LA CONFIANCE À L'ÈRE DU NUMÉRIQUE, DOUEIHI M., DOMENICUCCI J., DIR., 2018, PARIS, BERGER LEVRAULT/RUE D'ULM, COLL. AU FIL DU DÉBAT

Fabien Blanchot

ARIMHE | « RIMHE : Revue Interdisciplinaire Management, Homme & Entreprise »

2020/4 n° 41 | pages 103 à 116

ISSN 2259-2490

Article disponible en ligne à l'adresse :

https://www.cairn.info/revue-rimhe-2020-4-page-103.htm

Note de lecture

La confiance à l'ère du numérique, Doueihi M., Domenicucci J., Dir., 2018, Paris, Berger Levrault/Rue d'Ulm, Coll. Au fil du débat

Fabien BLANCHOT[30]

La confiance à l'ère numérique est le titre d'un ouvrage collectif dirigé par Milad Doueihi, directeur scientifique du Berger Levrault Institut, et Jacopo Domenicucci, professeur de philosophie. Il a été co-écrit avec onze autres contributeurs académiques. Sont successivement abordées la confiance connectée, c'est-à-dire la confiance interpersonnelle médiée par des interfaces numériques (partie 1), la confiance dans les artefacts, les agents artificiels, les dispositifs numériques (partie 2) et la manière dont évoluent les rapports de confiance dans un environnement numérique permettant stockage et vérification d'informations sur l'autre et les échanges passés (partie 3). De manière transversale, l'ouvrage s'intéresse aux mutations de la confiance induites par la « coopération numérique ».

La préface révèle l'origine de l'ouvrage qui est le résultat d'un partenariat entre le monde de la recherche et celui d'une entreprise : Berger Levrault, née à l'époque de Gutemberg et qui a migré dans les années 80 de l'imprimerie vers l'édition de logiciels. Ce partenariat a comme point de départ une problématique majeure pour l'entreprise : comment évoluer au contact de données sensibles « sans susciter un niveau élevé de défiance de la part des futurs usagers » (p.10) à l'égard des outils numériques développés ?

L'introduction générale, foisonnante mais décousue, établit une distinction claire entre sécurité numérique et confiance numérique. La première signifie que les canaux où circulent l'information sont sécurisés face à des attaques externes et que les messages transmis sont fiables. L'accent est mis sur le système technique et l'objectif est une suppression du risque. Cette dernière est illusoire dans la mesure où l'on ne connaît jamais tous les risques et que les concepteurs peuvent eux-mêmes manquer de fiabilité ou de bienveillance. La seconde est une relation entre agents dans un contexte de risques. Elle est nécessaire pour pallier les insuffisances de la sécurité et faciliter l'action. Sécurité numérique et confiance numérique sont donc distinctes et nourrissent des

[30] Professeur des universités, Université Paris-Dauphine PSL, DRM/CNRS - fabien.blanchot@dauphine.psl.eu

relations qui ne sont pas immédiatement évidentes et que l'ouvrage se propose d'explorer. Ainsi, il est d'emblée affirmé que la recherche de sécurité peut miner la confiance et que la confiance ne peut pas être considérée comme produite par la sécurité car elle a des aspects non cognitifs, émotionnels et normatifs, distincts du calcul rationnel des risques.

1. Une première partie consacrée à la confiance connectée

Dans le chapitre initial de la première partie, Laurent Jaffro, professeur de philosophie morale, défend une thèse qui requiert un regard approfondi sur le concept de confiance, en l'occurrence que les interactions en ligne sont confrontées à un problème de confiance systémique et épistémique, et non seulement à un problème de confiance pratique. La confiance pratique a trait à la fiabilité des agents (de leur conduite). Sa forme dominante, correspondant à une définition fréquente dans différents champs des sciences sociales, est la confiance décidée ou confiance-pari. Elle est interpersonnelle et se manifeste par une acceptation de dépendance à l'égard de la conduite future d'autrui. Elle n'existe que sous la condition d'une prise de risque, donc d'une action. Elle est attente normative plutôt que prédictive, et génère, selon qu'elle est confirmée ou déçue, gratitude ou sentiment de défection de celui qui trahit la confiance donnée. Elle se distingue d'autres formes de confiance parfois mentionnées dans la littérature, qui ne requièrent pas une action, sont moins volontaires et plus implicites, comme l'absence de suspicion ou comme la confiance assurée ou confidence, associée au danger mais pas à la prise de risque dans l'action (par exemple, la confiance dans le fait qu'il ne pleuvra pas). La confiance systémique concerne les environnements mêmes des interactions et pas seulement les interactions. Elle est donc intéressante pour étudier les situations où les agents ont affaire à des environnements numériques (des interfaces) et non pas directement à des personnes, les situations où les environnements sont au premier plan et les individus au second plan. Elle englobe familiarité (avec l'interface, dans les environnements numériques) et confiance assurée (à l'égard de cette interface, par exemple la fiabilité de l'information donnée par une encyclopédie en ligne ou la fiabilité de la confidentialité). La confiance épistémique a trait à la fiabilité des apparences (de la véracité de ce qui est dit, d'assertions). C'est une « forme de confiance dont le contraire est le scepticisme à l'égard de la connaissance, et qui ne s'exerce pas spécifiquement à l'égard d'autres agents, mais à l'endroit des apparences » (p.42). Fort de ces concepts, Laurent Jaffro alterne éléments analytiques et illustratifs de la confiance en ligne. Il montre le caractère hégémonique de la confiance épistémique pour tout ce qui touche à l'information en ligne, par exemple les encyclopédies : le problème est celui de la crédibilité (fiabilité des apparences) plutôt qu'une question de confiance décidée ou de sécurité. Il adhère à l'idée qu'un univers de confiance n'est pas un univers sécurisé, même si « un certain degré de sécurité de l'environnement est indispensable,

sans quoi la méfiance épistémique devient radicale et la confiance pratique ne peut se déployer » (p.49), et même s'il est un domaine où sécurité et confiance sont très liées, celui de la confidentialité et de l'identification. Il rappelle les obstacles habituellement associés à la confiance en ligne (le manque d'identité ou l'anonymat, le manque de caractéristiques personnelles ou la désincarnation et l'indétermination des normes et des rôles sociaux et professionnels), les moyens avancés pour les contourner (contrôle d'accès, traçage de l'identité, surveillance) et les discute. Il traite de la possibilité d'existence de la confiance pratique dans les interactions en ligne, où l'historique des interactions peut être absent, mais c'est aussi souvent le cas dans le « monde naturel », où « l'historique futur ne va probablement pas se constituer » empêchant l'adoption d'une stratégie de type « coopération-récitpocité-pardon » (*tit for tat*), en référence aux travaux de Robert Axelrod[31], pour installer la confiance, et où l'interaction physique est absente, qui n'empêche toutefois pas la réputation de jouer un rôle. Il critique le rapport étroit établi entre confiance et fiabilité, la seconde (*trustworthiness*) étant souvent considérée comme le déterminant principal de la première, et insiste sur l'importance de la signalisation dans la dynamique de la confiance : informer l'autre de la confiance qu'on lui donne, pour l'encourager à être digne de confiance. Ainsi, même si l'on n'a pas connaissance des vertus de l'autre (de sa « fiabilité substantielle »), on peut néanmoins lui donner sa confiance de manière rationnelle, la signalisation permettant de jouer sur le désir d'estime ou sur la sensibilité à la réputation du *trustee*. Laurent Jaffro souligne par ailleurs le rôle clé des indices ou des preuves dans la dynamique de la confiance, tout en pointant les difficultés de les obtenir (historique des relations, contacts physiques) dans les interactions en ligne et la nécessité de passer par des dispositifs « simulés » (par exemple, la réputation en ligne) qui requièrent, pour que la confiance pratique dans la fiabilité d'un agent puisse se développer, une confiance (épistémique) dans la fiabilité d'une preuve apparente. Se pose en effet la question de la fiabilité des systèmes mis en place (par exemple, un système de réputation en ligne). Et, encore une fois, il est rappelé que la sécurité ne peut ici remplacer la confiance, parce que la sécurité ne peut jamais être totale (il peut y avoir défaillance ou défection de ceux qui élaborent les systèmes de sécurité, de ceux qui surveillent le système, etc).

Le second chapitre, écrit par Thomas Simpson, professeur associé de philosophie et de politiques publiques, peut être considéré comme une extension du précédent. Il traite en effet des conditions et de la manière dont la communication par téléprésence, qui permet d'être présent auprès des autres mais à distance, peut conforter la « confiance reposant sur les indices », c'est-à-dire fondée sur des éléments de preuve pour juger qu'une personne est digne de confiance et que l'on peut donc considérer comme rationnelle. Les deux premières sections posent le décor. La première présente les technologies de

[31] Notamment : Axelrodt R. (1984), *The evolution of cooperation*, New-York, Basic Books.

téléprésence et montre leur évolution qualitative continue et en devenir : de la communication unidirectionnelle, avec les talkies-walkies, à la communication bidirectionnelle, avec les téléphones ; de la communication scripturale, comme le télégramme, à la communication verbale et non verbale, avec la vidéo ; de la communication à distance à l'action à distance, avec la robotique et les drones. L'évolution des technologies concerne la fidélité de reproduction du son et de l'image, la vitesse de transmission, la possibilité de communication réciproque et la précision de l'action à distance. La seconde section traite de la confiance fondée sur les indices. Il est rappelé qu'elle n'est qu'une forme de confiance parmi d'autres, la confiance pouvant faire référence à un état mental affectif, conatif et/ou cognitif ou encore à une action. Mais la détention d'indices sur la fiabilité d'autrui conditionne bien la confiance dans certaines situations, pour des raisons « pratiques » (en référence à la confiance pratique), les cas du baby-sitting et du recrutement sont notamment cités, ou « épistémiques » (en référence à la confiance épistémique), le cas du témoignage de locuteurs dans le cadre d'un audit est cité. Bien que ce ne soit pas central dans les développements faits par l'auteur, on peut noter que le niveau du risque encouru par le *trustor* en cas de confiance n'est pas mentionné comme un facteur explicatif du niveau de besoin d'indices, cette évacuation du niveau du risque étant un constat fréquent[32] et nous semble étonnant dans la littérature sur la confiance. Dans les situations où la recherche d'indices fonde la confiance, la communication est considérée comme centrale car elle permet de relever des indices, comme c'est le cas, par exemple, quand un *trustee* affirme quelque chose (c'est un indice d'un engagement de sa part à ne pas mentir même s'il ne fait pas de promesse explicite) ou fait une promesse, s'il appartient à une communauté où la fiabilité prévaut. Mais certains considèrent que la communication médiée numériquement ne permet pas de jouer ce rôle et, donc, ne permet pas le développement d'une confiance rationnelle. La section 3 présente leurs arguments, qui sont ensuite contestés. Un premier argument est que la communication médiée numériquement ne permet pas de relever les indices d'apparence (expressions, gestes, mots, apparence physique d'une personne) qui seraient importants dans la construction d'une confiance rationnelle. Cet argument, qui date du milieu des années 90, est de moins en moins vrai avec les technologies actuelles. Un second argument est que le contact physique, que ne permet pas la téléprésence, serait un élément (un indice) important pour construire une confiance dans un contexte où d'importants intérêts du *trustor* sont en jeu. Simpson conteste (contre-exemple à l'appui) que ce soit une condition toujours nécessaire sans pour autant rejeter l'hypothèse que « les formes de téléprésence ne soient pas propices à établir et entretenir des formes importantes de confiance » (p.77). Les sections qui

[32] Mais pas systématiquement : voir notamment Coleman J. (1990), Foundations of Social Theory, Harvard University, Press.

suivent s'écartent, sauf pour y revenir de façon conclusive, de la réflexion sur le lien entre confiance et téléprésence, pour se concentrer sur les usages de la téléprésence et leur signification. La section 4 insiste sur le caractère plus ou moins intrusif ou révélateur des technologies, selon le degré de téléprésence qu'elles permettent, ce qui n'est pas sans conséquence sur leur usage. Ainsi, le téléphone apparaît préféré quand il s'agit de se protéger ou de protéger autrui, par exemple quand il faut « cacher une partie de la vérité ou mentir pour arrondir les angles » ou dissimuler ceux qui sont à proximité physique du locuteur (p.79). En revanche, la vidéo est privilégiée lorsque l'intention est de renforcer l'intimité qui lie déjà des interlocuteurs. La section 5 défend quatre idées, étayées à des degrés divers. D'abord, « il y a un « pourquoi » pour chaque support de communication » (p.81). Le choix d'un support de communication dépendrait de la nature de la relation préexistente entre les personnes et de l'objet de l'échange. Il révélerait aussi une partie de l'acte illocutoire (l'intention) du locuteur. Ainsi, réunir toute la famille autour d'un ordinateur pour un appel Skype à un fils resté à l'étranger un jour de Noël serait une manière de réaffirmer l'intimité familiale. Ensuite, et de façon corrélative, « des informations en quantité n'équivalent pas forcément à une communication de qualité », (p.81), « une moindre quantité d'informations sert certains objectifs communicationnels mieux qu'une plus grande quantité ne le ferait » (p.82), ce qui conduit à une critique de la théorie de la richesse des médias qui est associée, de manière réductrice, à la prédiction que l'on utilise des formes de communication moins riches uniquement parce qu'on y est contraint. Par ailleurs, ce n'est pas parce que les technologies de téléprésence ne simulent pas parfaitement un contact en face à face qu'elles ne permettent pas une communication efficace. L'articulation de ces idées avec la confiance est finalement révélée : « la communication par téléprésence renforce la confiance entre ceux qui sont déjà proches, en montrant qu'ils sont toujours disposés à se révéler à l'autre » (p.84) ; réciproquement, « le degré de révélation et d'intrusion que cela implique est tel qu'il est inapproprié de le demander à un inconnu, à moins d'avoir une excellente raison » (p.85), d'une autre nature, généralement, que la construction ou le développement de la confiance. Enfin, « la télépresence est un moyen d'entretenir certaines formes de confiance. Elle est moins propice à la créer » (p.86). Des assertions intéressantes, mais qu'il semble nécessaire de considérer comme des conjectures plutôt que comme des résultats empiriques solides.

2. La confiance dans les dispositifs numériques traitée dans la partie 2

Le premier chapitre de la partie 2 est rédigé par Christopher Thompson, professeur associé de philosophie. Il pose la question de la possibilité de faire confiance aux artefacts. La définition retenue de la confiance est celle que l'auteur identifie dans la plupart des théories de la confiance et qu'il qualifie « d'authentique » ou de « traditionnelle » : une relation tripartite entre deux agents et une action, qui implique

que l'on compte sur quelqu'un (reliance) mais qui repose aussi sur une attitude participative du *trustor*, c'est-à-dire une attente normative qui génère un sentiment de gratitude ou de trahison, selon le comportement du *trustee*. L'artefact est défini comme un objet fabriqué intentionnellement afin de remplir une fonction particulière, ce qui le distingue de l'objet naturel et du naturefact défini comme étant un objet naturel auquel est assigné une fonction par un auteur qui le sélectionne à cette fin (par exemple, le rondin de bois pour servir de siège). Son utilisation est associée à une plus grande capacité d'action (agentivité), soit qu'il permet d'effectuer des actions autrement physiquement impossibles, soit qu'il permet d'agir à distance ou dans la durée (une boite de conserve, par exemple). Les artefacts sont liés à la confiance de deux manières : honorer sa confiance peut requérir l'usage d'artefacts et ces derniers peuvent être un objet de confiance. Les concepts étant posés, l'auteur présente les six significations qu'il est possible d'attribuer au fait qu'on puisse déclarer faire confiance à un artefact : une foi totale dans le fait qu'il va fonctionner d'une certaine façon, sans sentiment de trahison en cas d'observation contraire, ce qui ne constitue pas de la confiance « authentique » ; un anthropomorphisme, mais rares seraient les adultes qui commettraient cette erreur d'attribuer une réelle agentivité à des objets inanimés ; une forme d'heuristique, un raccourci pour faciliter notre raisonnement, tout en étant conscient que ce n'est pas de la confiance « traditionnelle » ; une confiance « authentique », fondée sur le fait que l'artefact serait un agent autonome doté d'états mentaux, c'est-à-dire capable de tenir compte de ses propres intérêts tout comme de ceux du *trustor*, et, surtout, d'autonomie réelle, c'est-à-dire capable de construire ses états mentaux indépendamment de ses concepteurs, mais aucun artefact ne disposerait d'une telle autonomie à ce jour ; une confiance « traditionnelle » envers les concepteurs et/ou fabricants de l'artefact, qui serait fondée sur les affirmations du fabricant (donc, d'ordre épistémique) ; enfin, une confiance « authentique » accordée aux artefacts et non réductible à ses concepteurs / fabricants, thèse défendue par ceux qui considèrent qu'on peut à la fois compter sur un artefact (reliance) et avoir une attente normative, qui prendrait la forme du sentiment d'être en droit de compter sur ce pourquoi l'artefact a été conçu. Mais Thompson conteste cette thèse. Pour ce faire, il fait un détour par les naturefacts : quand un arbre couché, faisant office de pont, vient à céder quand on l'utilise, il est peu probable qu'on se sente trahi par lui. Si, donc, on fait uniquement confiance aux objets qui ont un auteur, c'est que « c'est aux auteurs et non pas aux objets eux-mêmes, que nous faisons en fait confiance ». Au total, on dispose d'une analyse stimulante de la signification que l'on peut donner à l'usage du terme confiance pour faire référence aux relations entre les individus et des artefacts, en particulier les dispositifs numériques.

Le second chapitre, rédigé par Alexandre Mallard, Cécile Méadel, et Francesca Musiani, tous trois sociologues, s'intéresse à la confiance dans le système de monnaie

électronique décentralisé Bitcoin, utilisant la technologie « pair à pair » (P2P). Le Bitcoin est une approche qui se passe d'intermédiaires dédiés à l'échange de monnaies et à la régulation et, donc, de garants institutionnels de confiance (ou tiers de confiance). Ceux-ci seraient remplacés par une « confiance distribuée ». Le terme « distribué » est utilisé en référence au fait que le système Bitcoin est décentralisé au sein du réseau de ses utilisateurs (de leurs ordinateurs) dont certains (appelés mineurs) vérifient les transactions effectuées (conservées sous une forme agrégée appelée « bloc » et qui s'intègre elle-même dans une « chaine de blocs » dont le « stockage » est réparti entre tous les membres du réseau) et contribuent, de ce fait, à la construction collective de la confiance. Aucune définition précise *a priori* n'est donnée de la confiance, les auteurs préférant étudier les formes qu'elle prend en s'appuyant sur un corpus de texte d'experts. Quatre dimensions de ce corpus sont examinées, considérées comme liées à la mise en place de la confiance distribuée et correspondant à des définitions différentes de ce que représente la confiance. La première traite de la relation entre confiance et monnaie et compare le système Bitcoin avec la monnaie-marchandise et monnaie de crédit, ces deux derniers systèmes étant associés à deux conceptions distinctes de la confiance, l'une fondée sur les objets et l'autre sur les institutions sociales. Il est défendu que le système Bitcoin se distingue des deux autres systèmes monétaires, tout en s'en rapprochant sur certains points et que la confiance qui lui est portée repose sur son code P2P plutôt que sur la crédibilité des personnes et des institutions, élément de singularité de ce système monétaire. La seconde dimension traite de la relation entre confiance et infrastructure technique. D'abord, le réseau P2P sollicite la confiance de ses utilisateurs, parce qu'ils doivent accepter de partager une partie de leurs ressources informatique et reconnaître le rôle de « référentiel de la valeur » du système, ce qui repose notamment sur sa fiabilité technique. Ensuite, il sollicite leur confiance parce que l'architecture décentralisée, qui est une garantie contre les attaques et les erreurs, est néanmoins, de par sa complexité, sujette à de fausses rumeurs, à des informations erronées et à de la spéculation, et peut susciter la venue, en raison de l'anonymat qu'elle permet, d'utilisateurs souhaitant masquer des activités économiques frauduleuses. Enfin, le fait que tous les utilisateurs ne jouent pas le même rôle dans le système (certains co-développent le code ou contribuent à la vérification des transactions, d'autres s'en tiennent à des transactions) aurait des « conséquences structurelles sur la distribution de la confiance dans le Bitcoin » (p.128) : l'adoption de ce système difficile à comprendre par un participant « lambda » serait un acte de foi, requérant confiance envers le code, envers l'équipe de quelques développeurs inconnus, envers le protocole et envers les valeurs du système ; l'engagement des vérificateurs des transactions (les mineurs), encouragé par une récompense en bitcoins pour tout bloc généré correctement, serait une source de confiance pour tout le réseau. La troisième dimension traite de la relation entre confiance et le fait que le système Bitcoin combine anonymat et transparence. Le système permet

de relier les acteurs d'une transaction d'une manière fiable et anonyme, grâce à la cryptographie, à un dispositif distribué de vérification des blocs et l'enregistrement des transactions dans tous les ordinateurs du système. Ces trois éléments jouent conjointement le rôle d'un tiers de confiance. Mais la transparence induite par l'enregistrement distribué des transactions pose des défis discutés par les auteurs. Enfin la quatrième dimension traite de la relation entre confiance et intégration du Bitcoin à des systèmes monétaires et économiques déjà présents. La convertibilité du Bitcoin le soumet à des fluctuations qui peuvent affecter la confiance envers cette monnaie, suscitant le développement de dispositifs renforcés de régulation. *In fine*, cette contribution, si elle dessine les contours des formes de confiance qui peuvent être associées aussi système bitcoin, se révèle avant tout intéressante pour comprendre le système Bitcoin et ses enjeux.

Le troisième et dernier chapitre de la partie 2, rédigé par Boudewijn de Bruin, professeur d'éthique, et Luciano Floridi, professeur de philosophie et d'éthique, s'intéresse à la manière de traiter les risques associés à l'usage du *cloud* (informatique en nuage, informatique dématérialisée, outils en ligne) en adoptant une perspective éthique. La question de la confiance n'est ici pas du tout traitée (une seule occurrence du mot confiance dans le texte, dans une citation empruntée). Bien que l'article ne soit pas pleinement en cohérence avec le titre de l'ouvrage, il n'est néanmoins pas dénué d'intérêt, notamment si l'on considère que les risques mentionnés sont des obstacles à la confiance en ligne, et dispose d'une qualité appréciable : sa grande compréhensibilité, pour faire référence au concept d'Habermas cité par les auteurs. Avec le *cloud*, les ressources informatiques deviennent des services à la demande, ce qui comporte de nombreux avantages potentiels pour leurs usagers (collaboration plus facile sur des projets collectifs, besoin d'investissement réduit, baisse de coût, pas de besoin d'installer, de configurer, de mettre à jour des logiciels, compatibilité, accès à de grosses puissances de calcul, etc.) qui contribuent à expliquer le succès de l'informatique dématérialisée. Mais le *cloud* comporte aussi des risques (cyberattaques, espionnage, piratage, interruption des services) qui, en raison d'une vision réduite de la valeur de la propriété, requiert une pression didactique et restrictive sur les utilisateurs du *cloud* (les « *clouders* »). Le constat premier est que les nouvelles générations ont une attitude indulgente concernant le plagiat et le piratage en ligne : les informations disponibles sur internet sont vues comme « quelque chose qui est là pour être partagé et librement utilisé » (p.151) plutôt que comme la propriété d'autrui devant être respectée. Ce choix en faveur d'une liberté accrue signifie aussi renoncement (à des droits de propriété) et il est donc important, d'un point de vue éthique, qu'il soit aussi éclairé que possible. Pour cette raison, les « *clouders »* devraient rechercher eux-mêmes des informations et faire preuve, dans cette démarche, d'impartialité intellectuelle (ouverture), de sobriété intellectuelle (prise de recul sans tomber dans le scepticisme) et de courage intellectuel

(recherche active). Mais même si ces trois attitudes vertueuses étaient respectées, elles pourraient s'avérer insuffisantes pour un choix éclairé, si les acteurs de la filière ne communiquent pas correctement sur les avantages et inconvénients de leurs offres et ce de manière intelligible et accessible. Il est donc recommandé que ces acteurs fassent preuve « d'interlucency » (p.154) dans leur communication, une vertu épistémique axée sur autrui consistant à adapter les informations fournies au public ciblé et à vérifier activement que ce public les comprend (*via* des présentations détaillées et claires, des foires aux questions, etc.). Il est aussi recommandé que ça donne lieu à une réglementation contraignante. En revanche, il est préconisé que les gouvernements n'interfèrent pas dans le développement des technologies des sociétés d'hébergement et des fournisseurs de services en nuage, conformément au principe de neutralité technologique, parce qu'il est difficile de mesurer les risques *ex ante* et parce que cela pourrait étouffer la créativité et l'innovation. Les restrictions devraient plutôt concerner les entreprises utilisatrices de ces services, lorsqu'elles font courir des risques importants à leurs clients. Par exemple, les avocats ne devraient pas pouvoir stocker certaines données sur le *cloud*, dès lors que cela fait courir un risque à leurs clients (sans même que ces derniers le sachent). On ne parle pas de confiance dans ce chapitre, mais on souligne *de facto* le risque de trahison de confiance, sinon d'abus de confiance.

3. En partie 3, les défis de la société numérique

La dernière partie de l'ouvrage traite des défis de la société numérique. Un premier chapitre, rédigé par Tobias Matzner, professeur de *media studies*, aborde le thème de la vie privée (*privacy*) dans les médias numériques, en adoptant une approche arendtienne. L'auteur commence par rappeler que les médias numériques ont d'abord été perçus comme une source d'autonomie, un moyen de s'inventer une personnalité et une vie virtuelle complètement indépendantes de la vie réelle et, ainsi, une manière de s'écarter d'identifications discriminatoires ou insatisfaisantes de la vie réelle. Mais dans les faits, on peut constater une perméabilité forte entre les deux, tout ce que nous disons et faisons dans la vie réelle se retrouvant sur internet et dans des bases de données diverses, ce que renforce le développement des objets connectés. Les médias numériques sont ainsi une source d'hétéronomie « où les autres décident pour nous ce qui nous concerne, déterminant par-là même qui nous sommes » (p.176). La protection de la vie privée est censée prévenir cette hétéronomie, mais les théories modernes de la *privacy* considèrent que l'individu ne décide jamais totalement avec qui communiquer et quelles données révéler : il existe une menace « d'effondrement de contexte », c'est-à-dire non seulement un risque de dévoilement d'informations intimes et privées, mais aussi l'accès de la présentation de soi à d'autres publics que ceux visés. En outre, l'utilisation consciente et intentionnelle des médias numériques par les individus, leur prise de distance critique vis-à-vis des influences extérieures et leur mise en scène d'eux-mêmes

contrôlée, telles que prétendues par les théories modernes de la *privacy*, sont contrariées par les normes sociales imposées par les outils utilisés (les formulaires pour gérer son profil), par les interactions avec les autres (nos actions en ligne et les commentaires des membres de notre réseau, par le contenu automatique d'actualité que créé l'algorithme du site (Facebook, par exemple) en fonction de notre historique d'interaction et par notre identité construite que vend le site aux courtiers en données et aux agences de publicité. Les travaux d'Hannah Arendt sur les dangers de l'hétéronomie et la manière de contourner cette menace sont alors mobilisés, pour suggérer la manière dont on peut penser le rôle de la vie privée. Pour l'auteur citant Arendt[33], la solution au totalitarisme ne se trouve pas dans « l'*homo faber*, être humain souverain et conscient qui met en pratique ses idées en s'aidant d'outils et de ressources » (p.186), mais dans le pardon, en tant qu'il change la manière dont une personne apparaît aux yeux des autres. Si la vie privée se situe à un niveau social et est revendiquée par des normes, Matzner y voit une fonction analogue au pardon : elle « atténue l'influence que les autres apparences d'une personne peuvent avoir sur son apparence actuelle » (p.191). L'apparence étant entendue ici comme : « quelque chose qui est vu et entendu par les autres ainsi que par nous-mêmes » (p.188). Autrement dit, la vie privée « régule quelles apparences devraient être ignorées et inaccessibles dès le départ dans la mesure où elles influencent l'apparence actuelle » (p.191) en permettant que des choses du passé ne comptent plus dans l'apparence d'une personne. En ce sens, la vie privée n'implique pas que le sujet puisse librement mettre en scène sa présentation, mais elle protège contre la force totalisante d'autrui en permettant la pluralité des points de vue. Même si ce chapitre ne traite pas du rapport entre confiance et vie privée, un lien est postulé : « les accords en matière de *privacy* constituent une dimension de l'un des prérequis pour une société numérique : la confiance » (p.169). Préserver la vie privée constituerait donc un défi majeur.

L'avant dernier chapitre, rédigé par Jessica Feldman, chercheure spécialiste des médias, traite des techniques de surveillance de masse et d'informatique prédictive à travers lesquelles la liberté est supplantée par le contrôle, thème lié à celui du précédent chapitre. L'idée est défendue d'une proximité entre le totalitarisme et la surveillance numérique, dont le désir commun est de « contraindre le sujet politique à travers des épistémologies et des technologies de prédiction, et en prétendant à l'autorité » (p.197). Pour ce faire, l'auteure prend également appui sur les travaux d'Hannah Arendt, suivant lesquels le totalitarisme est lié à une croyance pathologique en une inévitable destinée de l'humanité fondée sur un raisonnement logique pour comprendre et prédire le comportement des autres sans attacher d'importance aux faits et à l'expérience. Cette pensée linéaire, logique plutôt qu'empirique, émanerait de la solitude et reflèterait un désir de contrôle à distance plutôt qu'un désir d'intimité ou de création d'un lien.

[33] Arendt H. (1998), *The Human Condition*, 2nd edition, Chicago, University of Chicago Press.

L'usage de ce cadre arendtien vise une meilleure compréhension de la relation entre déclin de la démocratie participative et développement d'une surveillance d'Etat généralisée et prédictive. D'une part, il suggère le rôle de la solitude dans la tendance au développement de la surveillance, à la différence des études sur la surveillance qui s'inscrivent dans le paradigme visuel dérivé des travaux de Foucault sur le panoptique de Bentham. D'autre part, il contribue à expliquer l'accent mis sur les formes d'écoute qui ne s'intéressent pas tant à ce qui est dit « c'est-à-dire aux faits et à l'expérience, qu'à la prédiction de schémas généraux basés sur les données à partir des schémas de communication d'une personne […] Cet accent mis sur les probabilités et les statistiques dans les écoutes coïncide avec l'idée d'Arendt sur la logique comme principe de fonctionnement fondamental du totalitarisme » (p.200). Par ailleurs, la surveillance ne doit plus être envisagée sous l'angle du panoptique mais sous l'angle de l'écoute, « car espionner une conversation a plus à voir avec la communication et les relations interpersonnelles qu'avec le contrôle exercé à distance sur les corps » (p.201). La dernière partie du chapitre s'intéresse donc aux effets des écoutes, en s'appuyant sur des entretiens avec des acteurs de mouvements sociaux. Trois effets sont identifiés. Le premier est un effet dissuasif, caractérisé par une autocensure des populations, qui renoncent à s'exprimer, *via* les médias technologiques, sur des sujets politiques, surtout s'ils sont hétérodoxes. Cette autocensure nuit aux mouvements sociaux, en réduisant la capacité des membres d'un groupe à nouer des liens, à s'exprimer ouvertement, à se faire confiance, et en augmentant, corrélativement, la paranoïa. En particulier, la peur de l'espion suscite la défiance plutôt que la confiance envers le nouveau-venu dans un groupe. Cet effet peut être sciemment recherché par des gouvernements, pour gérer le risque terroriste, le blanchiment d'argent et, plus généralement, tous ceux jugés « ennemis d'Etat ». Il serait annihilé par un sentiment de solidarité et de rejet lorsque les groupes ressentent la surveillance comme non justifiée. Le second effet est la peur que nos propos d'aujourd'hui soient utilisés contre nous après coup, du fait de leur enregistrement et de leur stockage. Il conduit à une expression plus limitée, plus contrôlée, moins créative, moins enjouée ; il empêche de penser à voix haute et génère anxiété. Le troisième est la peur de compromettre ses contacts et, le cas échéant, le sentiment de culpabilité associé. Au final, « la surveillance généralisée et omniprésente de la parole mène donc à une forme de censure interpersonnelle et collective qui exclut certaines des conditions fondamentales de la confiance et de l'amitié » (p.212).

Le dernier chapitre, rédigé par Maurizio Ferraris, professeur de philosophie, traite de la post-vérité. Ce phénomène est le fait du post-truiste, celui qui pratique la post-vérité, qui définit « la vérité (et non le mensonge) comme la non-correspondance entre la proposition et la chose » (p.215), qui fait du vrai une possession personnelle et qui conçoit donc la théorie de la vérité comme « la correspondance entre la proposition et les convictions personnelles de celui qui l'énonce » (p.216). Ferraris considère que la

post-vérité est un concept singulier, distinct de ceux d'intox et de mensonge, qu'elle s'exerce sur des sujets d'intérêt public (elle ne concerne pas les controverses privées) et que son champ de manifestation est le web qui permet l'expression de « millions de vérités individuelles » (p.217). Les prémices et la coloration de la post-vérité sont associées au postmodernisme philosophique, car c'est la critique de la vérité et de la réalité objective qui « a ouvert la voie au post-truisme comme subjectivation de la vérité » (p.216). C'est donc cet arrière-plan qui est présenté dans une première partie, de manière historique et critique, les fondements de la pensée postmoderne étant exposés puis résumés sous la forme de sophismes, et de manière extensive, en ce sens qu'il inclut un développement sur le populisme en tant que traduction politique du postmodernisme. La post-vérité est ici doublement associé au populisme : elle en reprend les erreurs et elle en constitue un instrument massif. L'ancrage postmoderne et populiste de la post-vérité étant posé, Ferrandis précise, dans une seconde partie, ce qui caractérise sa pratique, le post-truisme. En substance, il est violation des règles de la vérité, par « l'infraction systématique des quatre maximes conversationnelles énoncées il y a quarante ans par le philosophe Paul Grice » (p.229). Ces quatre maximes sont présentées et discutées dans le texte (p.229 à 235) : « sois sincère, fournis des informations véridiques pour autant que tu le saches » (maxime de la qualité), « ne sois ni réticent ni redondant » (maxime de la quantité), « sois pertinent » (maxime de la relation) et « évite l'ambiguïté, évite le bavardage » (maxime de la modalité). Premièrement, le post-truisme consiste à élaborer une vérité alternative sur des choses qui concernent la sphère de l'opinion publique, non pas pour duper, comme dans le cas du mensonge, mais pour émanciper le destinataire et/ou satisfaire ses intérêts en lui proposant une critique des vérités normatives d'une certaine autorité (la science, les experts). Deuxièmement, la post-vérité est en général redondante et comprend la création de sotises (*bullshit*), qui ne sont pas des productions conscientes de mensonges mais simplement des croyances erronées, qui se diffusent largement par viralité. Troisièmement, le post-truisme comprend des intox, des messages non pertinents, qui serviraient le plus souvent de prétexte pour exprimer une indignation et obtenir une reconnaissance plutôt que, comme dans l'intox classique, de renforcer la cohésion d'une société soudée (théorie du complot, péril juif, etc). Enfin, le bavardage, le *fashionable nonsense*, est prégnant dans la post-vérité, associé à l'idée postmoderniste que la parole exerce du pouvoir sur les choses, que notre représentation du monde dépend de notre vocabulaire. Pour autant, Ferraris ne désapprouve pas le post-truisme. Certes, il peut constituer un danger, par exemple si une communauté dans les hôpitaux ou les tribunaux se convainc du caractère insignifiant de la vérité ou si la science est discréditée au profit de la foi. Mais il reflète aussi le fait que l'humanité est plus que jamais « occupée à penser avec sa tête » (p.236). En conséquence, il ne s'agit pas tant de désapprouver la post-vérité que de développer une capacité de se mettre dans la tête des autres. La

vérification des faits et l'analyse de réputation font partie des moyens envisageables. Ils sont jugés insuffisants « parce que l'humanité n'est pas intéressée à savoir le vrai mais à avoir raison et à trouver confirmation de ses propres convictions… Lorsque l'on parle de post-vérite, on pense toujours à Trump mais on oublie deux choses essentielles. La première est que les électeurs de Trump pensent exactement comme lui, sinon ils ne l'auraient pas élu ; la deuxième est que tous ensemble sont convaincus non pas de dire des choses post-vraies, mais bien plus d'affirmer la pure et simple vérité et que les menteurs ce sont les Autres » (p.237). La dernière partie du chapitre traite donc de la manière dont on peut apprendre à penser avec la tête d'autrui et ainsi faire la vérité. Ferrandis y développe une théorie de la vérité fondée sur une distinction entre ontologie (ce qui existe indépendamment de ce que nous savons, sphère de la réalité), épistémologie (ce que nous savons sur ce qui existe, sphère de la vérité) et technologie (ce qui lie l'ontologie à l'épistémologie, l'être à la vérité, une action de mesure, de vérification d'un objet naturel ou social, une compétence qui assume la tâche de faire la vérité). Ce texte ne traite finalement en rien du rapport entre post-vérité et confiance. Toutefois, il ne fait guère de doute que la post-vérité peut susciter concomitamment de la défiance (des partisans d'une vérité plus orthodoxe) et de la confiance (de *followers* envers les post-truistes). En ce sens, il invite à explorer ces liens.
L'ouvrage s'achève par une « ouverture » rédigée par Jacopo Domenicucci et Milad Doueihi, au titre malicieux : *In Code, we trust ?*. Le point d'interrogation est judicieux. D'une part, la confiance dans le code informatique ne va pas de soi, face à un langage qui ne s'acquiert pas par socialisation. Il est en mutation permanente et peut être manipulé dans un contexte de post-vérité, de surveillance généralisée et de menace de la vie privée. D'autre part, cela requiert d'appréhender la confiance sous des angles nouveaux et non exclusivement en termes de relation interpersonnelle.

4. Point de vue critique sur l'ouvrage

La confiance à l'ère numérique est un ouvrage d'une grande richesse. Il traite de confiance médiée numériquement, mais aussi de confiance non médiée, car il faut bien un point de référence initial pour appréhender l'évolution des formes de confiance dans les contextes numériques. Il précise le rapport entre confiance et sécurité, entre confiance et rationalité, entre confiance et fiabilité, entre confiance et assurance, la différence entre confiance pratique, systémique et épistémique, entre confiance décidée ou confiance-pari et confiance assurée ou confidence et le concept de confiance distribuée. Il montre que les formes de confiance à l'œuvre varient selon les univers numériques auxquels on s'intéresse (les interactions en ligne, l'information en ligne, la téléprésence, les sites d'achat en ligne, un système de monnaie électronique, le *cloud*) tout comme les médias privilégiés peuvent varier selon le niveau de confiance préexistant entre des agents. Enfin, il évoque les obstacles à la confiance en ligne

(comme l'anonymat, les cyberattaques, l'espionnage, la surveillance généralisée, la difficile protection de la vie privée, la post-vérité) et, parfois, des moyens de les contourner (comme la signalisation, les indices, la cryptographie, « l'interlucency », des normes, la vérification des faits, la réputation en ligne).

La variété des angles d'attaque de la confiance numérique retenus est à mettre à l'actif des coordonnateurs et du profil des contributeurs qu'ils ont réunis, dont les champs de prédilection (éthique, philosophie, sociologie, information et communication, numérique) sont aussi divers que les lieux d'exercice (Allemagne, Etats-Unis, France, Italie, Norvège, Pays-Bas, Royaume-Uni). Cette variété aurait pu encore être enrichie par le regard de chercheurs en management dont les contributions sur la confiance et sur l'articulation entre confiance et digital ne manquent pas, notamment dans le champ du marketing, de la gestion des systèmes d'information, du pilotage des équipes virtuelles ou de l'économie collaborative. C'eût été aussi, peut-être, un moyen d'agrémenter l'ouvrage de résultats empiriques, qui font ici largement défaut. Par ailleurs, nous l'avons mentionné au fil de cet article, l'articulation explicite entre confiance et numérique n'est pas au rendez-vous dans tous les chapitres, ce qui est frustrant, mais aussi stimulant pour de futures contributions.

On notera finalement deux limites relatives à la construction de l'ouvrage. Si l'ordonnancement des chapitres est habile, il n'est réalisé que par juxtaposition. On peut regretter que les contributeurs ne dialoguent pas entre eux par des références croisées ou des renvois entre chapitres, alors même qu'il existe des zones de chevauchement dans les sujets qu'ils abordent. Il faut noter, c'est sans doute lié, que quatre des huit chapitres de l'ouvrage sont des reprises d'articles publiés antérieurement. Enfin, l'ouvrage ne comporte pas d'index des thèmes, concepts, mots-clés associés à chaque chapitre, ce qui aurait été appréciable étant donné l'étendue des thèmes traités et l'intérêt qu'il peut y avoir à comparer le traitement des thèmes et des concepts qui se retrouvent dans plusieurs chapitres.

Conclusion

L'ouvrage proposé n'est pas un ouvrage de vulgarisation. Il s'adresse à des chercheurs de toutes disciplines intéressés par la confiance mais aussi par les défis du numérique. Pour ceux-là, sa lecture sera assurément intellectuellement stimulante.