Interactions en ligne et concept de confiance

Laurent Jaffro

Brouillon du texte paru dans M. Doueihi et J. Domenicucci, *La confiance à l'ère numérique*, Paris, Berger Levrault/Editions rue d'Ulm, 2018, p. 33-62.

Interactions en ligne et concept de confiance

Laurent Jaffro, université Paris 1 Panthéon-Sorbonne et Institut universitaire de France

Ce chapitre porte sur la nécessité et les difficultés de l'application du concept de confiance dans le contexte des interactions en ligne. L'opinion à ce propos hésite entre deux appréciations opposées. Si le cyberespace est un univers de désinformation et de tromperie, dépourvu de protections suffisantes, comme on l'entend souvent dire, comment se fait-il alors qu'il ne soit pas déserté ? Faut-il que les gens soient crédules ou irrationnels à ce point ? Ou doit-on comprendre que des formes de confiance se développent cahin-caha dans les divers usages de l'internet, et que ces formes de confiance ont une certaine rationalité ? Il est certainement de l'intérêt des acteurs du commerce en ligne d'encourager la confiance et de décourager le scepticisme : mais, selon une vue commune, là où il y a du commerce, il y a aussi, de fait, de la confiance. De manière analogue, on peut douter que les « communautés » que constituent les groupes de discussion (ou « forums ») aient beaucoup d'affinité avec une communauté scientifique habituée à la critique et animée du souci de la vérité ; cependant, ces groupes explorent de nouvelles formes de l'espace public qui se dotent de leurs modes de régulation.

Les sciences sociales, les sciences de l'information et de la communication, et la philosophie, qui s'efforcent d'apprécier ce qu'il advient de la confiance dans le cyberespace ne doivent pas céder à la pression pragmatique des compagnies et firmes dont l'activité ne peut prospérer que si la confiance l'emporte sur la crainte ; ce n'est pas simplement parce qu'il y a un immense besoin de confiance dans ce domaine que le concept de confiance peut y être employé de manière pertinente. Même si la confiance dans la confiance, si l'on peut dire, est éminemment désirable aux yeux des parties intéressées au développement des réseaux sociaux, du commerce et

des services en ligne, le premier but des chercheurs, en tout cas dans des disciplines comme la philosophie, la science politique et plus généralement les sciences sociales, est de comprendre la confiance et non pas de contribuer à la propager.

Le chapitre procède de la manière suivante : Après une partie introductive qui mêle les considérations méthodologiques à la présentation du concept de confiance pratique, une deuxième partie défend la thèse selon laquelle une forme de la confiance, la confiance systémique, distincte de la confiance « par excellence » qu'est la confiance décidée, constitue un enjeu et un problème majeurs pour les interactions en ligne ; cette confiance systémique est étroitement liée à un concept de confiance épistémique qui ne se confond pas avec la confiance pratique. La troisième partie montre comment cette thèse peut jeter une nouvelle lumière sur les obstacles à la confiance en ligne et sur quelques problèmes à la fois théoriques et pratiques.

I. Concept et méthode

La confiance au sens assez déterminé qui intéresse particulièrement les sciences sociales et la philosophie se traduit par une conduite par laquelle un agent se place dans la dépendance d'une conduite d'un autre agent afin de réaliser une certaine fin, qu'il lui serait difficile ou plus compliqué d'atteindre sans cela, en escomptant plus ou moins sciemment que le signalement de cette dépendance incite l'autre agent à adopter cette conduite. On peut désigner cette forme de confiance sous les appellations de confiance décidée ou confiance-pari. Elle se distingue de la simple attente à l'égard d'une conduite d'autrui par la prise de risque que constitue l'acceptation de la dépendance à l'égard de cette conduite. La déception possible est à la mesure de l'investissement. Certes, autrui peut aussi « décevoir » une simple attente de ma part, mais, dans ce cas, il n'a pas à mes yeux le tort de faire défection. Car dans le cas de la simple attente le tort est mien, c'est celui d'une attente mal placée, tandis que dans l'autre cas, le tort est sien, et c'est celui d'une trahison de la confiance. Si l'on distingue ainsi attente prédictive et attente normative, la confiance est du côté de la seconde.[1]

1 Voir sur ce point Richard Holton: « Quand vous faites confiance à quelqu'un pour faire quelque chose, vous comptez sur cette personne pour qu'elle le fasse et vous considérez cette assurance [*reliance*] d'une certaine manière : vous avez tendance à avoir le sentiment de la trahison si elle est déçue et de la gratitude si elle est confirmée. » (Holton 1994, p. 67) Philip Pettit distingue aussi le genre du *reliance* de l'espèce qu'est le *trust*, en proposant des conditions très déterminées. Relèvent du premier une variété d'attitudes, dont les attentes à l'égard des conduites des personnes ou des institutions selon leurs rôles (par exemple à l'égard du médecin), mais aussi à l'égard de l'environnement (à l'égard des infrastructures routières, par exemple). Cela revient à considérer les

Il y a là une différence semblable à celle que l'on peut faire entre l'annonce ou description par un agent de sa conduite future et la promesse d'adopter cette conduite. Si *A* promet à *B* d'aller demain avec lui au cinéma, et qu'il se trouve que le lendemain *A* ne le fait pas, sauf empêchement majeur, *A* a le tort de ne pas tenir parole (et non pas le tort de l'avoir promis). Mais si *A* a simplement annoncé qu'elle irait demain avec *B* au cinéma, en décrivant simplement l'emploi du temps supposé de sa journée, et si le lendemain *A* ne le fait pas, le tort est encore sien, mais c'est le tort de la veille, à savoir celui d'avoir fait cette prédiction. Dans le premier cas, *B* protestera en rappelant à *A* la parole donnée ; dans le second cas, *B* reprochera à *A* d'avoir parlé inconsidérément.

L'analogie entre confiance décidée et promesse pourrait être poussée plus loin. Décider de faire confiance à une autre personne et attendre d'elle (dans le sens normatif de l'attente, typique de la confiance, et non pas simplement prédictif) qu'elle rende tel service, c'est en quelque sorte lui proposer de tenir une promesse qu'elle n'a pas faite elle-même. Je lui fais cette proposition tout bonnement en agissant d'une certaine manière qui anticipe le respect par elle de cette sorte de promesse. Il y a dans la confiance quelque chose de l'ordre de la promesse par procuration. Dans les termes d'une coopération, on peut dire que, quand l'agent *A* fait confiance à l'agent *B*, *A* adopte une conduite dont il escompte le complément de la part de *B*. Ce qui est particulier à cette confiance décidée est que *A* espère que cette adoption même motive chez *B* la réponse attendue.

Notez qu'alors que l'objet de ce chapitre est la confiance en ligne, premièrement, j'ai commencé par esquisser le portrait de la confiance, indépendamment du contexte de l'internet et des réseaux ; deuxièmement, je suis allé directement à la confiance « par excellence » qu'est la confiance décidée. J'ai adopté ainsi une manière de procéder qui est difficilement évitable, mais qui est contestable. Reprenons ces deux points.

Premièrement, il est assez naturel qu'on ait besoin de savoir d'abord ce qu'est la confiance avant de se demander ce qu'il advient d'elle dans les interactions en ligne. Le fait est que la littérature sur le sujet commence presque toujours par un propos général sur le concept de confiance, puis passe aux obstacles à son

---

personnes ou les choses ou les institutions comme « *reliable* » (fiables au sens de prévisibles), mais pas nécessairement comme « *trustworthy* » (fiables au sens de dignes de confiance) (Pettit 2004, p. 109-110).

application aux interactions en ligne, et valide ou non cette application avec des réserves importantes ou non[2]. Mais, ce faisant, on tend à négliger une approche que l'on pourrait dire phénoménologique (en un sens assez modeste du terme) qui consiste à partir d'une description des formes de la confiance et de la méfiance en ligne. Cependant, dans la mesure où les interactions en ligne tendent à simuler les interactions « naturelles » (ici le terme inclut le monde social et celui des technologies classiques), il n'y a rien d'étonnant à ce qu'on prenne comme étalon le concept de confiance qui est taillé à la mesure des interactions sociales ordinaires. Même s'il est difficile de procéder autrement, les études sur la confiance en ligne ne devraient-elles pas faire l'effort de partir davantage des données de nos expériences d'utilisateurs de l'internet et des réseaux ? Malheureusement, les études qui procèdent de cette manière-ci ont tendance à réduire la confiance à la sécurité informatique, car la question a tôt fait de prendre, dans cet environnement, le visage de celle de la sécurisation, de la certification, de la traçabilité, etc. Il faut alors retourner vers les concepts ordinaires de la confiance pour rétablir un questionnement plus ouvert. On doit se satisfaire d'une sorte de va-et-vient entre ce que l'on pourrait appeler nos premières idées sur la confiance et la considération des difficultés propres aux interactions en ligne.

Deuxièmement, la confiance « par excellence » n'est certainement pas la seule forme de confiance, et on peut se demander si elle est le meilleur candidat pour rendre compte de la question dans ce domaine. Ce point appelle un développement.

La structure du concept de confiance

Il y a en effet tout un ensemble de phénomènes sociaux et mentaux variés que l'on regroupe sous la rubrique générale de la confiance. Les recherches sur la confiance dans les sciences sociales et en philosophie tendent à privilégier certaines de ses formes comme si elles incarnaient mieux l'essence de la confiance. Une présentation plus nuancée est donnée par Thomas W. Simpson, qui insiste sur la pluralité des formes de la confiance.

2 C'est ainsi que procède Philip Pettit. Il commence par caractériser la confiance sous sa forme la plus discutée (la confiance « rationnelle » que l'on donne explicitement aux autres afin d'obtenir d'eux certaines conduites), en décrivant son « écologie » qui suppose le contact, le « face-à-face » ou d'autres supports d'une *evidence* (je reviens plus loin sur cette notion) sans laquelle la confiance serait irrationnelle, puis il exprime des doutes sérieux à l'égard de la capacité de l'internet à fournir cette écologie (Pettit 2004, p. 117). Cette manière de procéder ne prend pas assez au sérieux le sentiment qu'ont les utilisateurs d'entrer en contact avec les autres par divers moyens sur l'internet, par exemple grâce à des identités fictives qui n'ont pas pour seule fonction de cacher leur identité réelle, mais aussi de leur permettre d'échanger. Cependant le scepticisme de Pettit est explicitement relatif à l'état de l'internet et des modes de la « téléprésence » (p. 120) au moment où il écrit. Il semble ainsi prêt à admettre que les difficultés de l'idée de « confiance en ligne » soient plus empiriques que conceptuelles.

Simpson fait le portrait d'une confiance primitive et en quelque sorte matricielle, qu'il appelle *Ur-Trust*, qui consiste en des conduites plus ou moins spontanées qui préjugent plus ou moins implicitement de la coopération d'autrui, et qui comporte des *degrés*. Il distingue aussi des *espèces* de la confiance : cognitive (quand la confiance est proportionnée à l'*evidence* de la fiabilité), conative (quand elle consiste essentiellement en des intentions et résolutions), affective, et prédictive[3]. Certaines formes de la confiance consistent en des habitudes qui ne sont pas explicitées et qui n'ont pas besoin d'être rationalisées. Cette vision pluraliste de la confiance lui permet d'admettre, par exemple, qu'il y a bien une question de confiance (sous la forme d'une confiance *prédictive*) dans nos relations avec les machines, services et réseaux informatiques, et de ne pas limiter la discussion aux formes de la confiance à l'œuvre dans les interactions entre personnes, comme ont décidé de le faire assez arbitrairement certains chercheurs. Simpson observe aussi que cette confiance seulement prédictive, lorsqu'elle repose surtout sur la sécurisation, n'est pas vraiment de la confiance[4]. Au-delà de cet exemple, Simpson estime que la pluralité des formes de la confiance invite à un examen au cas par cas de la question de la confiance en ligne (selon qu'elle se pose dans le contexte d'un achat, d'une conversation, d'une recherche, d'un groupe de discussion, d'un réseau social au sens étroit du terme, de l'internet des objets, etc.). Si l'on ne passe pas en revue ces formes et leur présence dans une variété d'usages de l'internet et des réseaux, on risque fort de produire une analyse simpliste. Simpson a raison d'insister sur la structure plurielle et flexible de la confiance.

Dans le même esprit, je propose d'intégrer pleinement dans le concept de confiance cette diversité de formes et de degrés. Il me paraît commode d'envisager la diversité des formes de la confiance comme pouvant s'ordonner dans une « échelle » au sens que Robin G. Collingwood avait donné à ce terme, dans laquelle les différences de degré, en certains points, constituent des différences de nature[5]. En haut de l'échelle, on mettra les formes de confiance les plus volontaires, car la confiance « par excellence » se donne, s'accorde, est ainsi de l'ordre d'un

3 Simpson 2014.

4 « La sécurité rend possible la confiance simplement prédictive, mieux décrite sous le nom d'assurance [*reliance*]. » (Simpson 2014, p. 115)

5 Lorsqu'une notion générique comprend une variété d'espèces qui diffèrent entre elles à la fois en degré et en nature, et qu'elles se rangent ainsi, on a une *scale of form*s. Il en va ainsi des états solide, liquide et gazeux de l'eau ; la variable étant la chaleur, l'essence de l'eau étant incarnée aussi bien par chacun d'entre eux. Mais il est caractéristique des concepts philosophiques, selon Collingwood, que la variable soit identique à l'essence, de sorte qu'une forme incarne parfaitement l'essence et que les autres formes présentes sur l'échelle l'incarnent « moins » (Collingwood 1933, p. 57-61).

acte au sens fort du terme, de quelque chose que l'on décide en estimant que le résultat susceptible d'être ainsi atteint vaut le risque pris. Cette décision est sensible à des raisons. La volonté, ici, n'est pas exclusive de l'intelligence. Sur la partie inférieure de l'échelle on pourra placer les formes de la confiance les moins volontaires, les plus implicites, et les moins raisonnées, comme l'absence de suspicion[6].

Les formes distinctes, qui sont aussi des degrés, de la confiance s'arrangent peut-être ainsi, depuis ce qui n'est que faiblement confiance jusqu'à ce qui est supposé l'être vraiment, la flèche indiquant la direction du plus implicite et moins volontaire vers le plus volontaire et le plus explicite :

Petite échelle

absence de suspicions (familiarité, habitude) → assurance → confiance décidée

Au sommet de l'échelle, la forme « par excellence » de la confiance combine volonté et intelligence, décision et croyance. Elle est éminemment volontaire, puisque, dans un contexte de jeux répétés, elle s'amorce par la décision d'agir de manière confiante, c'est-à-dire sur la prémisse qu'autrui est digne de cette confiance. Si autrui ne se conduit pas en retour d'une manière qui conforte cette décision, j'ai des motifs de ne pas courir une nouvelle fois ce risque. Si autrui répond comme je l'espère, ma décision étant confortée par un début de preuve de sa fiabilité, je pense gagner à la répéter lors d'une nouvelle occasion d'interaction, de sorte que la fiabilité d'autrui devient progressivement l'objet d'une croyance et non pas simplement d'un pari. Le nœud de la confiance est cependant plus serré que cela, et ne dépend pas entièrement de la seule bonne fortune dans les réponses d'autrui. En effet, en signalant à autrui et mieux encore à des tiers ma décision de faire confiance, j'incite autrui à considérer les inconvénients d'une défection et, comparativement, l'avantage qu'il y a à installer la coopération. Je dispose donc de moyens d'influencer la conduite d'autrui, car il sait que j'aurai plus tard les preuves de sa conduite qui me manquent au moment où je prends l'initiative. Si, comme le rappelle Weckert[7], on peut voir une analogie entre la confiance décidée et ce que William James disait de la croyance religieuse, à savoir qu'elle est un « saut dans l'obscurité », il convient de préciser que l'obscurité s'éclaircit à mesure que l'interaction se répète.

6 Il ne s'agit pas de n'importe quelle absence de suspicion, en tout cas pas d'un simple zéro de suspicion, car il est requis que l'on soit dans un contexte d'interaction entre des agents qui est susceptible de fournir des raisons d'avoir des doutes ou de prendre des précautions.

7 Weckert 2005, p. 113.

Cette forme « haute » de la confiance est-elle à l'œuvre dans les interactions en ligne ? Weckert considère que la confiance décidée est présente dans les équipes virtuelles que l'on peut constituer dans les échanges en ligne, notamment lorsqu'il s'agit de projets qui relèvent du « *web* collaboratif »[8]. Plus récemment, une étude de cas conduite par Vincent Véchambre voit dans une communauté virtuelle d'hospitalité et d'hébergement un phénomène de confiance interpersonnelle, une « rencontre des volontés et pas un simple service d'accueil »[9].

Mais ce que j'ai appelé la petite échelle peut être replacé dans une échelle plus vaste, non plus de la confiance, mais cette fois plus largement de l'interaction dans l'incertitude, sur laquelle les formes extrêmes *distinctes* sont aussi des formes *opposées*[10] :

Grande échelle

défiance complète → prudence (méfiance, circonspection) → coopération → confiance

Mon propos n'épouse que partiellement la conception de Collingwood : je reprends ce qu'il dit de la structure des concepts philosophiques et je l'applique à la confiance pratique ; comme il le fait pour d'autres concepts, je situe en haut de la petite échelle la forme à la fois la plus volontaire et la plus intelligente. Cependant, je ne peux pas suivre Collingwood sur un point majeur : je crois qu'estimer que la forme la plus haute incarne le mieux la confiance est une conception dominante, mais simplificatrice, sinon erronée. Elle encourage un certain mépris théorique pour les formes inférieures. C'est un biais que l'on rencontre dans la plupart des études sur la confiance[11], du moins quand celles-ci ne la confondent pas purement et simplement avec la sécurisation.

Nécessité pratique et théorique du concept de confiance

Une question que l'on ne peut éviter est celle de l'intérêt de l'interrogation sur la confiance. La notion est certainement importante sur le plan pragmatique – d'où son succès dans les sciences de gestion –, mais est-elle

8 Citant à l'appui le travail de Patricia Wallace (Wallace 1999), Weckert estime que « si l'environnement en ligne est de nature à contribuer à ce que les gens agissent comme s'ils avaient confiance, la confiance authentique peut émerger » (*ibid.*).

9 Véchambre 2013, p. 41.

10 Sur la différence entre « distinction » et « opposition », voir encore l'*Essay on Philosophical Method* (Collingwood 1933, p. 74-76). Selon Collingwood, la coalescence entre distinction et opposition est caractéristique des concepts philosophiques. Appliqué à la confiance, cela signifie que l'opposition entre confiance et défiance est aussi une relation entre deux formes d'interaction dans l'incertitude.

11 Louis Quéré parle du « cas paradigmatique » de la confiance interpersonnelle et met en cause à juste titre son hégémonie : « Il ne semble pas que dans tous les cas de confiance il y ait une composante de délégation de pouvoir à autrui pour l'obtention d'un résultat concernant le bien-être de celui qui fait confiance, pas plus que l'engendrement d'une obligation d'honorer la confiance faite. C'est d'ailleurs pourquoi dire que l'on se fie à des objets, à des systèmes ou à des institutions ne pose problème que si l'on se donne une conception très stricte de la confiance. » (Quéré 2001, p. 141-142)

réellement utile sur le plan théorique ? La position déflationniste répond « non ». La confiance peut être conçue comme un ensemble de conduites, d'attitudes, de croyances dont on peut supposer l'existence dès lors que des personnes s'engagent dans un échange dont l'issue n'est pas immédiate, mais différée, par exemple un échange commercial. C'est un point non controversé que l'acceptation du risque associé à une telle interaction fortement diachronique témoigne qu'il y a de la confiance en ce sens[12]. Mais selon l'interprétation déflationniste, la confiance n'est pas la cause ou la raison de cette acceptation. Cette acceptation est une condition suffisante pour qu'on puisse parler de confiance, et non pas un effet ou une conséquence d'une confiance. Faire appel à la confiance pour rendre compte de cet échange n'explique rien. Le risque est accepté parce qu'il est rationnel pour les personnes concernées de l'accepter, ou parce qu'elles ont des croyances ou des habitudes qui les conduisent à l'accepter. Donner un rôle explicatif à la confiance est de ce point de vue superflu. Il y a une analogie avec ce qu'on appelle « bonne ambiance » dans une soirée entre amis. Il suffit que la soirée se passe très agréablement pour que l'on soit justifié à parler de bonne ambiance. Mais il serait stupide d'expliquer le fait que la soirée se passe très agréablement comme l'effet de la bonne ambiance. Comme nous sommes assez superstitieux, nous avons tendance à donner un rôle causal à la confiance alors que la notion résume un ensemble de conditions qui sont remplies dès lors que des personnes interagissent d'une certaine manière, que l'on caractérisera comme amicales, parfois, ou même comme seulement coopérative. En outre, les auteurs qui donnent un rôle explicatif à la confiance ont tendance à opposer confiance et choix rationnel. Pour eux, le choix rationnel est celui de la méfiance et conduit à ne pas s'exposer au risque en prenant une initiative qui augmente la vulnérabilité. Par suite, si cette initiative est prise, c'est qu'une autre source de l'action est intervenue : la « confiance ». Mais, pour le déflationniste, cette nouvelle explication est *ad hoc*. On tient pour acquis que le la fluidité des échanges sociaux y compris lorsqu'ils supposent l'acceptation de risques importants ne peut pas s'expliquer par le choix rationnel, les croyances et les habitudes. Mais cela ne va pas du tout de soi.

Ce déflationnisme est séduisant, mais il manque une dimension importante des phénomènes de confiance. « Confiance » (ou ses équivalents en français ou en d'autres langues) n'est pas seulement un terme théorique

12 En introduction à son analyse du « capital social », Robert Putnam rappelle le mot de Kenneth Arrow : « Toute transaction commerciale a virtuellement en elle un élément de confiance, en tout cas toute transaction conduite dans la durée. » (Arrow 1972, p. 357 ; cité par Putnam 2000, 288) Dans cet article, Arrow mettait le *trust* aux côtés de la véracité (*truthfulness*) dans une liste des vertus qui constituent un contexte favorable pour l'efficacité économique (p. 354).

qu'utilisent les sciences sociales ou les techniques comme la politique ou le marketing. C'est aussi un terme du langage ordinaire qui fait l'objet d'un usage intensif dans les pratiques. Car la signalisation de la confiance, au premier chef l'explicitation « je te fais confiance », bien qu'elle ne soit pas une condition nécessaire des phénomènes de confiance, joue très souvent un rôle causal. Quand *A* agit en signalant la présomption que *B* va répondre favorablement, cette signalisation accompagnée d'autres facteurs (comme le souci de la réputation) peut inciter *B* à répondre favorablement. En ce sens, la confiance affichée par *A* fait partie de ce qui est susceptible d'expliquer la conduite de *B*, et en retour d'expliquer également la conduite de *A*. Cela ne signifie pas qu'une théorie de la coopération en termes de choix rationnel doive céder la place à une théorie de la confiance qui aurait une physionomie entièrement différente (elles figurent toutes deux sur la « grande échelle », que j'ai présentée plus haut). Cela signifie plutôt que la théorie de la coopération doit être complexifiée et notamment doit intégrer le rôle de la signalisation de la confiance.

Je peux décrire telle interaction comme ayant été accomplie dans la confiance. Par exemple, je peux décrire la relation entre un patient et son médecin comme une relation de confiance (dans les deux sens, d'ailleurs). Par là, je n'explique pas pourquoi et ni comment la relation a réussi, mais j'indique qu'elle a réussi. Mais dans l'exemple du patient et de son médecin, la signalisation de la confiance ne joue guère de rôle. S'il fallait qu'un patient explicite sa confiance pour augmenter ses chances d'obtenir du médecin la conduite attendue, ce serait assez bizarre, car la conduite attendue est déterminée et motivée suffisamment par les obligations professionnelles du médecin. Inversement, dans les situations où cette signalisation joue un rôle important, le vocabulaire de la confiance ne sert pas seulement à cette redescription d'un succès. Son adoption est importante pour *obtenir* le succès, de sorte qu'il n'est pas absurde de donner un certain rôle explicatif à la confiance.

Ce sont des différences très fines. Prenons l'exemple du facteur qui distribue le courrier. En l'absence des destinataires, mais en présence des voisins, il a tendance à confier les colis aux voisins. Peut-être a-t-il l'habitude de procéder ainsi. Le fait-il en insistant auprès des voisins sur le fait qu'il leur fait confiance ? Selon que l'on répond oui ou non à cette dernière question, on identifie une forme distincte de confiance : avec ou sans signalisation. On voit ici que les différences de degré peuvent donner lieu à des différences d'espèces, selon l'esprit de l'échelle collingwoodienne des formes. On a donc besoin du concept de confiance, y compris

sous sa forme « par excellence » ; mais cela ne veut pas dire qu'elle soit la seule pertinente.

Un problème d'application

Un dernier point de méthode : Ce chapitre traite d'un problème d'application qui suppose que nous avons un concept de confiance. Le terme « notion » serait peut-être plus prudent, puisqu'il s'agit en réalité d'une famille conceptuelle aux formes variées et aux frontières floues. Mais gardons le mot « concept » avec cette réserve à l'esprit, qui est bien satisfaite par l'idée d'une échelle de formes. On peut réagir de plusieurs manières à un problème d'application de concept. Pour qu'un tel problème surgisse, il suffit que le concept paraisse ne pas s'appliquer ou s'appliquer avec beaucoup de difficulté dans tel domaine.

Une première réaction consiste à conclure que le concept ne s'applique pas ou très mal dans ce domaine. Il n'est pas pertinent et la tentative de l'appliquer n'est pas intéressante.

Une deuxième réaction est de conclure qu'il faut réviser en profondeur le concept parce qu'il *devrait* s'appliquer, même si c'est au prix de quelques adaptations. Cette deuxième réaction peut être motivée par l'existence d'une articulation entre un concept de sens commun, plutôt lié à des pratiques, et un concept philosophique, plutôt lié à des théories, elles-mêmes révisables.

Une troisième réaction consiste à accepter l'inadéquation relative du concept, à essayer de la faire servir à une meilleure compréhension, et à être attentif aux évolutions du concept qui sont à l'œuvre dans les pratiques concernées, surtout lorsqu'elles sont elles-mêmes des pratiques sociales relativement nouvelles et très évolutives, car liées à des environnements technologiques eux-mêmes en mutation constante.

Une quatrième réaction est de considérer qu'il existe un autre concept, voisin, avec lequel le premier est parfois confondu, et dont l'application est pertinente.

Le présent chapitre combine les deux dernières stratégies. Le concept voisin est celui de la confiance épistémique, cette 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. Ma thèse est que l'application du concept de confiance aux interactions en ligne est complexifiée par des problèmes de confiance

épistémique.

Cela permet de mettre en valeur des questions que la littérature sur la confiance numérique tend à négliger : celles de la confiance que l'on peut dire systémique (qui concerne les environnements mêmes des interactions et pas seulement ces interactions), et plus généralement celle de la confiance dans nos activités et dans leurs contextes. Ensuite seulement je reprends le dossier des questions prétendument centrales de la confiance en ligne. Il me semble en effet que dans les usages divers de l'internet l'utilisateur a d'abord à faire à des environnements numériques (le plus souvent des interfaces visuelles, graphiques et textuelles), et non pas directement à des personnes. Dans les relations interpersonnelles ordinaires, les autres personnes sont au premier plan, et les environnements sont au second plan. Cette différence justifie que l'on ne néglige pas les problèmes de la confiance systémique dans le cas des interactions en lignes. Comme les environnements numériques simulent ou prolongent pour la plupart l'environnement dit naturel, la question de la confiance épistémique, qui est une question de fiabilité des apparences et de nos propres pouvoirs (alors que la question de la confiance pratique est celle de la fiabilité des agents), ne doit pas non plus être négligée. C'est ma proposition principale.

II. La condition de confiance systémique et les problèmes de la confiance épistémique

Sans qu'il soit possible dans l'espace de ce chapitre de développer une phénoménologie ou une généalogie de la confiance cyberspatiale, il importe de situer l'émergence du questionnement sur la confiance numérique dans les environnements technologiques concrets de l'utilisation quotidienne. Un changement important dans l'expérience de l'utilisateur au sein de l'informatique personnelle est apparu très tôt dans l'histoire de cette technologie. Les interfaces graphiques se sont substituées au *shell* et à la ligne de commande. Le projet même d'une informatique personnelle dans les années mille neuf cent quatre-vingt supposait ce changement. Il rendait possible un écart très important entre les connaissances de l'administrateur et les connaissances requises de l'utilisateur – cet écart allait se creuser en proportion de la popularisation de la technologie. Le succès d'une informatique personnelle passait par le développement d'environnements de travail « conviviaux » ou transposant l'expérience familière du bureau.

Une différence majeure entre l'expérience d'un environnement numérique et celle de l'environnement naturel (le terme inclut ici les artéfacts, car il renvoie à l'expérience quotidienne) est que l'étendue, la solidité et la sécurité de la première sont à proportion de certaines connaissances théoriques autant que techniques, tandis que l'exercice plein des compétences dans un environnement naturel (à l'exception de contextes très techniques, par exemple industriels) ne suppose pas du tout de la part des acteurs la connaissance analytique des processus et mécanismes qui les sous-tendent. Sauf dans le cas d'une pathologie, il n'y aurait aucun sens à conseiller à quelqu'un d'être prudent lorsqu'il dresse une table pour un repas au motif qu'il reste ignorant des processus psychophysiologiques à l'œuvre derrière les gestes qu'il accomplit à cette fin (en revanche, dresser une table suppose la maîtrise de certains codes). En ce domaine, il y a des tas de choses que nous savons faire sans savoir ni avoir besoin de savoir comment exactement nous les faisons, et *a fortiori* sans avoir besoin de vérifier la fiabilité des processus sous-jacents.[13] Les ferions-nous si nous avions besoin de vérifier que nous ne sommes pas trompés tandis que nous les faisons ? On peut sérieusement en douter. Un grand nombre d'actions ne peuvent être accomplies aisément que sous condition de non seulement de confiance pratique, mais aussi de confiance systémique, expression qui rassemble sous ma plume à la fois ce que Luhmann appelle « familiarité » et ce qu'il appelle *confidence* (qu'il distingue de *trust*)[14]. S'appuyant sur Luhmann, Nissenbaum insiste sur le fait que la confiance accroît le pouvoir d'agir, en réduisant le spectre des options et par là la nécessité de calculs stratégiques très complexes[15]. La confiance permet la focalisation pratique. Il y a des choses qu'on fait seulement ou qu'on fait mieux en étant en confiance, c'est-à-dire en étant dispensé de procéder à des

13 Dans l'histoire de la philosophie moderne, c'est le scepticisme (par exemple celui d'un David Hume) qui a attaqué les assurances de la métaphysique et de la théologie relatives aux conditions d'arrière-plan de nos opérations mentales. Mais ce scepticisme même concédait que ses doutes n'avaient pas lieu d'être dans la vie de tous les jours. Les personnes étaient appelées à la vigilance dans la conduite de leur esprit et à la critique rationnelle dans l'appréciation de ce qui était proposé à leur croyance ; mais elles n'étaient pas incitées à la folie d'une mise en doute quotidienne des conditions mêmes de la vie de l'esprit.

14 Dans un article initialement paru en anglais (in Gambetta 1988), Luhmann distingue trois formes, *familiarity*, *confidence*, *trust*, et montre comment le *trust* est à la fois distincts des deux premières formes et conditionné pratiquement par elles. « La familiarité est un fait inévitable de la vie ; la confiance décidée (*trust*) est une solution aux problèmes spécifiques posés par le risque. Mais cette dernière doit être produite dans un monde familier et des changements peuvent avoir lieu dans les traits familiers du monde qui auront un impact sur la possibilité de développer la confiance dans les relations humaines. C'est pourquoi nous ne pouvons pas négliger les conditions de la familiarité, et ses limites, quand nous entreprenons d'explorer les conditions de la confiance décidée. » (Luhmann 2001, p. 18) Dans cette traduction de Louis Quéré, *trust* est rendu par « confiance décidée », pour être distingué de *confidence*, traduit par « confiance assurée ». La différence principale entre familiarité et confiance assurée est que seconde suppose un contexte pratique de danger, que ne suppose pas la première. La différence principale entre confiance décidée et confiance assurée est la différence entre risque et danger : que la confiance décidée est corrélative de l'appréhension d'un risque, qui n'existe que sous la condition d'une prise de risque, donc d'une action. Dans la même veine, Nissenbaum distingue entre la confidence dans les communications et systèmes informatiques eux-mêmes et le *trust* à l'égard des agents, qui est lui-même de l'ordre de l'action (Nissenbaum 2001, p. 105).

15 La confiance réduit également le besoin de recours aux protections juridiques formalisées. Voir Fukuyama 1997, p. 150.

vérifications ou des estimations plus nombreuses[16]. Mais précisément est-il raisonnable de s'en dispenser dans n'importe quel environnement ?

Le développement des interfaces conviviales, qui imitent un univers familier, n'a pas inauguré un âge de la confiance, même une fois passée la phase d'apprentissage. La familiarité avec les machines, même virtuose, ne vaut pas confiance, simplement parce que des connaissances élaborées dans le domaine de l'informatique demeurent indispensables à la sécurisation de cette expérience. Si l'on passe des machines aux réseaux et aux relations entre les clients et les serveurs, le degré d'élaboration requise augmente énormément. Or précisément la facilitation de l'expérience de l'utilisateur repose essentiellement sur une stratégie de la part des firmes informatiques qui consiste à le dispenser des connaissances requises, voire, dans le cas de systèmes très fermés, à organiser l'inaccessibilité de ces connaissances, souvent au nom d'un principe de sécurité par l'obscurité[17].

La continuité entre les interactions naturelles et les interactions en ligne est fictive. Cette fiction est savamment entretenue par la convivialité des environnements informatiques, et s'est renforcée au début de notre siècle. Sauf dans les cas (certes très fréquents) où l'internet est utilisé comme un moyen de communication interpersonnelle ou entre une personne et une institution, ce sont souvent des robots, machines ou interfaces qui se substituent aux autres personnes, voire les simulent, et sont considérés implicitement, de manière presque superstitieuse comme des sortes d'agents[18]. En outre, l'exploitation industrielle, en pleine expansion, des données massives (*big data*) constitue un phénomène sans équivalent dans les interactions sociales familières, de sorte que les questions de la confiance se trouvent fortement dépaysées lorsqu'elles sont posées dans un univers où les réponses aux requêtes d'un individu suivent des algorithmes qui se substituent à des décisions humaines.

16 Nissenbaum 2001, p. 106-107.

17 A contrario, le nouvel âge de l'accès libre aux sources, y compris au code source de logiciels ou de systèmes d'exploitation, sera-t-il celui d'une nouvelle confiance ? Il est certainement trop tôt pour répondre à cette question. Mais on peut avancer que, dans l'hypothèse où les sources deviendraient publiques, comme la quasi totalité des utilisateurs seraient incapables de les auditer, on aurait toujours besoin de confier à un tiers compétent le soin de s'assurer de leur fiabilité du point de vue de la sécurité informatique – la question de la confiance à l'égard de ce tiers restant vive. Il y a là un exemple d'un phénomène de régression de la confiance sur lequel je reviendrai dans la dernière section.

18 Un exemple très récent de ce type de simulation est fourni par l'assistant personnel qui se présente comme un haut-parleur doté de microphones et animé par un logiciel spécialisé dans le traitement de requêtes exprimées dans une langue naturelle, fonctionnant de manière bidirectionnelle, capable d'analyse et de synthèse vocales, de telle façon qu'il peut répondre oralement à des questions posées oralement. Ce genre d'équipement peut susciter des inquiétudes car il pourrait être susceptible d'être transformé en outil d'espionnage. Mais le point intéressant, ici, est qu'il présente un écart extrême entre simplicité et facilité pour l'utilisateur, d'un part, et, d'autre part, complexité et obscurité des technologies sous-jacentes, appuyée sur le traitement des données massives (produites par les requêtes de millions d'utilisateurs sur un moteur de recherche).

Il y a eu un premier âge de l'informatique personnelle où tout le problème était de rapprocher l'expérience de l'utilisateur de son expérience naturelle. La machine matérielle et surtout logicielle devait être adaptée aux gestes familiers, y compris à ceux qui avaient été acquis dans le cadre d'une autre technologie, bientôt désuète, comme le clavier QWERTY ou équivalent. Avec les smartphones et les tablettes, les interfaces homme-machine que sont le clavier ou la souris ont pris un sérieux coup de vieux. Les gestes humains s'adaptent désormais à l'écran tactile. Cette familiarité, virtuose chez les plus jeunes, n'est pas de même nature que celle du premier âge. Elle repose sur une appropriation corporelle des outils qui n'a plus guère à être médiée par la simulation de l'environnement naturel. Mais le problème de confiance systémique est exactement le même : on ne me donne accès qu'à ce qui *facilite* mon expérience d'utilisateur, mais non pas à ce qui la *sécurise* ou l'*insécurise*. Avant d'en dire un peu plus sur les relations entre confiance pratique, confiance épistémique, et sécurité informatique, il importe d'insister sur une distinction conceptuelle.

Confiance pratique et confiance épistémique : la différence

Voici un exercice qui montre que deux concepts de confiance doivent être distingués. Considérons les questions suivantes :

(1) Les choses sont-elles ce qu'elles semblent ? Cette personne est-elle ce qu'elle prétend être ?

(2) Le monde va-t-il persister dans l'existence ?

(3) Ces belles pages *web* sont-elles la vitrine d'une entreprise commerciale sérieuse ?

(4) La marchandise que j'ai commandée sera-t-elle livrée ?

Ces questions se ressemblent. Muni d'une définition de la confiance épistémique comme un ensemble d'attitudes qui tiennent les *apparences* pour fiables, tandis que la confiance pratique tient des *agents* pour fiables, on remarque que la plupart de ces questions concernent la confiance et la méfiance épistémiques. Il semble cependant que la question (4) puisse concerner également la confiance pratique, car elle porte sur une coordination d'actions et est relative à mes attentes normatives à l'égard d'autres agents. Il peut être intéressant d'analyser cette hésitation.

(4) concerne la confiance ou la méfiance pratiques notamment lorsque l'on se trouve dans les circonstances suivantes : On pense qu'on achète ou qu'on va acheter une marchandise sur l'internet ; que c'est à un marchand que l'achat est fait. On peut alors avoir confiance dans l'enseigne, ou faire plus ou moins confiance au marchand, ou avoir des doutes sur sa capacité à tenir les délais de livraison annoncés, ou encore on peut être trop méfiant pour ne pas utiliser la clause de rétractation. C'est typiquement un problème de confiance pratique parce que le contexte est fortement diachronique : je prends immédiatement et facilement des engagements dont je ne saurai que plus tard si j'avais vraiment raison de les prendre[19].

Maintenant, changeons les paramètres et imaginons que l'on se trouve dans des circonstances où l'on décrit soi-même ce qu'on fait non pas en termes d'achat d'une marchandise, mais comme consistant à « se faire arnaquer sur internet »[20]. On n'est pas du tout prêt à décrire ce que l'on fait comme un achat sur internet. On pense qu'on est peut-être (ou probablement, ou certainement) en train de se faire arnaquer. C'est peut-être un faux marchand, une fausse marchandise, un faux achat. Alors il n'y a pas de place ici pour la confiance ou la méfiance pratiques, mais seulement pour la méfiance épistémique. Je me demande si la marchandise que j'ai commandée sera livrée, simplement parce que je subodore que ce qui m'est proposé n'a que l'apparence d'une transaction marchande et que l'annonce d'une livraison est un mensonge.

C'est la différence entre un usage métaphorique et un usage propre du terme « arnaque ». Si le produit est trop cher ou s'il est livré trop tard, voire n'est pas livré en raison d'une incapacité ou du manque de sérieux de l'entreprise, c'est métaphoriquement une arnaque. On se situe encore dans un problème de confiance pratique. Si l'on a des motifs de croire que c'est réellement une arnaque, on sort de ce problème. Il faudrait ne pas maîtriser du tout le lexique de la confiance et de la méfiance pratiques pour dire que l'on fait peu confiance ou que l'on se méfie de quelqu'un que l'on tient littéralement pour un arnaqueur.

Mais cette démarcation est trop forte et injuste, y compris dans sa traduction dans la théorie. Si j'abordais

19 Weckert estime que pour cette raison le commerce en ligne suppose plus de confiance que le commerce direct : « Normalement quand j'achète des marchandises j'obtiens presque immédiatement ces marchandises et le reçu, de sorte que cela ne requiert pas beaucoup de confiance. L'achat en ligne requiert plus de confiance. » (Weckert 2005, p. 114) Cependant, on pourrait objecter à Weckert que la diachronicité forte caractérise aussi des situations très classiques du commerce ou des échanges de service.

20 Il faut garder à l'esprit qu'une arnaque exploite généralement la crédulité et non pas nécessairement la confiance (pratique). Les escrocs savent tirer parti de la méfiance autant que de la confiance. Les amorces du type « votre sécurité est compromise » sont l'équivalent en ligne des pratiques des voleurs à la fausse qualité, qui par exemples se font passer, auprès de personnes très méfiantes mais aussi très crédules, pour des policiers occupés à réprimer ou prévenir la délinquance.

théoriquement ces questions en décrétant que les conditions de la confiance pratique ne sont pas réunies, au prétexte qu'il y a des arnaques au sens littéral dans les interactions en ligne, je négligerais arbitrairement tous les aspects sous lesquels l'expérience des utilisateurs est dans la continuité des interactions interpersonnelles ou des interactions entre personnes et institutions. Je commettrais l'équivalent d'un sophisme sceptique bien connu, qui consiste à inférer de l'expérience particulière de la tromperie des sens au caractère essentiellement et universellement trompeur des sens. Que peut-on alors retenir de cet exercice ? Que la confiance pratique suppose des paramètres assez finement ajustés, notamment que les apparences ne soient pas soupçonnées au point extrême où le problème de confiance épistémique deviendrait vraiment envahissant.

Un exemple de problème épistémique de crédit

Les problèmes de confiance en ligne sont-ils d'abord des problèmes de confiance épistémique ? Oui. Sont-ils seulement des problèmes de confiance épistémique ? Non. Il y a cependant des régions dans les usages de l'internet où les problèmes de confiance épistémique sont véritablement hégémoniques. C'est le cas de tout ce qui touche à l'*information* en ligne.

Le degré de crédit que nous accordons à un dictionnaire ou une encyclopédie sous la forme d'un livre peut être garanti par notre jugement et notre expérience, ou par l'autorité qui est attachée à la maison d'édition, etc. Une contrefaçon de ce type d'ouvrage, si elle se borne à la copie illégale, n'est pas moins fiable. Mais l'information peut se périmer partiellement ; je peux me procurer une nouvelle édition. La capacité d'une maison à produire une nouvelle édition d'un ouvrage de ce type est une manifestation de son sérieux. On pourrait multiplier ainsi les considérations sur les fondements de la crédibilité d'une source d'information. La donne n'est pas substantiellement modifiée lorsqu'il s'agit des services de dictionnaire ou d'encyclopédie en ligne proposés de manière fermée par une maison d'édition. Les choses semblent se compliquer lorsque ce type de service repose sur une communauté ouverte de collaborateurs ; mais là encore on ne sort pas d'un problème classique de crédibilité et d'autorité, qui est distinct des questions de confiance et de méfiance pratiques. Bref, lorsque l'on se demande si telle source d'information est digne de confiance, on se demande si elle mérite d'être crue, et non pas si sa volonté va s'ajuster avec la nôtre. Certes, on pourrait soutenir que le crédit est assimilable à la

délégation à l'action d'une institution ou d'une personne du soin de vérifier une information, et qu'il y a quelque chose de l'ordre de la confiance pratique dans cette délégation. Mais il demeure que l'attitude qui permet de se prémunir contre les effets de la crédulité est la méfiance épistémique, et non pas la confiance pratique « par excellence », la confiance décidée.

Lorsque les réseaux sont utilisés plutôt comme des sources d'information que comme des vecteurs d'échanges financiers, commerciaux, ou d'autres interactions pratiques, on retombe dans un problème traditionnel d'autorité qui appelle des solutions qui reposent non pas sur des mécanismes de sécurisation, mais sur des vertus épistémiques, comme le suggère une publicité d'un acteur majeur des réseaux sociaux, qui invite à la méfiance dans la propagation de nouvelles ou de prétendues informations : « Méfiez-vous des titres… Examinez attentivement l'URL… Effectuez des recherches sur la source… Faites attention aux mises en forme inhabituelles… Tenez compte des photos… Contrôlez les dates… Vérifiez les preuves apportées… Consultez d'autres articles… »[21]

Contre la confusion de la confiance et de la sécurisation

Revenons à la question des rapports entre confiance et sécurité. Par la confiance « par excellence », un agent manifeste qu'il est véritablement un agent, soit qu'il accepte les risques auxquels l'expose son action, y compris le risque de la défection ou trahison d'autrui, simplement parce que la confiance de cet agent est acceptation de la liberté d'autrui. Mais il est aussi essentiel à la confiance de pouvoir être rationnelle, au sens où elle répond à des raisons perçues, qui rendent le risque acceptable. Selon John Weckert, il y a un lien étroit entre confiance et reconnaissance de l'autonomie, alors qu'à l'inverse la surveillance traduite la défiance[22].

Pour cette raison, il est difficile de se satisfaire de l'approche qui est adoptée le plus souvent par les sciences de l'information et de la communication et qui rapporte exclusivement la question de la confiance à celle des technologies de sécurisation. Par exemple, Şerif Bahtiyar et Mehmet Ufuk Çağlayan définissent la confiance comme « l'attente de sécurité d'une entité de la part d'un service selon l'information disponible relative à

21 https://fr-fr.facebook.com/help/188118808357379 (consulté le 1er mai 2017).
22 « Si on a confiance en moi, on ne me surveille pas pour s'assurer que je fais ce qu'on me dit de faire. » (Weckert 2005, p. 98)

l'évaluation de sécurité de cette entité »[23]. Il me semble que cette définition n'est pas une définition de la confiance en ligne en général, mais seulement de la confiance à l'égard de la sécurité d'un service. Or nos attentes à l'égard des services en ligne ne concernent pas seulement la sécurité, sauf lorsqu'il s'agit de services dédiés essentiellement à cet aspect.

Contre ce type d'approche, Nissenbaum estime aussi qu'il ne s'agit pas de « sécuriser » la confiance, mais de la « nourrir »[24]. Elle invite à comprendre la différence subtile entre la sécurisation et « les preuves, les signes, les signaux et les indices », jamais certains, qui fondent la confiance sans la garantir[25]. La différence est aussi que les fondements de la confiance n'annulent jamais la responsabilité épistémique. Un autre argument intéressant de Nissenbaum contre les séductions d'une sécurisation de la confiance est que le danger est susceptible de venir des *insiders* autant que des *outsiders* : il n'y a pas que les intrus (individuels) qui nous fournissent des raisons de nous méfier, mais les moyens organisationnels de la surveillance des activités, par exemple à des fins commerciales, par le *web-tracking*[26]. Bref, un univers de confiance n'est ordinairement pas un univers « sécurisé ». Nissenbaum, qui critique la vision de la « fiabilité comme sécurité ou de la confiance par la sécurité »[27], a raison sur ce point. Mais en même temps un certain degré de sécurité de l'environnement est indispensable, sans quoi la méfiance épistémique devient radicale et la confiance pratique ne peut pas se déployer.

L'exemple du Web of Trust

Je ne voudrais pas adhérer à ce consensus des philosophes contre la sécurisation intégrale de la confiance au point de disqualifier toutes les questions pratiques et théoriques que les informaticiens traitent sous la rubrique « confiance et sécurité ». Il est un domaine majeur, et dont l'importance pour le développement des services en ligne est de plus en plus manifeste, dans lequel le cousinage de la confiance et de la sécurité ne peut être considéré que d'un bon œil : celui des communications qui intègrent des solutions aux problèmes de la

23 Bahtiyar et Çağlayan 2013, p. 291.
24 Nissenbaum 2001, p. 121. Selon elle, la confiance est exposition au risque, vulnérabilité assumée.
25 Nissenbaum 2001, p. 122.
26 Nissenbaum 2001, p. 125-129. Sur ce point, Nissenbaum est heureusement contrainte de contrevenir à sa décision de ne pas parler de la *confidence* systémique.
27 Nissenbaum 2001, p. 103.

confidentialité et de l'identification. Je prends l'exemple de GnuPG, qui fournit notamment une réponse au problème de certification de l'identité.

Cette technique de sécurisation des communications donne lieu, ici, à une transitivité de la confiance. La validation d'une clef publique, qui normalement conditionne sa signature, est non problématique lorsqu'il s'agit de la clef d'une personne que l'on connaît par un contact personnel et auprès de laquelle on peut vérifier qu'il s'agit bien de sa clef. La liaison entre l'identité numérique et la personne physique est ainsi assurée (une clef de *B* est dite valide pour *A* quand *A* s'est assuré qu'elle est bien la clef qu'elle prétend être, soit la clef de *B*). Mais si nous souhaitons correspondre de manière cryptée (ou avec des signatures numériques) avec des personnes que nous ne connaissons pas personnellement, comment faire ? Nous pouvons déléguer à des personnes dont nous avons validé les clefs publiques la validation des clefs de personnes avec lesquelles nous ne sommes pas en contact. Plus exactement, nous pouvons faire en sorte que soient automatiquement valides pour nous les clefs publiques validées par une personne dont nous avons validé la clef publique, de sorte que nous pouvons les signer sans angoisse. Cependant, il saute aux yeux que cette délégation est très risquée. Il y a un nouveau problème qui tient au fait que les autres peuvent valider de manière inconsidérée les clefs de personnes avec lesquelles elles ne sont pas personnellement en contact et auprès desquelles elles n'ont pas fait de vérification. Il faut répéter qu'ici « valider une clef » signifie s'assurer qu'elle est la clef de son propriétaire prétendu.

C'est sur ce point qu'intervient une solution caractéristique du *Web of Trust* : elle consiste pour nous à assigner de manière privée des niveaux de *trust* aux clefs publiques des autres.[28] Assigner un niveau de *trust* à une clef, c'est noter pour soi-même le degré auquel on fait confiance au propriétaire de cette clef pour ne pas signer inconsidérément les clefs des autres. Le *trust* en ce sens n'est pas la même chose que la validation. Cette dernière consiste pour *A* s'assurer de la liaison entre une clef et *B* qui en est le propriétaire. Un utilisateur responsable ne signe que les clefs validées avec soin. Le *trust* consiste pour *A* à considérer *B* comme capable d'une pratique fiable de validation. Les niveaux de *trust* de GPG sont les suivants : par défaut, *unknown* ; *none* quand on juge que *B* n'a pas cette capacité ; *marginal*, quand on juge que *B* a connaissance de la nécessité de ne signer que les clefs validées avec discernement ; *full*, quand *A* accorde la même valeur à la signature de *B* qu'à

28 Cette présentation exploite un manuel, *The GNU Privacy Handbook* (Ashley, Copeland *et al.* 1999, chap. 3).

la sienne propre.

Comme normalement la signature d'une clef suppose sa validation (un utilisateur ne devrait signer que les clefs validées, c'est-à-dire que les clefs dont il s'est assuré de la liaison avec leur propriétaire prétendu), le fait qu'une clef soit signée indique qu'elle a été validée, à la condition que l'on ait confiance dans la capacité du signataire à ne signer que ce qui a été validé correctement. Cet usage du terme *trust* peut surprendre, mais il me semble correct, car il s'agit bien de la confiance d'une personne à l'égard d'une autre personne. En un sens voisin, l'acrobate de cirque a confiance dans son partenaire pour le réceptionner avec compétence et efficacité. Dans le cas du *Web of Trust*, ce n'est pas l'habitude qui est la source de cette confiance, mais une décision (la décision informée d'assigner un certain niveau de confiance).

L'expansion du « réseau de la confiance » est possible (sans entrer dans les détails, ni indiquer ce qui relève du paramétrage individuel, comme la limite assignée au chemin de confiance, qui *personnalise* ce réseau) parce que GPG considère comme valide pour *A* une clef qui a été signée soit par un propriétaire de clef signée par *A* et à qui *A* a assigné le niveau de confiance « *full* », soit par plusieurs propriétaires de clefs dont *A* a signé les clefs et à qui *A* a assigné le niveau « *marginal* ». Comme les personnes avec lesquelles *A* interagit ont aussi leur propre *Web of Trust*, des chemins de confiance sont tracés qui permettent à *A* de bénéficier des validations des autres, dans des limites qu'il a pu fixer. La confiance ici est transitive.

GnuPG emploie la notion de confiance (plus exactement, une forme de la confiance qui est la délégation) pour résoudre le problème pratique d'assurance de l'identité. Généralement, on considère que le manque profond d'assurance de l'identité dans les interactions en ligne empêche le développement de la confiance. Ici, ce sont des pratiques de confiance qui viennent compenser ce manque. GPG est exemplaire parce qu'il enseigne que la sécurisation ne peut pas dispenser de l'exercice d'une responsabilité autant épistémique que pratique, ici celle par laquelle chacun peut définir son réseau de confiance. GPG suppose des connaissances techniques (plus exactement une maîtrise technique et du soin) minimales (mais non minimes).

La communication des clefs publiques emprunte plusieurs voies, dont le dépôt sur des serveurs. Il est intéressant de noter que l'acte de prendre l'initiative de communiquer sa clef publique dans un courriel sans que

le destinataire ne l'ait demandé est un geste de confiance au sens de la confiance « par excellence ». Comme le remarque l'auteur du guide dont s'autorise cette présentation : « Si vous publiez votre clef, alors vous faites en sorte qu'il est beaucoup plus acceptable pour les autres de publier leurs clefs. En outre, vous facilitez pour les autres l'ouverture d'une communication sécurisée avec vous puisque vous avez pris l'initiative et que vous montrez que vous employez. »[29] On touche ici, au-delà des questions de confiance systémique, à une mise en œuvre de ce que Pettit appelle confiance « dynamique »[30].

Cependant la forme de confiance ici est spécifiée par les actions qu'elle permet : la pratique de la signature numérique ou du chiffrement du courriel ou de fichiers. L'usage de GnuPG à des fins de communication suppose d'abord que plusieurs personnes désirent s'engager dans des actions particulières qu'elles peuvent décrire, par exemple, comme des actions de sécurisation des courriels. Le *Web of Trust*, pour cette raison, ne peut prétendre être le paradigme d'une solution au problème global de la confiance dans les interactions en ligne.

III. Obstacles à la confiance en ligne et problèmes épistémiques

Cette dernière section poursuit l'exploration des relations entre les dimensions épistémique et pratique du problème de la confiance en ligne en reprenant le *locus classicus* des études sur le sujet : les obstacles à la confiance en ligne. Dans son article déjà cité, Nissenbaum discute une liste d'obstacles qui reprend en partie celle qu'avait dressée Philip Pettit et que l'on retrouve aussi chez d'autres auteurs :

(a) Le manque d'identité, l'anonymat.

(b) Le manque de caractéristiques personnelles, la désincarnation.

(c) L'inscrutabilité des contextes, indétermination des normes et des rôles sociaux et professionnels.

Ces premiers manques correspondent à une invisibilité que Pettit dramatisait par cette formule : « Sur l'internet nous portons tous l'anneau de Gygès[31]. » Cela constitue une exagération au regard des capacités d'observation des conduites sur l'internet qui sont désormais accessibles non seulement aux firmes, mais aux personnes.

29 Ashley, Copeland *et al.* 1999, chap. 4.
30 Voir, ci-dessous, la section « Confiance et fiabilité ».
31 Pettit 2004, p. 118.

Du côté des experts en informatique et communication, on a tendance à prétendre que des mesures de sécurité peuvent remédier à cela. Nissenbaum en dresse cette liste :

(1) Le contrôle d'accès.

(2) Le traçage de l'identité.

(3) La surveillance.

La thèse de Nissenbaum est que les mesures de sécurité, aussi nécessaires soient-elles, ne sont pas des solutions qui suffisent à installer la confiance. Dans ce qui suit, je discute certains des « obstacles », en particulier (a), celui de l'identification, et j'attire l'attention sur la dimension épistémique des difficultés et des solutions.

Avant d'aller plus loin, il convient d'insister sur le caractère contingent de certains obstacles, qui sont très dépendants de l'état des pratiques et des techniques. Weckert voit dans l'obstacle (b), la désincarnation (*disembodiment*), à la suite de plusieurs auteurs, une difficulté majeure pour la confiance en ligne. La communication sur l'internet ne mobilise pas le langage corporel (sauf dans des situations très spécifiques de communication vidéo), mais « s'appuie presque exclusivement sur la langue écrite »[32]. C'est ce qui frappait le plus lorsque l'on abordait la phénoménologie des interactions en ligne à un stade antérieur du développement des usages de l'internet. Mais qu'en est-il dans un univers qui a recours non seulement aux « avatars », mais aussi aux « émoticônes » ou « emoji » ? Ne sont-ils pas les moyens de construction d'une identité apparente, y compris visuelle ? Il y a un apprentissage possible de nouveaux moyens de signifier les intentions, d'adapter la communication aux exigences sociales, voire de trouver une forme de courtoisie ou de civilité. La désincarnation ne saurait être considérée comme une propriété essentielle de la vie numérique ; elle est plutôt une sorte de handicap que des pratiques très évolutives et adaptatives, autant que de nouvelles techniques, parviennent à compenser en partie. Cependant, cette identité construite est aussi une identité largement contrôlable par soi, ce qui la distingue plus fortement des modalités de la présentation de soi en face à face[33].

Il est difficile de considérer tous les « obstacles » à la confiance en ligne comme intangibles et définitifs. Il en va ainsi, par exemple, de l'obstacle de l'anonymat : des identités numériques peuvent avoir suffisamment de

32 Weckert 2005, p. 107.

33 Je remercie Jacopo Domenicucci de sa suggestion sur ce dernier point.

stabilité pour permettre le recueil d'une expérience des interactions en même temps qu'elles cachent des identités réelles. De plus, cette situation n'est pas très différente de celle que l'on rencontre dans le monde social moderne. Les recherches sur la confiance en ligne ne devraient-elles pas moins sous-estimer l'importance de l'adoption de rôles et le sérieux des identités fictives sur l'internet ? Un ouvrage de Patricia Wallace paru il y a presque deux décennies, qui adoptait une perspective de psychologie sociale, repérait déjà les transformations des modes de présentation de soi dans le cyberespace[34]. Bref, ce sont des difficultés qui peuvent être atténuées, sinon levées, par un ajustement des pratiques et des technologies. Par exemple, le développement du commerce en ligne aurait pu paraître peu probable il y a quelques décennies. Et ses dispositifs sont susceptibles d'être encore améliorés[35].

L'absence d'historique et les difficultés d'identification

Il est temps d'aborder le cœur de l'application de la confiance pratique aux interactions en ligne. S'agissant de ce que Pettit appelle *evidence of file*, l'historique des interactions, Simpson estime qu'il peut faire complètement défaut dans des interactions en ligne avec des personnes que l'on n'a jamais croisées, ce qui constituerait une différence majeure avec le monde naturel. Mais il me semble, au contraire, que le problème de la confiance pratique se pose de manière pertinente dans ce type de situation, semblable à celle de l'état de nature hobbesien. Chacun ne connaît que très peu de personnes parmi celles avec lesquelles il est appelé à interagir. C'est la condition de l'homme moderne[36]. C'est cette condition qui appelle comme solution la stratégie du *tit for tat*, comme l'a montré notamment Robert Axelrod[37], sans laquelle la confiance, par exemple

34 Wallace 1999.

35 Simpson estimait en 2011 que des mesures pratiques et des solutions techniques peuvent encore améliorer très substantiellement les systèmes de réputation dans le cadre du commerce en ligne. Par exemple, par la technique, aujourd'hui assez diffusée, qui consiste à labelliser comme confirmés les avis de personnes qui ont acheté le produit qu'elles évaluent (Simpson 2011, p. 29-30). Remarquons que cet outil de la confiance par la réputation auprès de tiers suppose une confiance épistémique dans cette labellisation.

36 A la suite de Francis Fukuyama, Weckert (art. cit.) estime que les interactions en ligne rendent difficiles l'installation d'un cadre de confiance, en l'absence de communauté particulière qui partage des normes et des valeurs (voir Fukuyama 1997, p. 36, p. 151). Par ailleurs, André Orléan, dans une analyse qui ne porte pas sur les interactions en ligne, a insisté sur l'inscription sociale du recours à la réputation, dans des « réseaux sociaux », et sur la dépendance de la confiance à l'égard de normes et de « ressources non économiques » (Orléan 2000, p. 59-77). Mais on doit nuancer ce type d'opposition. Il y a une fiction de communauté qui accompagne la plupart des usages de l'internet. De plus, il peut y avoir une communauté de confiance qui ne repose pas sur des valeurs partagées. L'objection de Fukuyama serait parfaitement pertinente s'il s'agissait du concept « épais » de la confiance (*thick trust*, l'expression est de Williams 1995, p. 120), par exemple celui de la *fides* interpersonnelle, de l'amitié, ou des liens familiaux, tels que pouvaient les entendre (ce n'est qu'un exemple) les anciens Grecs et Romains. Mais le concept moderne de confiance ne suppose pas tant, et est même indépendant d'une communauté morale substantielle.

37 Axelrod 1992.

dans des contextes commerciaux, diplomatiques, ou autres, ne pourrait jamais s'installer. Le problème propre à la confiance en ligne n'est pas celui de la nouveauté et de l'étrangeté des autres. Certes, on ne connaît pas leur historique passé, mais la différence avec les interactions naturelles est que, dans de nombreux cas, l'historique futur ne va probablement pas se constituer. Les personnes avec lesquelles je commence à interagir, si je ne les connais pas autrement, peuvent s'évaporer. Simpson le remarque : « Ce qui est caractéristique de la vie en ligne, c'est la facilité qu'il y a à quitter n'importe quelle relation »[38].

Il n'est pas certain que les usagers entendent renoncer à leur rêve de furtivité. L'inquiétude relative à la surveillance de l'internet conduit nombre d'utilisateurs habiles à s'assurer que leur identité ne sera que très difficilement traçable. Par exemple, ces utilisateurs peuvent avoir recours à *The Onion Router* qui se superpose à l'internet et procède au chiffrement des communications TCP. Est-ce à dire que, de manière assez paradoxale, ces utilisateurs cherchent à établir la confiance dans l'usage de l'internet grâce à l'anonymat ? Ce dernier ne fait-il pas obstacle à la confiance ? Je crois qu'il faut plutôt en conclure que cette confiance unilatérale (analogue à celle de la personne qui est « en confiance » parce qu'elle peut voir sans être vue) n'a que très peu de rapports avec les formes de la confiance dont nous parlons ici. Elle a plus d'affinités avec une sécurisation extrême qu'avec une société de la confiance.

Turilli, Vaccaro et Taddeo ont réagi à la thèse selon laquelle, puisque les conditions nécessaires de la confiance ne sont pas remplies dans un environnement en ligne, la confiance en ligne ne peut pas vraiment exister. Selon ces auteurs, la littérature qu'ils critiquent s'est accordée sur deux conditions nécessaires : (a) l'existence d'un arrière-plan culturel, moral et institutionnel partagé, qui permettrait une sorte de pression normative incitant au retour de la confiance, notamment par la sanction sociale de la défection ; (b) l'assurance quant à l'identité diachronique du *trustee*. La stratégie de Turilli *et al.* consiste à nier que la condition (a) soit une condition nécessaire et à soutenir que la condition (b) est bien remplie : dans un environnement en ligne le *trustee* peut avoir une identité diachronique et être par là sujet à l'évaluation de sa réputation. S'appuyant sur des travaux de psychologie sociale[39], les auteurs montrent que donner de l'importance à la condition (a), qui réduit

38 Simpson 2011, p. 32.
39 Yamagishi *et al.* 1999.

excessivement la prise de risque, revient à confondre la confiance avec l'assurance[40]. On peut leur donner raison sur ce point, avec cette réserve que les choses ne sont pas noires ou blanches dans l'application d'un concept dont la structure est celle d'une échelle de formes.

La condition d'identification est plus difficile à apprécier. Contre Pettit *et al.*, les auteurs rappellent que la condition d'interaction physique n'est pas présente non plus dans les échanges commerciaux classiques de longue distance (notamment à une époque où le téléphone etc. n'existait pas), ce qui n'empêche pas la confiance. De même, la réputation indirecte (auprès de tiers) joue efficacement sans cette interaction physique[41]. Ces deux points peuvent être concédés. Cela montre seulement que l'interaction physique n'est qu'un moyen parmi d'autres de s'assurer de l'identité.

S'agissant de la condition (b), les auteurs pensent qu'elle est remplie, soit par l'identification (liant une identité en ligne et une identité réelle) que permettent des technologies, soit même en l'absence d'une telle identification. Dans le premier cas, des technologies ou des procédures formalisées de vérification de l'identité peuvent prendre le relais de l'interaction physique. C'est le deuxième cas, celui de l'absence d'identification, qui est la plus problématique. Dans cette situation où l'identification ne peut pas être complète, la réputation, qui, selon ces auteurs, consiste dans l'enregistrement de l'historique des activités, est attachée à une identité en ligne seulement. On peut objecter à Turilli *et al.* qu'ils tendent à confondre la sensibilité à la réputation (qui concerne aussi bien le *trustor* que le *trustee*) et la surveillance des interactions. Bien plus, les procédures de traçage sur lesquels les auteurs s'appuient sont précisément de celles qui sont susceptibles de miner la confiance.

Confiance et fiabilité

Ces auteurs s'appuient sur une approche la confiance qui fait de la fiabilité au sens fort (*trustworthiness*) le déterminant principal de la confiance : celle-ci est fondée sur, et garantie par, la fiabilité. Il me semble que cette vision est réductrice et ne tient pas compte du fait que la confiance décidée peut s'accorder sans être

40 « Quand il n'y a pas de risque ou quand il est largement réduit par des structures sociales et morales, la confiance est remplacée par l'assurance ou par un certain degré d'assurance. » (Turilli *et al.* 2010, p. 337)
41 Turilli *et al.* 2010, p. 338-339.

initialement conditionnée par une estimation de la fiabilité, et sans être pour autant nécessairement irrationnelle (cependant, elle deviendrait irrationnelle si, dans un jeu répété, elle se maintenait en dépit de l'évidence de la défection). C'est l'erreur de la théorie cognitiviste de la confiance que défend Russell Hardin, selon qui « les déclarations "je crois que tu es digne de confiance" et "je te fais confiance" sont équivalentes »[42]. Elles ne sont pas équivalentes. Car *A* peut faire confiance à *B* en le traitant comme s'il était digne de confiance sans nécessairement croire qu'il est digne de confiance. Et *A* peut croire que *B* est digne de confiance et cependant ne pas lui faire confiance. De plus, si la confiance était réductible à la croyance dans la fiabilité, en tant que croyance elle répondrait à la fiabilité perçue et serait susceptible d'être « vraie » ou « fausse » de ce point de vue. Mais la dynamique de la confiance est telle qu'elle peut *rendre* l'autre fiable.

En ce sens, Pettit a raison de caractériser le *trust* (ce que j'ai appelé la confiance « par excellence ») comme une *reliance* dynamique, dont les deux conditions sont (a) que les personnes sur lesquelles l'on compte soient conscientes, grâce à ma signalisation, du fait que je compte sur elles et (b) que, par cette signalisation de ma confiance, j'espère leur fournir un motif d'agir comme je l'attends[43]. Il est essentiel à la confiance interpersonnelle de *se signaler* afin d'encourager la dynamique de la confiance. Traiter quelqu'un comme digne de confiance n'est pas la même chose que le traiter comme prévisible. C'est l'investir d'une forme de responsabilité à l'égard de mon attitude de confiance. Et à l'intérieur même de la sphère du *trust* proprement dit (qui est corrélative de la fiabilité au sens fort), Pettit distingue entre la fiabilité substantielle des vertueux (« confiance primaire », dans son vocabulaire) et celle qui est présumée indépendamment de toute information sur les vertus ou autres dispositions, et qui repose sur le désir de l'estime, la sensibilité à la réputation (« confiance secondaire »). Grâce au circuit de la réputation, la confiance peut fonctionner bien au-delà de la communauté des amis et indépendamment des vertus. Ce que la confiance placée dans l'autre signale, c'est l'estime dans laquelle l'autre est tenu, une reconnaissance que l'autre peut aussi désirer. La signalisation de la confiance est ainsi une manière d'« amorcer » une réponse d'autrui qui montre en retour qu'elle est justifiée[44].

Le pari de la confiance est lui-même une incitation susceptible d'induire la conduite supposée. C'est pourquoi

42 Hardin 2002, p. 10.
43 Pettit 2004, p. 110.
44 Pettit parle de « bootstraps operation » (Pettit 2004, p. 115).

une théorie de la confiance qui prétendrait que la fiabilité qu'elle suppose doit être établie, prouvée d'abord, afin que la confiance soit rationnelle, reposerait sur une compréhension incorrecte de son objet. La relation entre *trust* et *trustworthiness* est donc plus complexe qu'il n'y paraît. Si la confiance ne répond pas simplement à la fiabilité perçue, à quoi répond-elle ? Pas à des preuves de sécurité, en tout cas.

Problème de la simulation

La confiance, même sous sa forme « par excellence », se conforte ou se défait en fonction d'indices, qui constituent des raisons d'avoir ou non d'avoir confiance. Sous cet aspect, une conception plausible de la rationalité de la confiance est « évidentialiste » : la confiance répond à de l'*evidence* au sens anglais de ce terme (au sens de l'indice ou de la preuve). Précisément parce que les grands types d'*evidence* (dont la liste est dressée dans Pettit 2004) font assez largement défaut dans les interactions en ligne, les acteurs de divers services et en particulier du commerce en ligne ont développé des dispositifs qui tendent à combler ce manque, qui tiennent lieu d'*evidence*. Un problème important est que ces indices peuvent faire l'objet d'une simulation.

Simpson insiste sur ce problème de « *mimicry* » des preuves, indices et raisons de la confiance, pour qui voudrait fonder la confiance rationnelle de manière évidentialiste. La réputation en ligne, exprimée dans l'historique ou *track-record* des interactions entre usagers ou entre services et usagers, paraît être un moyen très prometteur de la constitution d'un tel fondement de la confiance. La réputation n'est-elle pas une sorte d'extension ou de relais de l'expérience personnelle, qui reste la référence en matière d'*evidence*[45] ? L'inventivité technologique peut même permettre d'améliorer la qualité des indices de fiabilité liés à la réputation en ligne, et c'est de ce côté que les acteurs de l'internet devraient placer leurs espoirs. D'autres sources d'*evidence*, comme celle du contact physique (« *evidence of face* », selon l'expression de Pettit), ne sont pas du tout opératoires dans le contexte des interactions en ligne. Sur un forum, une photographie de profil

45 On peut considérer que la réputation est l'aspect interpersonnel d'une expérience qui peut être intrapersonnelle sous la forme de l'historique des interactions. Nissenbaum assimile la réputation à l'« expérience des autres », qui peut compenser l'absence d'historique d'une « interaction directe » (Nissenbaum 2001, p. 110). La réputation est une sorte de bilan d'expérience par procuration. Cela est analogue à la distinction entre information directe et information indirecte sur les conduites des autres. En l'absence de répétition du jeu au niveau personnel (interaction à un seul coup, absence d'expérience), la réputation est une sorte de résumé de jeux répétés au niveau collectif. De cette manière, les ressources de la confiance « par excellence » peuvent être mobilisées. Dans le cas du commerce en ligne ou des interactions dans un groupe collaboratif, par exemple, la stratégie du *tit for tat* peut être mise en œuvre collectivement.

personnel fournit une *evidence of face* « quasi nulle »[46]. Dans les échanges naturels, même si nous pouvons être dupes d'un air sympathique, nos capacités de détection peuvent opérer. La conception évidentialiste de la confiance est sans doute la manière la plus solide de rendre compte de la rationalité et de l'irrationalité des conduites de confiance. Le fait que sa mise en œuvre dans le contexte cyberspatial comporte de nombreuses difficultés, notamment celle de la simulation, n'est pas une raison de l'abandonner[47].

Il est intéressant de noter ici l'articulation entre une difficulté de fondation de la confiance pratique et une méfiance épistémique relative à cette fondation. La confiance (pratique) dans la fiabilité d'un agent est conditionnée par une confiance (épistémique) dans la fiabilité d'une preuve apparente. Dans le cyberespace, ses preuves apparentes peuvent être interprétées de manière suspicieuse, non comme des preuves de la confiance, mais comme des instruments d'une tromperie organisée. Ils ne sauraient donc instituer de toute pièce la confiance, mais la supposent à leur tour. Ce problème de la simulation est, sous cet aspect, une forme particulière d'un problème plus général, celui de la régression.

La confiance épistémique dans les fondements de la réputation

Avant d'aborder le problème de la régression, prenons un exemple de cette intrication typique d'une question de confiance épistémique dans une question de confiance pratique. Selon Simpson, les dispositifs d'évaluation qui constituent la réputation d'une personne auprès de tiers, tels que ceux qu'emploient de grandes firmes du commerce en ligne, permettent de compenser les manques d'*evidence* : « Un historique fidèle [*truthful*] de la fiabilité [*trustworthiness*] passée est une excellente preuve de la disposition d'une personne à être digne de confiance, et cette disposition est un bon indicateur inductif de la fiabilité future. » Le problème, selon l'auteur, est alors le suivant : « Mais les systèmes de réputation fournissent-ils un historique fidèle de la fiabilité passée ? Ceci constitue un métaniveau d'analyse : les systèmes de réputation sont-ils eux-mêmes dignes de

46 Simpson 2011, p. 32.

47 Simpson lui-même s'appuie sur cette conception, lorsqu'il remarque que la personnalisation des résultats des requêtes sur les moteurs de recherche, ou encore des informations sur les réseaux sociaux, ou des suggestions d'achat sur les sites marchands, rend ces services non dignes de confiance selon le principe évidentialiste. En effet elle renforce le biais de confirmation, par lequel nous sommes spontanément plus intéressés par et plus attentifs à ce qui confirme nos opinions qu'à ce qui conduirait à les reconsidérer. (Simpson 2014, p. 117-119)

confiance ? »[48] Ce dernier questionnement est celui de la confiance épistémique.

Simpson pense que la difficulté vient du fait qu'il peut y avoir une stratégie de couverture, qui consiste à construire une bonne réputation à moindre coût pour mieux tromper ensuite. Il y a d'autres stratégies également présentes dans le monde réel. Il y a des stratégies facilitées dans le monde en ligne, comme celle du changement d'identité, pour se refaire une réputation. Comme le remarque Simpson, cela ne signifie pas que les systèmes de réputation soient fondamentalement viciés, car souvent nous n'hésitons pas à accorder notre confiance lors d'interactions en ligne. Mais n'y a-t-il pas ici une certaine confusion entre une question de confiance épistémique et une question de confiance pratique ? C'est la confiance pratique, sous la forme d'une confiance-pari, qui règle en pratique une difficulté qui pourrait bloquer la coopération et qui correspond à une bonne stratégie dans un jeu non coopératif : il vaut mieux renoncer si on s'expose trop à coopérer. Les personnes qui ont tendance à faire confiance surmontent ce blocage, certes à leurs risques et périls, mais aussi souvent pour leur plus grand profit. Cependant cette confiance décidée qui est à l'œuvre ne supprime pas l'autre problème, qui est celui de la confiance épistémique à l'égard des systèmes de réputation. On a agi sans tenir compte en pratique du risque de tromperie. Cette décision n'a rien à voir avec une assurance quant au crédit de certaines informations.

Bref, ce que Simpson appelle « analyse de métaniveau de la fiabilité des systèmes de réputation »[49] correspond à un problème de confiance épistémique. Celui-ci peut être lié à des conduites stratégiques telles que celle qui consiste à ne pas donner d'avis négatifs de peur des représailles. Il peut être aussi de l'intérêt d'une firme de commerce en ligne de minorer les avis négatifs ou de les rendre moins visibles.

Problème de la régression

Il s'agit de la régression de la nécessité de la confiance dans un contexte de sécurisation. Pour Weckert et plusieurs auteurs, surtout du côté des philosophes et des politistes, la sécurisation ne suffit pas à assurer la confiance, elle la présuppose, et peut la détruire en encourageant l'illusion qu'elle est superflue. Weckert montre la régressivité indéfinie du problème de la confiance systémique en procédant à une expérience de

48 Simpson 2011, p. 33.
49 Simpson 2011, p. 34.

pensée. Imaginez un réseau informatique totalement sécurisé par les meilleures technologies de protection contre les intrusions, d'authentification et de chiffrement. À l'intérieur de ce réseau, certains utilisateurs pourraient se conduire en délinquants. Ajoutez un système de surveillance pour faire la chasse à cette délinquance. Encore faudrait-il avoir confiance dans ce système de surveillance. Supposez alors que la confiance dans ce système est garantie parce que les surveillants eux-mêmes sont surveillés : « Cela repousse seulement la confiance à un autre niveau, et ainsi à l'infini[50]. » Quant à une automatisation complète, qui rendrait la sécurisation immanente à toutes les activités et semblerait dispenser de cette régression, elle suppose de toute façon la confiance dans les concepteurs de ces automatismes.

De cette expérience de pensée, Weckert devrait conclure seulement que la sécurisation suppose la confiance, de telle sorte qu'on ne peut prétendre remplacer complètement la confiance par la sécurisation. Il conclut que « la confiance ne peut pas être engendrée par la seule sécurité », ce qui est en effet suggéré par cette expérience de pensée, mais aussi qu'« en fait, trop de sécurité étouffe la confiance ». On ne peut qu'être d'accord avec cette deuxième conclusion, mais il me semble qu'elle est la conclusion d'une autre chaîne d'arguments. Ces arguments qui sont très présents dans la littérature philosophique sur la confiance, quand elle critique certaines simplifications de sciences de l'information et de la communication, reviennent généralement à montrer que trop de sécurité rendrait la confiance superflue. Il y a là une apparente contradiction. D'un côté, on a cette thèse (1) selon laquelle la sécurité suppose la confiance et ne peut en dispenser. De l'autre côté, une thèse (2) selon laquelle la sécurité rend superflue la confiance. La contradiction n'est qu'apparente si on distingue bien confiance pratique et confiance systémique. La thèse (1) porte sur le lien entre sécurité et confiance systémique. La thèse (2) sur le lien entre sécurité et confiance pratique.

La régression n'est pas un problème pour une philosophie de la confiance ; elle l'est pour une philosophie de la sécurité qui prétendrait remplacer entièrement la confiance par la sécurité. Bien au contraire, la confiance, au moins sous sa forme systémique, est ce qui met un terme à ce type de régression. Il y a des questions relatives aux fondements de nos pratiques que nous ne nous posons pas. Nous n'avons pas de raisons de soupçonner que ces fondements soient fragiles. Cependant, une philosophie de la confiance qui prétendrait s'appliquer aux

50 Weckert 2005, p. 109.

interactions en ligne en soutenant qu'elles présupposent une confiance systémique qui elle-même est acquise et est solide aurait quelque chose de très naïf, sinon de mensonger. Les doutes sur la sécurité des environnements sont justifiés au regard de l'expérience. Par exemple, l'utilisation d'un réseau privé virtuel (VPN) ne sécurise qu'au prix d'une exposition à un nouveau risque (celui de l'observation des communications par les personnes qui ont pleinement accès au serveur du VPN), de sorte qu'un certain discernement est recommandé dans le choix du prestataire. Le problème est donc que la confiance que les interactions en ligne supposent reste pour une large part à instituer. Elle n'est pas « acquise » au sens bizarre où ce terme tend à signifier ce qui est donné et va naturellement de soi. Elle est « acquise » (ou à acquérir) au sens où elle n'est pas naturelle.

Conclusion

Il y a une analogie entre cette discussion et ce que disait David Hume des circonstances de la justice dans son *Traité de la nature humaine* (III, II, 2). Selon Hume, la justice (entendue au sens social et économique, qui règle le travail, la propriété, l'échange) est une institution qui ne remplit de fonction et ne correspond à un besoin que lorsque deux sortes de circonstances sont réunies. Il faut que les agents ne soient ni extrêmement bienveillants, altruistes, ni extrêmement égoïstes. Il faut aussi que les biens soient relativement rares, c'est-à-dire qu'on ne se trouve ni dans la surabondance, ni dans la grande pénurie. Laissons de côté l'interrogation sur la pertinence économique de ces affirmations. Essayons de décrire dans des termes analogues le besoin d'une institution de la confiance : pour qu'elle soit à la fois possible et utile, cette institution suppose que les interactions ne soient pas complètement sécurisées, mais aussi qu'elles ne soient pas non plus, si l'on peut dire, complètement insécurisées ; les interactions ont lieu dans l'incertitude, mais celle-ci est cantonnée à la sphère de l'action et ne contamine pas toutes les conditions et tout l'environnement. Il faut également que les agents ne soient ni extrêmement crédules et téméraires, ni extrêmement soupçonneux et précautionneux.[51]

Les objectifs du marketing ou des acteurs institutionnels et commerciaux du cyberespace et ceux des sciences sociales et de la philosophie ne sont pas les mêmes. Là où ces dernières visent à comprendre ce qu'est la confiance et quelles en sont les conditions, les premiers visent à l'implémenter en rassurant les usagers, le

51 Ici encore, avec la considération de l'incidence des différences de degrés dans les paramètres de la confiance, qui selon leur ajustement peuvent la faire apparaître ou la dissiper, on voit que la suggestion selon laquelle le concept général de confiance a la structure d'une échelle des formes au sens de Collingwood n'est pas absurde.

problème pragmatique prenant le pas sur le problème théorique. Les uns et les autres ne portent pas le même regard sur des instruments de la confiance en ligne tels que les certifications du type *WebTrust* ou sur les agents conversationnels animés[52] ; pour la recherche universitaire, ce sont plutôt des palliatifs que des instruments de la confiance ; pour les acteurs du commerce en ligne, tous les moyens sont bons pour attirer le chaland et il n'y a pas lieu d'être très regardant sur les conditions de la confiance.

Ces palliatifs, remèdes ou incitations à une confiance assez fragile dans le commerce en ligne, posent à leur tour des problèmes de confiance, mais il ne s'agit pas nécessairement d'une régression du problème de la confiance pratique. En effet, lorsque l'on s'interroge sur la fiabilité d'une certification – par exemple, pour le *web*, un label « site de confiance » –, ou bien quand on devient attentif aux artifices par lesquels les services en ligne simulent une interaction naturelle, c'est plutôt un problème d'une nature différente qui surgit : un problème de confiance épistémique. Plus les interactions en ligne « émulent », comme disent les informaticiens, un environnement social naturel, plus on glisse de la question de la confiance pratique à celle de la confiance épistémique. On passe graduellement de l'évaluation des risques liés à son action à une autre espèce de questionnement : les doutes sur les apparences.

L'avenir de la confiance pratique dans le contexte des interactions en ligne dépend vraisemblablement, non seulement des solutions technologiques, mais aussi et surtout du développement de vertus épistémiques, appuyées sur une culture technique générale. L'art de ne pas s'en laisser conter reste une condition de la confiance cyberspatiale.

Bibliographie

Mike Ashley, Matthew Copeland, Joergen Grahn, & David A. Wheeler, *The GNU Privacy Handbook*, The Free Software Foundation, 1999, chap. 3 [https://www.gnupg.org/gph/en/manual/x334.html] et chap. 4 [https://www.gnupg.org/gph/en/manual/x547.html] (consulté le 3 août 2017)

Kenneth J. Arrow, *Gifts and Exchanges*, *Philosophy & Public Affairs*, 1, 4 (1972), p. 343-362.

52 Sur ces deux outils, voir Weckert 2005, p. 110.

Robert Axelrod, *Donnant donnant. Une théorie du comportement coopératif*, trad. Michèle Garène, Paris, Odile Jacob, 1992.

Şerif Bahtiyar & Mehmet Ufuk Çağlayan, « Security Similarity Based Trust in Cyberspace », *Knowledge-Based Systems*, 52 (2013), p. 290-301.

Robin G. Collingwood, *An Essay on Philosophical Method*, Oxford, Clarendon Press, 1933.

Francis Fukuyama, *La confiance et la puissance. Vertus sociales et prospérité économique*, trad. Pierre-Emmanuel. Dauzat, Paris, Plon, 1997.

Diego Gambetta (éd.), *Trust : Making and Breaking Cooperative Relations*, Oxford, Blackwell, 1988.

Russell Hardin, *Trust and Trustworthinesss*, New York, Russell Sage Foundation, 2002.

Richard Holton, « Deciding to Trust, Coming to Believe », *Australasian Journal of Philosophy*, 72, 1 (1994), p. 63-76.

Niklas Luhmann, « Confiance et familiarité. Problèmes et alternatives », trad. Louis Quéré, *Réseaux*, 108, 4 (2001), p. 15-35.

Helen Nissenbaum. « Securing Trust Online. Wisdom or Oxymoron? », *Boston University Law Review*, 81, 3 (2001), p. 635-664.

André Orléan, « La théorie économique de la confiance et ses limites », *Cahiers de Socio-Économie*, « La confiance en question », dir. R. Laufer et M. Orillard, Paris, L'Harmattan, 2000, p. 59-77.

Philip Pettit, « Trust, Reliance and the Internet », *Analyse und Kritik*, 26 (2004), p. 108-21.

Robert Putnam, *Bowling Alone. The Collapse and Revival of American Community*, New York, Simon & Schuster, 2000.

Louis Quéré, « La structure cognitive et normative de la confiance », *Réseaux*, 108, 4 (2001), p. 125-152.

Thomas W. Simpson, « Computing and the Search for Trust », in Richard H. R. Harper (éd.), *Trust, Computing and Society*, Cambridge, Cambridge University Press, 2014, p. 95-119.

Thomas W. Simpson, « e-Trust and reputation », *Ethics and Information Technology*, 13, 1 (2011), p. 29-38.

Matteo Turilli & Antonino Vaccaro & Mariarosaria Taddeo, « The Case of Online Trust », *Knowledge, Technology & Policy*, 23, 3-4 (2010), p. 333-345.

Vincent Véchambre, « Confiance numérique », *Communication*, 32, 2 (2013), mis en ligne le 11 avril 2014, consulté le 8 août 2017. DOI : 10.4000/communication.5074

Patricia Wallace, *The Psychology of the Internet*, Cambridge, Cambridge University Press, 1999.

John Weckert, « Trust in Cyberspace », in Robert J. Cavalier (éd.), *The Impact of the Internet on Our Moral Lives, Albany*, State University of New York Press, 2005, p. 95-117.

Bernard Williams, « Formal Structure and Social Reality » (1988), dans *Making Sense of Humanity*, Cambridge, Cambridge University Press, 1995.

Toshio Yamagishi & Masako Kikuchi & Motoko Kosugi, « Trust, Gullibility, and Social Intelligence », *Asian Journal of Social Psychology*, 2, 1 (1999), p. 145-161.