'adéquation de leurs résultats être sujette à controverses. N'est-ce pas le cas des algorithmes qui ne cessent d'irriguer nos pratiques quotidiennes ? Ne sont-ils pas aujourd'hui l'objet de disputes quant à leur rôle dans nos sociétés³⁹ ? C'est selon moi le grand mérite des études critiques des effets des algorithmes : elles sont parvenues à rendre *négative* l'invisibilité des rapports constitutifs des méthodes informatiques de calcul. En façonnant l'objet « algorithme », ces enquêtes critiques ont également façonné une nouvelle forme de pouvoir, lui-même suggérant la mise en place de nouveaux contre-pouvoirs.

Ce souffle critique initial fut à coup sûr rafraîchissant. Mais le drame algorithmique subséquent a tôt fait de calibrer le problème en des termes déléterés. Une fois devenus des objets flottants et insondables, les algorithmes ont progressivement suggéré des prises de positions peu conciliables : faut-il dorénavant croire les industriels de la Silicon Valley qui prônent le bienfait de la rationalité inscrite dans les algorithmes ou les académiciens et journalistes activistes qui contestent les biais, les exclusions et les aliénations que ces entités participent à faire survenir ? Faut-il soutenir celles et ceux qui critiquent la « vie algorithmique » et ses vues monopolistiques sur l'intelligibilité du réel⁴⁰ ou celles et ceux qui critiquent cette critique au nom d'une nouvelle rationalité désormais supérieure⁴¹ ? Des tranchées se creusent, le conflit s'enlise. Comment sortir de cette impasse et réinsuffler une politique des algorithmes ?

La réponse des auteurs STS précités – déjà esquissée plus haut – a le mérite d'être clair : afin de mettre en débat les algorithmes et les discuter politiquement, il est primordial de rendre compte de leur fonctionnement. L'argument, qu'au fond je partage, revient ainsi à affirmer que la mise en visibilité des constituants internes des algorithmes participerait à alimenter les délibérations quant au rôle et à la place des algorithmes dans nos sociétés. Comme le dit bien Cardon : « Cet examen est indispensable s'il on veut débattre publiquement des calculs que nous voulons et de ceux que nous ne voulons pas, contrôler leurs agissements et leur opposer des calculs alternatifs »⁴². Mais comment s'y prendre, concrètement ? Sur quoi baser ces enquêtes permettant elles-mêmes de suggérer des négociations sur et avec les algorithmes ?

Là se situe, à mon avis, la limite des récentes études sociales des algorithmes : les matériaux qu'elles utilisent pour rendre visibles les composantes des algorithmes contribuent, en fait, à *invisibiliser encore davantage* ces composantes. La faute à leur emphase sur les fonctionnements des systèmes algorithmiques, parti-pris commode mais tend malheureusement à passer sous silence ce qui participe à l'ajustement de fonctionnalités. C'est le grand danger de commencer l'analyse *par l'objet technique déjà constitué* : lorsque l'on part de ce qui persiste et fonctionne – par exemple une méthode informatique de calcul –, on s'interdit de rendre compte du mouvement de transformation, du « chemin de médiations », des nombreux détours qui ont permis l'assemblage provisoire, et toujours à maintenir, de cette persistance fonctionnelle⁴³. Et c'est précisément dans les mouvements de prélèvement et d'assemblage qui précèdent et rendent possible une persistance algorithmique que se cachent, peut-être, les objets dont nous avons besoin pour réinsuffler une politique des algorithmes.

« Vous avez dit 'attachements' ? », dans Madelaine Akrich, Yannick Barthe, Fabian Muniesa et Philippe Mustar (eds), *Débordements : Mélange offert à Michel Callon*, Paris, Presses des Mines, 2013, pp. 179-190.

³⁸ Jérôme Denis reformule bien cet argument en avançant qu'il suffit « d'un minimum d'expérience des enquêtes en milieux professionnels pour comprendre, ne serait-ce qu'intuitivement, que l'invisibilité des pratiques n'est pas toujours un problème et, inversement, que la visibilité n'est pas toujours une solution ». Jérôme Denis, *Le travail invisible des données. Éléments pour une sociologie des infrastructures scripturales*, Paris, Presses des Mines, 2018, p. 109.

³⁹ Sur la place croissante des algorithmes dans nos vies quotidiennes, voir notamment Massimo Mazzotti, « Algorithmic Life », *Los Angeles Review of Books*, 22 janvier, 2017. Consulté le 6 novembre 2018.

⁴⁰ Voir par exemple Éric Sadin, *La vie algorithmique : Critique de la raison numérique*, Paris, L'échappée, 2015.

⁴¹ Voir par exemple Brian Christian et Tom Griffiths, *Algorithms to Live By : The Computer Science of Human Decision*, New York, Henry Holt, 2016.

⁴² Dominique Cardon, 2015, op. cit., p. 13.

⁴³ Sur cette question de l'analyse des objets techniques, voir Bruno Latour, *Enquête sur les modes d'existence : Une anthropologie des Modernes*, Paris, La Découverte, 2013, pp. 212-235. Sur la vulnérabilité des objets, notamment techniques, et leur aspiration à la maintenance, voir Jérôme Denis et David Pontille, « Material Ordering and the Care of Things », *Science, Technology, and Human Values*, Vol 40, No 3, 2015, pp. 338-367.

Mais quels sont ces matériaux qui ne parviennent pas à rendre visible les « chaînes opératoires »⁴⁴ sous-jacentes aux techniques algorithmiques et qui donc participent – malgré eux – à la perpétuation de l’invisibilisation négative des composantes des algorithmes ? Il y a tout d’abord les articles scientifiques publiés dans les revues académiques, souvent utilisés comme sources primaires par les récentes études sociales des algorithmes⁴⁵. Il serait bien sûr incorrect de dire que cette littérature est erronée : au contraire, elle atteste de ce qui est sur le point de, peut-être, devenir scientifiquement vrai. Mais comme l’ont montré de nombreuses – et maintenant canoniques – études sociales sur les sciences⁴⁶, ces publications scientifiques tendent à rendre compte de résultats de processus, non des activités pratiques et incertaines qui ont mené à ces résultats. Difficile dans ces conditions de déceler les « invisibles » qui rendent possibles des persistances algorithmiques puisque ces invisibles sont précisément ce qui soutient la formulation des énoncés publiés par les revues spécialisées. Michael Lynch résume bien ce problème inhérent à l’analyse des publications scientifiques :

« Les éléments de méthode fournissent des maximes de conduite étape par étape à destination de praticiens déjà compétents assimilés à un mélange de sens commun et des pratiques de recherche non formulées ... Ces pratiques non formulées sont nécessairement omises des études sur les sciences qui s’appuient sur les résidus littéraires des laboratoires. »⁴⁷

En somme, comme elle s’adresse à des personnes déjà compétentes et habituées, la littérature scientifique ne prend pas soin de préciser les moyens alambiqués mais pratiques permettant d’obtenir cette littérature.

De plus, compte tenu de l’exigence de conviction propre à ce genre littéraire, les méthodes informatiques de calcul telles que présentées dans la littérature scientifique tentent de convaincre de leur pertinence. Un énoncé publié dans une revue scientifique est en effet toujours orienté contre d’autres énoncés et vise à obtenir l’adhésion des lecteurs. D’où l’importance des techniques de captation qui visent à « disposer le texte de façon que, quel que soit l’endroit où le lecteur veuille se rendre, il ne puisse prendre qu’un seul chemin »⁴⁸. Ces habitudes de conviction et le supplément de nécessité qu’elles établissent – primordiales pour garantir *in fine* l’objectivité des constructions scientifiques certifiées – tendent à épuiser les comptes rendus scientifiques portant sur les algorithmes de la myriade d’éléments disparates qui ont contribué à leur mise en existence textuelle. Lorsque l’on se base sur ces documents pour analyser les méthodes informatiques de calcul, ce sont donc les hésitations, les doutes, les équipements et écritures « infra-ordinaires »⁴⁹ qui tendent à échapper au regard de l’analyste.

Qu’en est-il des innombrables manuels et autres outils pédagogiques destinés à apprendre aux étudiants comment développer des algorithmes ? Participent-ils eux-aussi à la perpétuation de l’invisibilisation négative des composantes des algorithmes ? Ces matériaux sont certainement cruciaux pour inculquer aux étudiants et aux nouveaux venus les éléments de base des méthodes informatiques de calculs, chose indispensable à leur analyse. Pourtant, comme l’a bien montré Suchman, ces ressources pédagogiques sont par définition des comptes rendus *normatifs* de comment un travail devrait être fait, non pas des comptes rendus *descriptifs* de comment ce travail est effectivement fait. C’est une précision cruciale mais souvent oubliée : « Ces comptes rendus représentent des idéalizations de typifications. En tant que

⁴⁴ La notion de « chaîne opératoire » a été introduite par André Leroy-Gourhan dans sa publication de 1943, *L’homme et la matière : Évolution et techniques*, Paris, Albin Michel, 1943. La notion incite à considérer les objets techniques comme les résultats de matériaux, gestes et habitudes diverses dispersés dans le temps et l’espace. Cette notion a surtout été utilisée en ethnologie, archéologie et paléontologie. Pour une discussion de cette notion, voir la collection de textes réunis par Hélène Balfet, *Observer l’action technique : Des chaînes opératoires, pour quoi faire ?*, Paris, CNRS, 1991.

⁴⁵ C’est le plus souvent le cas des analyses de Cardon, notamment dans Dominique Cardon, 2012, op. cit., mais également – quoique par des voies détournées (en se référant par exemple à des textes qui se basent sur de la littérature scientifique) – dans Dominique Cardon, 2015, op. cit. C’est également courant dans le travail de Mackenzie, surtout dans Adrian Mackenzie, 2017, op. cit.

⁴⁶ Karin Knorr-Cetina, *The Manufacture of Knowledge. An Essay on the Constructivist and Contextual Nature of Science*, Oxford, Pergamon Press, 1981 ; Bruno Latour et Steve Woolgar, *Laboratory Life : The Construction of Scientific Facts*, Princeton, Princeton University Press, 1986 (1979) ; Michael Lynch, *Art and Artifact in Laboratory Science : A Study of Shop Work and Shop Talk in a Research Laboratory*, London ; Boston, Routledge Kegan & Paul, 1985.

⁴⁷ Michael Lynch, 1985, op. cit., p. 3.

⁴⁸ Bruno Latour, *La science en action : Introduction à la sociologie des sciences*, Paris, La découverte, 2005 (1989), p.140. Souligné dans le texte.

⁴⁹ Le terme « infra-ordinaire » a été proposé par Pérec, en opposition à « extra-ordinaire » : Georges Pérec, *L’infra-ordinaire*, Paris, Seuil, 1989. Le terme fut ensuite repris en sociologie, notamment par Murielle Lefebvre, « L’infra-ordinaire de la recherche. Écritures scientifiques personnelles, archives et mémoire de la recherche », *Sciences de la société*, No 89, pp. 3-17.

telles, elles dépendent pour leur écriture de l'effacement des contingences et des différences »⁵⁰. Ces matériaux sont ainsi intéressants pour la vision qu'ils donnent de certaines recettes péremptoires, non pas pour la vision qu'ils donnent des pratiques effectives et incertaines⁵¹. Malgré ces indéniables qualités, c'est tout le problème du travail de Mackenzie : en s'appuyant principalement sur des manuels et des cours sur le *machine learning*, il base ses analyses sur la façon dont certains pédagogues actuels souhaiteraient que les algorithmes en apprentissage automatique viennent au monde, et non pas de la façon dont ces dispositifs viennent effectivement au monde.

Quid maintenant des échanges d'emails et des rapports de séances utilisés par Neyland dans son étude de la mise en place d'un algorithme de sécurité ? Ces matériaux sont intéressants en ce qu'ils émanent des coulisses d'un projet algorithmique avant que celui-ci ne se solidifie en tant que dispositif fonctionnel. Mais le problème vient cette fois-ci du traitement que Neyland fait de ses matériaux : alors qu'il y avait là, peut-être, matière à documenter les dispositifs permettant de provisoirement surmonter la fragilité, les détours et l'incertitude entourant la formation d'une méthode informatique de calcul, l'auteur rend compte de cette formation comme d'une ligne qui file droit. Sur la base de ses matériaux, il montre en effet que des solutions techniques émanent d'alignements de positions mais il ne prend pas soin de rendre compte des objets, habitudes, et situations par lesquelles ces alignements s'épaississent. La trame du développement de ce dispositif algorithmique apparaît ainsi rectiligne et quasi-nécessaire.

En résumé, ces trois récentes études sociales des algorithmes ont – malgré leurs qualités – le défaut de perpétuer l'invisibilisation d'une myriade d'entités dont l'assemblage a pourtant participé à produire les algorithmes considérés. Les matériaux de leurs enquêtes – pour le cas de Cardon et MacKenzie – ou le traitement de leurs matériaux – pour le cas de Neyland – les empêchent d'apprécier les discontinuités qui ont mené aux continuités observables qui « fonctionnent », suggérant ainsi que le résultat final est le même que le mouvement qui y a mené : efficace et inévitable. Dans cette atmosphère de nécessité, il est difficile d'identifier des interstices capables de suggérer des prises critiques sur et avec les algorithmes.

Retour au laboratoire et nouveaux objets

À ce stade de l'article, je n'ai pas d'autres choix que de demander au lecteur de me suivre – du moins momentanément – lorsque je suggère que les tentatives *STS* récentes pour sortir du drame algorithmique participent malgré elles à l'invisibilisation négative d'éléments potentiellement utiles au projet politique consistant à alimenter le débat public sur les algorithmes et leurs effets bien réels. Si nous acceptons cette hypothèse – en elle-même discutable –, la prochaine question pourrait être : « Mais comment donc rendre visibles ces objets, dispositifs et habitudes “infra-ordinaires” potentiellement capables de suggérer de nouvelles prises critiques sur et avec les algorithmes ? »

À mon avis, il est possible de compter sur un genre analytique propre aux *STS*, labellisé « étude de laboratoire ». L'histoire est bien connue : durant les 1970, surtout aux États-Unis, à la confluence de l'ethnométhodologie⁵² et du « programme fort »⁵³, plusieurs analystes des sciences commencèrent à étudier les processus de formation d'énoncés certifiés en recourant aux méthodes ethnographiques immersives. Ironie de l'histoire, l'argument alors avancé afin de justifier le recours aux méthodes ethnographiques était très similaire à celui que je tente ici d'exposer (d'où l'extrême banalité de mon argument) : selon ces ethnographes, ni les études en termes d'effets, ni celles en termes de fonctionnements ne sauraient rendre compte des activités complexes qui constituent les processus de formation d'énoncés scientifiques⁵⁴. Une façon audacieuse de procéder au vu de l'importance des productions de l'activité scientifique pour le collectif était ainsi de rentrer dans les laboratoires et de documenter les situations dans lesquelles ces productions scientifiques prennent progressivement forme.

Populaires – bien que peu tentées – jusqu'au milieu des années 1990, les études de laboratoire ont eu droit à leur lot de critiques. Celles-ci condamnèrent tout d'abord l'absence de portée politique de leur

⁵⁰ Lucy Suchman, « Making Work Visible », *Communications of the ACM*, Vol 38, No 9, 1995, p. 61. Ma traduction.

⁵¹ À noter qu'il est crucial de rester attentif aux effets performatifs de ces dispositifs pédagogiques. Ce sujet est bien développé dans les études sociales sur la finance, notamment dans Donald MacKenzie, Fabian Muniesa et Lucia Siu (eds), *Do Economists Make Markets? On the Performativity of Economics*, Princeton, Princeton University Press, 2007.

⁵² Harold Garfinkel, Michael Lynch et Eric Livingston, « The Work of a Discovering Science Construed With Materials From the Optically Discovered Pulsar », *Philosophy of the Social Sciences*, vol. 11, n° 2, 1981, p. 131-158 ; Michael Lynch, « Technical Work and Critical Inquiry: Investigations in a Scientific Laboratory », *Social Studies of Science*, vol. 12, n° 4, 1982, p. 499-533.

⁵³ David Bloor, *Knowledge and Social Imagery*, London, Routledge and Kegan Paul, 1976.

⁵⁴ Voir par exemple Bruno Latour et Steve Woolgar, op. cit., pp. 15-42.

positionnement pratico-centré⁵⁵. Il y eu ensuite la tristement célèbre « guerre des sciences » qui dénonça le danger d'abandonner toute ambition réaliste. Plus récemment, c'est l'emphase sur l'incertitude et la fragilité des processus scientifiques qui fut l'objet de critiques, cette emphase ayant potentiellement opérée comme munition argumentative à celles et ceux désireux d'attaquer certains consensus scientifiques, notamment en ce qui concerne la question du dérèglement climatique⁵⁶. Difficile de dire si ces critiques ont eu un réel impact sur la désuétude des études de laboratoire. Quoiqu'il en soit, cette méthode d'enquête fut progressivement délaissée à partir des années 2000 au profit de positions plus directement activistes qui constitueront ce qui sera plus tard appelé le « tournant normatif » au sein des *STS*⁵⁷. Était-ce légitime de condamner les études de laboratoire pour leur difficulté à embrayer sur des questions politiques, pour leur frilosité quant aux questions rationalistes, ou pour leur propension à suggérer des controverses artificiellement maintenues ? Peut-être. J'ai toutefois l'impression qu'elles ne méritaient pas le désintérêt qu'elles ont récemment subi. Dans le sillage de Doing, il me semble en effet que les études de laboratoire souffrent plus généralement d'une « erreur comptable »⁵⁸ : même si les enquêtes initiales – citées et re-citées – ont bel et bien permis de considérer les sciences à l'aune de prises nouvelles, ces enquêtes n'ont pas, et de loin, exploité tout leur potentiel. En somme, malgré les débats qu'elles ont pu susciter au sein des *STS*, les études de laboratoire n'en sont qu'à leur début. Quoiqu'on en dise, les pratiques constitutives des faits scientifiques restent pour beaucoup méconnues, rendant leurs effets difficilement saisissables.

Allonger la liste des existants qui contribuent à la formation d'énoncés certifiés afin de mieux les apprécier : admettons qu'il s'agit là d'une appétence des études de laboratoire, genre analytique qui, en fait, n'en est qu'à ses débuts. À quoi cette méthode d'enquête pourrait-elle aboutir pour le cas de la *computer science*, cette jeune discipline aux contours flous de laquelle émane parfois, entre autres, des méthodes informatiques de calcul souvent appelées « algorithmes » ? Et comment ces éléments d'analyse pourraient-ils revitaliser la critique politique des algorithmes ?

C'est pour mieux se faire une idée du potentiel heuristique et politique des études de laboratoire pour l'analyse des processus algorithmiques que je vais à présent puiser dans les matériaux d'une étude menée au sein d'un laboratoire d'informatique⁵⁹. Ces résumés de terrain seront présentés chronologiquement et une partie de leur contenu sera considérée en fin d'article. Commençons donc par la première entrée et voyons où tout cela nous mène :

28 octobre :

Il faut bien l'admettre, je n'y vois pas grand-chose. Deuxième jour de terrain au sein de ce laboratoire d'informatique spécialisé dans le traitement de l'image digitale et mes observations restent désespérément fades. Je peux bien compter combien de *computer scientists* font partie du laboratoire, me renseigner sur d'où ils viennent et décrire, un peu, leur environnement architectural. Mais pour ce qui est de leurs pratiques, je m'y prends mal, mal, mal. Pas de blouse, de souris ou de paillasse ; pas d'effervescence qui pourrait me servir de couverture. Je suis visible, trop visible... L'endroit est silencieux, plutôt statique et les chercheuses et chercheurs sont la plupart du temps sagement installés en face de leurs moniteurs. Ils s'affairent, certes, mais ces affaires restent trop loin de moi. Comment m'insérer dans leurs agencements ?

3 Novembre :

⁵⁵ Langdon Winner, « Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology », *Science, Technology, & Human Values*, Vol 18, No 3, pp. 362-378.

⁵⁶ Edward Woodhouse, David Hess, Steve Breyman et Brian Martin, « Science Studies and Activism: Possibilities and Problems for reconstructivist agendas », *Social Studies of Science*, Vol 32, No 2, 2002, pp. 297-319 ; Stephen Turner, « The Third Science War », *Social Studies of Science*, Vol 31, No 1, 2003, pp. 123-149.

⁵⁷ Sur ce « tournant normatif », voir David J. Hess, « Local and Not-so-Local Exchanges: Alternative Economies, Ethnography, and Social Science », in Jeff Juris et Alex Khasnabish (eds.), *Insurgent Encounters: Transnational Activism, Ethnography, and the Political*, Durham, Duke University Press, 2013, pp. 151-170.

⁵⁸ Park Doing, « Give Me a Laboratory and I Will Raise a Discipline: The Past, Present, and Future Politics of Laboratory Studies », in Edward J. Hackett, Olga Amsterdamska, Michael Lynch et Judy Wajcman (eds), *The Handbook of Science and Technology Studies. Third Edition*, Cambridge (Mass.), MIT Press, 2008, p. 291. Voir aussi Malte Ziewitz et Michael Lynch, « It's important to Go to the Laboratory: Malte Ziewitz Talks with Michael Lynch », *Engaging Science, Technology, and Society*, Vol 4, 2018, p. 370.

⁵⁹ Les éléments qui suivent sont issus d'une enquête résumée dans Florian Jaton, « We Get the Algorithms of Our Ground Truths », op. cit.

C'est officiel, j'intègre un projet d'article à partir de lundi prochain : joie sur la Terre ! Vais-je finalement parvenir à rendre compte du travail de ces *computer scientists* ? J'ai bien fait de demander à la directrice du labo de ne plus me donner de privilège oisif : je vais devoir maintenant contribuer aux productions de son laboratoire. Mais serai-je à la hauteur ? Elle sait que mes connaissances sont infimes (bien que je prenne des cours pour me former en vitesse) et doit donc avoir une idée en tête. Mais un ethnographe participant à un projet d'article en traitement du signal, est-ce bien raisonnable ?

7 novembre :

La première séance de travail avec mes « désormais-collègue » est terminée. L'équipe est plutôt sympa, même si BJ, GY et CL se demandent certainement pourquoi je m'obstine à vouloir plonger dans leur univers en les aidant pour ce projet d'article. D'ailleurs, quel est l'objet de cet article ? Est-ce un nouvel *algorithme* ? Peut-être, un peu. Il faut dire que mes collègues parlent plus volontiers d'un nouveau « modèle », mais ils ne voient pas non plus d'objection à appeler ça un algorithme. Après tout, si j'y tiens... Et cet algorithme-à-faire, pourrait-il un jour être inclus dans un système plus large et de ce fait irriguer le quotidien de centaines de milliers d'utilisateurs ? C'est ce dont l'équipe rêve, en sachant bien que ce type d'« application » fructueuse d'un algorithme publié dans les *Proceedings* d'une conférence est extrêmement rare. Mais mine de rien, ça arrive parfois : ne parle-t-on pas de l'« algorithme de Viola et Jones » ? Surtout que notre projet flirte avec les problématiques de la compression d'images : « *This is were the beef is* », me dit d'ailleurs GY. Et la conférence en question n'est-elle pas sponsorisée par Google, Facebook et Microsoft ? En tous les cas, pour l'heure, tout ça est encore très lointain, le problème immédiat étant l'absence d'une chose que mes collègues nomment *ground truth*. Je comprends que c'est pour les aider à remédier à cette absence que la directrice m'a affilié au projet. Soit. Il est encore trop tôt pour comprendre ce qu'est véritablement une *ground truth*, terme pour le moins peu phonique. Pour l'instant, je vais me contenter de faire ce que BJ, GY et CL m'ont confié, à savoir identifier des plateformes Web possédant des images libres de droit et télécharger sur le server du labo les contenus qui feraient écho, de près ou de loin, à ce que notre modèle-algorithme devra être capable de détecter (en l'occurrence des différentiels de saillance). Des opérations en elles-mêmes peu palpitantes mais qui laissent tout de même entrevoir les coulisses de la mise en place d'un algorithme porteur de grands espoirs.

15 novembre :

Je n'ai pas pu m'empêcher de creuser cette histoire de *ground truth*, quitte à temporairement repousser mon travail de collecte de données sur server (mais le travail sera fait pour la prochaine séance). Et à force de creuser – notamment en en parlant durant les pauses à la cafétéria –, l'histoire devient plus claire : les *ground truths* sont partout ! Quel aveugle j'ai été : tout était pourtant là, dès le début sous mes yeux ! Plus une journée ne se passe sans qu'un membre du labo ne se plaigne ou ne s'enthousiasme pour une *ground truth*. « Nos résultats sont prometteurs : regardez ces performances ! », nous dit MS lors du *lab meeting* hebdomadaire. Sauf que je me rends maintenant compte que la possibilité d'existence de ces beaux résultats réside dans la présence d'une *ground truth* acceptée et utilisée par toute une communauté de recherche afin d'évaluer les performances de leurs modèles de détection d'ombre. « Difficile de vraiment savoir si le modèle fonctionne parce que pour ce sujet-là, les métriques d'évaluation sont en fait assez mauvaises », nous dit RK – toujours au même *lab meeting* – au sujet de sa nouvelle méthode de segmentation en super-voxel. Forcément, s'il n'y a pas consensus sur la façon dont une base de données *ground truth* définit les termes d'un problème à résoudre, il devient compliqué de comparer les performances des algorithmes censés résoudre ce problème... Mais également durant les précieuses pauses : « Tu connaîtrais pas des gens en faculté des Lettres par hasard ? », me demande NK. « Parce que j'ai besoin d'experts pour mettre en place mon modèle de reconnaissance manuscrite ». « Tu veux dire, pour faire une *ground truth* ? », tenté-je. « Oui c'est ça. Je pourrais le faire mais comme je ne suis pas du domaine, on risque de me tomber dessus... ». Encore une manifestation de cet objet « *ground truth* ». Je commence à prendre peur : j'espère qu'il s'agit là d'une extase momentanée. Si ce n'est pas le cas, mes épaules risquent de ne pas être assez larges...

12 décembre :

Le projet d'article se poursuit, non sans discussion. Nous disposons maintenant de 800 images en haute résolution qui quadrillent bien, du moins potentiellement, ce que l'algorithme sera supposé détecter. Mais est-ce suffisant ? Les jeux de données utilisés pour la formation d'algorithme en traitement de l'image n'ont-ils pas besoin d'être gigantesques ? Apparemment non, pas forcément. Surtout pour ce type de projet où il s'agit principalement de voir si un nouveau concept de détection « prend » au sein de la communauté de recherche (et potentiellement aussi dans l'industrie). Et de toute façon, le temps commence déjà à manquer : il faut absolument lancer l'expérience de *crowdsourcing* avant la pause de Noël, histoire de pouvoir traiter les résultats à la reprise du mois de janvier. Et comment concevoir ce questionnaire Web qui sera rempli par des hommes et des femmes du monde entier, pour beaucoup issus de l'Asie du Sud ? Il est important de ne pas se planter car les réponses que ces inconnus vont donner constitueront non seulement la base permettant d'entraîner le modèle – le *training set* – mais aussi ce qui permettra de l'évaluer – l'*evaluation set*. Sans compter qu'il faudra aussi rémunérer les participants à cette expérience. Pas grand-chose, certes, mais comme notre budget est d'environ 1000 EUR et qu'il nous faudra demander l'avis d'environ 300 personnes – c'est le standard actuel –, les tâches que nous pourrons leur demander ne doivent pas être trop chronophages. Et de toute façon, si les tâches sont trop alambiquées, les participants feront n'importe quoi (BJ en a déjà fait l'expérience). Toutes ces questions impactent sur le design de l'application Web qui se chargera de collecter les *logs* et les stocker sur nos serveurs. Ou plus précisément, toutes ces questions impactent sur GY à qui incombe la tâche de coder cette application Web en PHP et Javascript, deux langages de programmation fait pour les *Web designers*, pas pour les *computer scientists* (c'est du moins l'avis de GY) !

2 février :

Le projet a pris un peu de retard. Rien d'alarmant mais la finalisation du questionnaire-application Web par GY lui a pris plus de temps que prévu (il travaille aussi sur d'autres projets d'articles). Ce qui n'a pas été sans déplaire à CL, chargée avant Noël d'écrire le script MATLAB permettant de mettre en base de données les centaines de milliers de coordonnées géométriques $[x_1, y_1 ; x_2, y_2]$ bientôt extraites des opérations des participants au *crowdsourcing* (grâce précisément à l'application Web de GY). Et personne ne souhaite écrire ce genre de script durant les fêtes de fin d'année... Quoi qu'il en soit, nous disposons maintenant d'une jolie base de données nous permettant de visualiser le résultat du travail, sous-payé, de ces 300 individus (en fait un peu moins car j'apprends que GY avait également mis au point un système permettant d'identifier des *bots*, mettre de côté leurs résultats, et ne pas rémunérer leurs actions). Et c'est là que le travail barbant commence pour de bon. CL et GY ont déjà passablement œuvré pour le projet : ça sera donc à BJ – en première années de doctorat – et moi – en première année de doctorat *en sciences sociales* – de travailler ces résultats pour chaque image. Nous savions bien ce à quoi devait aboutir ce travail de traitement de données : on en avait discuté au début du projet, et les résultats de nos discussions avaient même orienté la conception de l'application Web. Mais nous découvrons maintenant la charge de travail rébarbatif que cela représente. Segmenter au pixel-près des formes complexes d'images naturelles afin d'obtenir des matrices en niveaux de gris représentant les valeurs relatives de saillance telles qu'exprimées par les 300 usagers : on en aura pour des semaines ! Pas étonnant que personne, dans la communauté de recherche en *image processing*, n'ait voulu cadrer ce problème en ces termes. À quand la sous-traitance de *ground truth* ? BJ me dit que ça existe déjà, notamment à l'Université de Hong Kong (où il a été formé). Et que ça n'est pas beau à voir.

4 mars :

Il nous aura fallu 1 mois pour finir de traiter toutes ces données. C'est que pour notre santé mentale, BJ et moi avons également fait d'autres choses, moins rébarbatives (notamment, pour ma part, un projet d'article sur ce projet d'article). Mais nous y sommes, officiellement : nous avons à notre disposition une nouvelle base de données *ground truth* rassemblant des données non transformées – les *input-images* – et leurs corolaires transformés – les *output-targets*. Et les termes du problème

que notre algorithme va devoir résoudre sont ainsi définis : il va devoir être en mesure de transformer au mieux les *input-images* en *output-targets*. Comme il est d'usage de le faire, nous avons aussi divisé aléatoirement la *ground truth* en deux ensembles : le *training set* qui servira à mettre en place l'algorithme, et l'*evaluation set* qui servira à tester ses performances et à les comparer avec d'autres algorithmes déjà publiés par d'autres laboratoires. Est-ce maintenant que les choses sérieuses commencent ? Les milliers de lignes de code, les mathématiques complexes, etc. ? Un peu. Il faut dire qu'on va tenter de réutiliser au maximum les modèles que BJ, GY et CL ont déjà mis au point : la date de soumission du manuscrit est en effet au 31 mars, et il faut se dépêcher... J'espère être en mesure de participer au tripatouillage du *training set* !

28 mars :

Les derniers tests effectués sur l'*evaluation set* sont plutôt concluants. En vertu des mesures statistiques usuelles en termes de *precision* et de *recall*, les résultats de notre algorithme sont meilleurs que ceux d'algorithmes déjà publiés. Encore heureux puisque notre algorithme est natif de la *ground truth* qui nous sert pour l'évaluation des résultats ! Tout le monde est d'accord, l'inverse aurait été invraisemblable. Et de toute façon, ce qui compte vraiment n'est-il pas au final la mise au point de notre nouvelle *ground truth* ? Car si le papier – que GY est en train de finir de rédiger dans l'urgence – est retenu, toutes celles et ceux qui voudront battre les performances de notre algorithme devront se référer à notre article qui présente aussi une nouvelle *ground truth* : citations garanties ! Pour autant qu'il soit retenu bien sûr. Et pour ce genre de grande conférence, la moyenne est d'environ 80% de refus. En tous les cas, il est là, sous mes yeux, sous forme de 5 fichiers Matlab renfermant des milliers de lignes de code : notre (prototype d') algorithme. Je comprends maintenant un des intérêts du formalisme mathématique : infiniment plus simple à lire et à présenter. Ici, même le pseudocode ne ferait pas l'affaire.

19 juin :

La réponse est tombée il y a de ça une semaine : l'article n'est pas retenu. Triste. Mais pour quelle raison ? BJ m'a envoyé le compte rendu des *reviewers* et on se voit pour en discuter. Le papier est irréprochable au niveau mathématique, ce qui est une bonne nouvelle pour GY qui s'est chargé de l'analyse. Mais les trois *reviewers* sont d'accord pour dire que le point de l'article – qui consiste à dire que notre algorithme *surpasse* les performances des algorithmes déjà publiés – ne tient pas. La faute à nos évaluations de performances qui ne testent les algorithmes déjà publiés que sur notre nouvelle *ground truth*, donnant ainsi un avantage certain à notre algorithme. Il est vrai que l'argumentation du papier était peut-être mal montée. Était-il vraiment important de miser sur les performances de notre algorithme alors que ces performances n'étaient pas vraiment commensurables avec les autres algorithmes ? Comme notre algorithme émanait du *training set* d'une nouvelle *ground truth*, il était de base conçu pour résoudre un nouveau problème. Mais beaucoup d'autres auteurs ont fait ça avant nous, et ont été publiés, à tel point que cette « erreur logique » pouvait légitimement être prise comme une sorte d'obligation. D'autant qu'en ces temps de montée en puissance des techniques d'apprentissage automatique, les critères de sélection d'articles gravitent de plus en plus autour de la performance. Soit, en vertu de leur obnubilation pour les performances chiffrées, les *reviewers* ont vu juste. Mais ne pouvaient-ils pas au moins saluer le fait de se lancer dans une redéfinition de problématique ? Ne pouvaient-ils pas admettre que la communauté de recherche en *image processing* ne peut pas continuer à travailler éternellement avec de vieilles *ground truths* ? Car – pour BJ – ça n'est pas très compliqué de battre les performances d'algorithmes déjà existant : si vous lui donnez l'infrastructure nécessaire (qui par ailleurs coûte très cher), il parviendra à mettre en forme un algorithme d'apprentissage automatique et l'affaire sera dans le sac. C'est d'ailleurs ce que font souvent celles et ceux qui sont publiés pour ce genre de grandes conférences. Ah les beaux résultats ! Mais pour quel problème ? Celui mal défini par cette *ground truth* qui date de 2006 ? De bons algorithmes qui résolvent de mauvais problèmes définis par de mauvaises *ground truths*. De mauvais algorithmes, en somme. Ces *reviewers* n'ont-ils pas encore compris que pour avoir de meilleurs algorithmes, il nous faut de meilleurs *ground truths* ? Ils ont trop vite oublié leurs années doctorales où toutes ces questions sont des problèmes quotidiens. À force de n'exiger que

des résultats de performances fiables, ils s'enferment dans des problèmes insipides. Mais bon, selon BJ, tout n'est pas perdu. On resoumettra l'article en modifiant la problématique. Car il est vrai qu'en l'état, le document n'est pas facile à suivre.

Malgré sa vocation avant tout illustrative, ce bref plongeon au sein de la science informatique en train de se faire donne à sentir l'épaisseur ontologique de ces méthodes informatiques de calcul appelées « modèles » ou « algorithmes ». Car l'algorithme de CL, GY et BJ existe dès le début du projet, en novembre, sous la forme d'espérances et de désirs. Mais au 28 mars, n'existe-t-il pas davantage ? Oui, car à cette date, il a bel et bien été assemblé et peut dès lors effectivement détecter des différentiels de saillance au sein d'images numériques. Mais le 19 juin, la non-publication de l'article faisant état de ses composantes et de ses performances n'a-t-elle pas ensuite amoindri son existence ? Certainement, car à ce stade, il en est limité à 5 fichiers Matlab stockés dans le serveur interne du labo, alors même qu'une fois associé à son article, il aurait pu devenir capable d'intéresser un de ces employés de Microsoft parcourant les grandes conférences en traitement du signal à la recherche de nouvelles idées et talents. Difficile ainsi de réduire l'existence d'un algorithme à son squelette informatique : c'est enveloppé d'une multitude de documents, d'évènements, d'objets et de désirs que la réalité de ce type d'entité gagne – ou perd – en intensité.

Cet élément pointe lui-même vers les plus curieux objets de cette brève incursion ethnographique : les *ground truths*, ces bases de données référentielles qui mettent en lien des *input-data* et des *output-targets*. Si ces bases de données semblent capables de définir les termes et les solutions de problèmes à même d'être résolus de façon computationnelle, elles ne sont pas toujours données : pour le cas qui nous intéresse ici, GY, CL, BJ et moi-même avons dû la façonner durant un processus complexe engageant un certain nombre d'habitudes, de ressources et de valeurs. Et il semble que cette démarche n'aurait pas pu être complètement évitée : cette *ground truth* apparaissait comme condition préalable au développement d'un nouvel algorithme. Ceci constituait même l'intérêt du projet (du moins en juin, pour BJ) : proposer une *ground truth* inattendue de laquelle pourrait dériver des algorithmes inattendus. Pour l'ethnographe, mais aussi peut-être – pourquoi pas ? – pour les *computer scientists* se dessine alors un continent fascinant et pour l'heure quasi inexploré : l'étude des tenants et aboutissants de ces bases de données *ground truths*, véritables matrices historiques mais contingentes de bon nombre d'algorithmes. Et là, précisément, germe également un potentiel politique, impossible à présumer sans se rendre sensible à la manufacture des algorithmes : si bon nombre d'algorithmes dérivent de *ground truths*, se saisir des *ground truths* revient également, dans une certaine mesure, à se saisir des algorithmes. D'autant plus que ce type de bases de données semble concerner plusieurs groupes de personnes ayant a priori peu de choses en commun : pour notre cas, la formation de la *ground truth* liait, entre autres, des doctorants en traitement du signal soucieux de proposer des algorithmes innovants, des travailleurs du clic enrôlés par les nouvelles formes de précarité, des professeurs bien implantés à qui il incombe d'opérer des sélections d'articles pour le bon déroulement d'une conférence sponsorisée par les géants du numérique et un ethnographe désireux de rendre compte de la constitution des processus algorithmiques. Et c'est peut-être par l'identification de ce type d'affection transversale que les discussions sur les algorithmes pourraient embrayer sur des affaires capables d'être saisies politiquement ; c'est peut-être rendant visible ce qui nous relie aux processus algorithmiques que ceux-ci commenceront à être attentivement considérés. Reste encore, bien sûr, à persévérer dans l'enquête.