

PROTECTION DE LA VIE PRIVEE ET PATHOLOGIES DU WEB

Par Patrick Waelbroeck, maître de conférences en économie, Institut Mines-Télécom : Télécom ParisTech ; CNRS LTCI

Table des matières

| | |
|--|----|
| Protection de la vie privée et pathologies du web | 2 |
| Introduction..... | 3 |
| Pourquoi s'intéresser à l'économie des identités numériques ? A-t-elle un impact sur l'économie réelle ? | 3 |
| Pas d'interactions directes entre agents économiques dans le modèle de référence (équilibre général)..... | 4 |
| Organisation du chapitre..... | 5 |
| Les identités numériques : deux approches sociologiques..... | 5 |
| Le laboratoire d'identité..... | 6 |
| Gestion active d'identités multiples, manifestations et projection de soi | 7 |
| Boucle interactive entre la construction et la projection de soi..... | 8 |
| Valeur économique des données personnelles et de la gestion active des identités numériques..... | 8 |
| Economie de la protection de la vie privée | 9 |
| Marchés à plusieurs versants, valorisation des données et le gratuit | 11 |
| Asymétrie d'information et réputation..... | 11 |
| Prospectives :..... | 12 |
| Réalisation de valeurs données, propriété intellectuelle et marché des données | 13 |
| Big data: anticipation, prévention et responsabilisation | 14 |
| Réaction forte du privacy | 15 |
| Relation avec le pouvoir : l'empowerment | 16 |
| Bibliographie..... | 17 |

INTRODUCTION

Une enquête réalisée en octobre 2011 pour le Journal du Net révèle que l'internaute utilise en moyenne 14 identités numériques que ce soit pour les achats en ligne (4), la messagerie électronique (3) l'administration (2), les réseaux sociaux (2), la consultation de comptes bancaires en ligne, la messagerie instantanée, les forums. La majorité des internautes préfèrent des identifiants spécifiques à chaque usage. Les internautes gèrent de plus en plus de mots de passe pour sécuriser leurs comptes, plus d'un tiers utilise plus de 5 mots de passe.

Quels sont les éléments économiques à retenir pour analyser ce concept difficile à saisir qu'est la notion d'identité numérique ?

L'identité numérique serait la collection des traces que nous laissons derrière nous (ID, adresse IP, email, mots de passe, pseudo, données personnelles, administratives, bancaires, photos, avatars, tags, liens, publications) ainsi que le reflet de ces traces stockées et analysées par les moteurs de recherche et navigateurs web. L'identification (authentification), l'identité en ligne et les données à caractère personnel sont une condition d'accès aux différents services en ligne. Nous laissons des traces dès que nous lançons un navigateur internet. Ces traces participent à nos différentes identités que nous gérons de manière plus ou moins actives. Aux deux extrémités du spectre, on trouve d'une part l'anonymat complet qui permet la séparation des différentes pièces d'un puzzle de l'identité et l'impossibilité de reconstituer l'image complète. A l'autre extrémité, certains décident de révéler leur nom réel ainsi que de nombreux éléments de leur vie privée dans les réseaux socio-numérique tels que Facebook.

Pourquoi s'intéresser à l'économie des identités numériques ? A-t-elle un impact sur l'économie réelle ?

La première question à laquelle ce chapitre tente de répondre porte sur l'importance économique de la notion d'identité numérique. Ce point soulève deux autres questions : 1) Est-ce juste un phénomène important économiquement ? 2) Est-ce qui se passe en ligne a une influence sur l'économie réelle ?

La réponse à ces deux dernières questions est affirmative, et ce, pour trois raisons. Premièrement, la manière dont les gens gèrent ou non leurs identités en ligne influence de manière fondamentale le développement des communautés en ligne et la quantité ainsi que la fiabilité des informations qui y circulent. D'une part, les communautés telles que celles présentes sur Amazon.com ou ebay.com son très actives pour générer des informations sur les produits et sur les vendeurs de ces plateformes d'échanges. Ces informations conduisent souvent à des recommandations personnalisées qui ont une grande valeur économique dans la mesure où les internautes peuvent prendre des décisions informées à la fois en ligne ou dans le monde physique.

Deuxièmement, les événements qui se produisent en ligne peuvent avoir un impact important dans la vie réelle. Pour ne prendre qu'un exemple, les sites de rencontres en ligne permettent aux gens de prendre contact en ligne pour éventuellement se rencontrer réellement. Les effets entre le monde réel et celui en ligne peut se faire ressentir dans les deux sens, puisqu'une différence de point de vue lors d'une rencontre physique peut se répercuter en ligne ; inversement, des photos ou des messages échangés sur Facebook peuvent avoir des répercussions énormes sur la vie privée des gens ; de nombreux tabloïds illustrent quotidiennement ce point. Enfin, de plus en plus souvent, une rencontre physique est précédée par une « prise de contact virtuelle » (par exemple dans le cadre d'un recrutement ou d'un contact commercial). La connaissance de l'identité de l'internaute et sa « e-réputation » sont donc de plus en plus essentielles pour le recrutement professionnel. Le développement de sites de réseaux professionnels tel que LinkedIn témoigne de cette évolution.

Troisièmement, la manière dont les gens gèrent leurs identités peut avoir un impact sur des compagnies telles que Google ou Facebook qui exploitent les données personnelles. En effet, dans un environnement de confiance, les internautes seraient prêts à céder leurs informations privées en échange de services personnalisés et d'offre adaptées à leurs besoins et à leurs envies. Ces données personnelles représentent un enjeu financier considérable. Google est valorisé à 200 milliards de dollars au début de 2012, principalement grâce aux montants que les annonceurs sont prêt à payer pour accéder à des audiences ciblées grâce aux données personnelles. De même, Facebook a été valorisé à 100 milliards de dollars lors de son introduction en bourse grâce à la manière dont la compagnie commercialise les données personnelles de ses utilisateurs. Ces réseaux sociaux sont financés par la publicité et dépendent donc fondamentalement du comportement des internautes en matière de traces qu'ils laissent ou non lors de leur parcours de navigation. Dans le cas extrême où les internautes utilisent des pseudonymes ou naviguent de manière anonyme, de tels sites deviennent moins attractif auprès des annonceurs qui cherchent à cibler leur offres. Etant donnée le manque d'informations, les annonceurs ne peuvent plus aussi facilement pratiquer des prix en fonction de leur cible : la discrimination par les prix en est donc rendu difficile. Ainsi, la gestion active des identités remet en question la métaphore de l'oignon qui postule que l'identité d'une personne est constituée de couches successives laissées par les influences socio-cultures passées. Il suffirait alors pour mieux cibler les internautes de peler les différentes couches pour arriver au noyau. Hui et Png (2004) et Rochelandet (2010) proposent une revue de cette littérature sur l'économie de la protection de la vie privée. Ce chapitre remet en question cette vision simpliste du ciblage publicitaire.

Enfin, il y a d'autres enjeux autour du phénomène des big data. De plus en plus de compagnies développent des services basés sur l'exploitation massive de données personnelles collectées en ligne. Dans quelle mesure ces données sont-elles représentatives de la réalité des individus concernés ?

Pas d'interactions directes entre agents économiques dans le modèle de référence (équilibre général)

Malheureusement, la littérature économique sur la question des identités numériques est complètement inexistante, hormis quelques rares exceptions, et ce, pour deux raisons. Premièrement, les économistes sont mal équipés pour analyser la formation des goûts. Ainsi, la théorie économique postule l'existence d'une fonction d'utilité qui mesure le niveau de bien-être ou de satisfaction d'un individu, mais n'explique pas comment celle-ci est construite (construction de l'identité sociale, construction des goûts et des références culturelles). Deuxièmement, les identités numériques se construisent par l'interaction avec les autres, et là encore, la théorie n'est pas encore adaptée pour analyser les interactions directes entre agents économiques.

Ce double constat est partagé par Kirman (2004) qui souhaite que les sciences économiques s'intéressent plus à la question des interactions et donc également de manière implicite à celle de l'identité. En effet, pendant plus de trente ans, les sciences économiques ont évolué autour du paradigme de l'équilibre général. Dans ce modèle, tous les consommateurs et les entreprises sont des atomes sans identités et isolés qui n'interagissent qu'à travers un système de prix. Ils ont également tous la même information. Ainsi, la notion d'identité est à la fois inutile et n'est donc par abordée. D'autre part, les modèles issus de la théorie des jeux supposent que les gens connaissent parfaitement l'identité des autres individus avec qui ils interagissent. Ces interactions souvent locales, donc impliquant des notions d'identités, ont des conséquences en terme de comportement moutonniers, de réseaux, et de multiplicités d'équilibre en fonction des croyances de chacun. Le comportement agrégé d'une population devient alors différents de comportements individuels, ce qui remet en question la notion d'agent représentatif en économie.

Organisation du chapitre

Les utilisateurs révèlent leurs informations personnelles pour développer des relations étroites et significatives à travers des échanges réciproques. Ce désir de communiquer des renseignements personnels est contrebalancé par le risque de perdre l'anonymat. Ceci peut conduire à des risques en termes de protection de la vie privée, liés au fait que des inconnus bien informés peuvent s'immiscer dans la vie personnelle. Il existe une connexion entre le virtuel et le monde matériel que certains cherchent à atténuer. Ceci peut conduire à la déception et la stigmatisation (en raison de l'âge ou les handicaps physiques par exemple) et donc au mensonge. Le monde physique apparaît donc comme une contrainte sur la gestion des identités numériques. La réputation peut conduire des gens à jongler entre des identités multiples. Les gens élaborent leurs identités en tant que réponse à des motivations personnelles et à des demandes externes issues des contextes sociaux.

Ces nombreuses motivations derrière la gestion des identités numériques sont bien résumées par Kaplan (2010) et la Fing :

- la construction de soi
- la maîtrise de l'information (contrôle)
- la commodité (personnalisation du service, rappel des mots de passe, bookmarks,...)
- la valorisation de soi

Ce chapitre est organisé en deux grandes sections. Dans la section 2, je résume deux approches sociologiques de la notion d'identités numériques qui sont à la fois en opposition et complémentaires. Dans la section 4, j'analyse de manière critique différentes littératures économiques qui touchent directement ou indirectement à la notion d'identité. Je m'efforce de faire le lien avec les approches sociologiques. La section 4 conclut le chapitre.

LES IDENTITES NUMERIQUES : DEUX APPROCHES SOCIOLOGIQUES

L'identité numérique semble résulter d'un processus cognitif d'identification qui résulte du traitement d'un signal parfois imparfait comme un pseudo ou un avatar, ou encore une image sur un site de rencontre. Le processus d'identification en ligne est plus complexe que celui que l'on effectue tous les jours dans le monde physique où l'enveloppe corporelle donne un signal relativement précis, même si certaines personnes portent des masques qui cachent leur véritable nature. La manière d'écrire, le pseudo utilisé sont autant de signaux imparfaits qui trahissent une identité. Les différents degrés d'anonymat associés à la disparition de l'enveloppe physique créent un environnement médiaté où de nouvelles formes d'identités émergent.

La notion d'identité est étroitement liée à celle de la protection de la vie privée au sens où les gens tentent de contrôler l'information qui leur arrive et qu'ils émettent. Même si beaucoup d'informations ne sont pas a priori personnelles, elles peuvent le devenir a posteriori avec le data mining (exploitation de données pour chercher des régularités) et les big data (mise en relations de données publiques, privées, d'entreprise) qui permettent de recouper des informations. Par ailleurs, la notion d'identité est une notion objective, alors que la réputation est une notion subjective. Cependant, ces deux notions coexistent sur Internet. L'identité numérique devient alors le plus petit commun dénominateur des traces et de leurs interprétations.

Deux visions des identités numériques s'opposent dans la littérature principalement sociologique et du marketing. La première suppose que les gens utilisent les outils internet pour construire leur identité, ils essaient donc des avatars et des comportements différents de leur personnalité réelle. Ainsi, selon Turkle (2006), les gens utilisent des personnages différents de leur véritable identité. Ces identités sont fonction de la technologie, du contexte social (par exemple issu des normes sociales établies dans les communautés en ligne) et du contexte culturel. Les utilisateurs de

technologies de communications peuvent être en situation de co-présence liée à la disponibilité de plusieurs technologies de communication simultanément ou de la présence de plusieurs identités simultanément. La seconde vision suppose, au contraire, que les internautes se représentent en ligne comme ils sont dans la vie réelle, mais qu'ils gèrent de manière active les informations qu'ils divulguent au reste d'une ou plusieurs communautés. Dans les deux cas, les internautes peuvent avoir des identités multiples.

Le laboratoire d'identité

Turkle (2006) montre que les technologies de communication actuelles créent une nouvelle forme d'identité attachée à la technologie utilisée et au contexte social : téléphone portable, internet, messages instantanés. Les utilisateurs "cyclent" de plus en plus entre ces différents outils et créent ainsi un état continu de co-présence. Ainsi, une personne en voyage peut vérifier les nouvelles et la météo dans la ville où elle habite tout en téléphonant à quelqu'un très loin géographiquement : on peut être quelque part et autre part au même moment. Les mondes virtuels comme second life offrent la possibilité de se renouveler dans des contextes sociaux différents et changeants. Le choix des avatars dans de tels mondes virtuels est souvent lié à des enjeux de sexualité ou d'intimité, car l'anonymat permet de générer son propre outil de construction d'identité. Ainsi, les réseaux sur internet représentent un nouveau filtre à travers lequel on peut se découvrir et explorer ses limites. Les identités en ligne peuvent prendre de nombreuses formes : le choix d'un avatar sur World Of Warcraft, la participation à des forums de discussion ou des communautés de connaissances, la construction de playlists de musique, Ces outils permettent également de valider et d'être validé afin d'être reconnu et établi. Au final, l'utilisateur recherche à construire et à tester son identité.

La technologie possède un caractère absorbant et continu qui permet de construire et de détruire des identités multiples. Cependant, la technologie peut modifier les comportements et les identités si les utilisateurs participent à un concours de beauté par exemple, où les gens cherchent à savoir ce que les autres pensent avant de prendre leur décision. La recherche de la validation des autres peut créer un problème dans la construction de l'identité. De manière plus générale, les ordinateurs et les nouvelles technologies changent la manière dont nous réfléchissons, dont nous envisageons le droit de la protection de la vie privée, la manière d'interagir dans la vie réelle lorsque l'on est habitué à des rencontres virtuelles.

Selon Kaufmann (2004), l'identité est le résultat d'un travail où l'individu opte pour une image de lui-même. Kaufmann exploite la théorie de Hazel Markus (psychologie sociale) et de Sheldon Stryker (sociologie). Il développe la notion d'identité narrative pour proposer l'identité Immédiate, contextualisée et opératoire (ICO). L'identité se développerait comme une double hélice : la première reflète les normes, les habitudes, et la seconde est caractérisée par la subjectivité, les inventions et les ambitions.

Kafai (2010) illustre bien ce concept de laboratoire d'identités. Kafai et ses co-auteurs étudient la manière dont les jeunes adolescents construisent leurs avatars et leurs identités dans un jeu social virtuel appelé Whyville. Les utilisateurs peuvent acheter des parties de visage pour constituer une interface avec le reste de la communauté. Ils s'intéressent plus particulièrement aux ressources utilisées pour construire leurs visages ainsi qu'au choix des caractéristiques du visage et aux contraintes institutionnelles et sociales. Le statut social est établi par la "beauté" du visage. Les utilisateurs choisissent les caractéristiques de leurs avatars pour de nombreuses raisons qui diffèrent du simple désir d'avoir un avatar qui leur ressemble : beauté esthétique, envie de s'affilier à un groupe ou à une autre personne, caractéristiques que l'on peut avoir dans la vie réelle, par adhésion ou en signe de protestation contre une mode ou par goût du déguisement. Certains utilisateurs modifient fréquemment le visage de leur avatar ou créent même plusieurs avatars ("alts"). Les utilisateurs créent ces "alts" pour pouvoir collecter plus rapidement de la monnaie

virtuelle. Les contraintes sociales semblent très fortes dans Whyville : l'apparence de l'avatar est l'objet de nombreux fils de discussion, ce qui crée une norme sociale même si de nombreuses possibilités de personnalisation sont offertes par l'interface.

En conclusion, les normes sociales structurent les identités; les gens tentent de reproduire leurs normes culturelles, mais sont contraints par les communautés auxquelles ils appartiennent.

Gestion active d'identités multiples, manifestations et projection de soi

A l'opposé de la notion de laboratoire d'identités, la projection de soi (pour reprendre un terme utilisé par la Fing) considère que les internautes cherchent à communiquer, partager, s'exposer, obtenir des services personnalisés et gèrent ainsi de manière active leur véritable identité en divulguant plus ou moins d'informations aux autres. Par exemple, Merzeau (2009) considère que la pratique de ciblage et de mise en adéquation entre l'offre et la demande évolue d'une phase d'identification d'un type (stable et reproductible) vers un modèle de "token" (idiosyncratique et contextuel). Les données personnelles ne représentent plus des probabilités mais des attestations de présence. On passe ainsi du modèle de cible au modèle de crible où ne sont retenues que les informations validées par les utilisateurs.

Selon Cardon (2008), les internautes divulguent beaucoup d'informations personnelles et explorent des nouvelles communautés. Cette extériorisation traduit la tension entre être (sexe, âge, statut matrimonial, ...) et faire (oeuvres, projets, productions). L'expression de l'identité caractérise la tension entre la vie réelle (quotidienne, professionnelle, amicale) et une projection de soi.

Cardon pense que la visibilité peut prendre quatre formes : se cacher, se voir (site de rencontre); se voir caché (avatar); montrer caché (skyblog, Friendster); tout montrer, tout voir.

Il existe donc un fort lien réel-virtuel : sites de rencontre, réseaux physiques existants se renforcent en ligne, les plateformes post-it où des éléments de géo-localisations sont souvent affichés suppriment la frontière entre le réel et le virtuel. Cependant, cette approche ignore le problème des identités multiples et surtout le choix endogène de révéler de l'information aux autres membres de la communauté.

Coutant et Stenger (2010) soulignent également que les frontières hors-ligne et en-ligne sont très poreuses : les individus reconnaissent que leur profil en ligne représente une version améliorée de soi pas forcément en contradiction avec le monde réel. Les auteurs critiquent l'analyse de Cardon. En premier lieu, l'axe réel/projeté traduit mal l'aspect interactif du processus identitaire et le fait que l'individu ne contrôle pas toujours ce processus. L'axe être/faire est en contradiction avec la littérature sur les techniques de soi et l'écriture de soi (rite de passage, actes extrêmes)

Les différentes manifestations de soi sont bien illustrées par le comportement des internautes sur les sites de rencontre. Ainsi, Chester et Bretherton (2007) montrent que sur ces sites les caractéristiques qui sont désirables en ligne soient similaires à celles qui le sont dans le monde physique. Les gens sont prêts à se présenter sous leur facette réelle. Les autres stratégies sont considérées comme malhonnêtes, ce qui n'empêche pas la plupart des membres de ces communautés de déguiser leur identité. Selon Toma C. L., Hancock, J. T., Ellison N. B. (2008), 86% des membres de sites de rencontre en ligne croient que les autres utilisateurs ne communiquent pas leur véritable apparence physique. Enfin, il existe des contraintes technologiques liées au fait que certains sites qui regroupent les utilisateurs par catégorie d'âge. De nombreux utilisateurs essaient de contourner ces contraintes en se présentant plus jeunes qu'ils ne le sont vraiment. Au final, l'évaluation de l'identité d'autrui devient un processus complexe lié au décryptage de signaux parfois contradictoires. Dès lors, certains indices tels que les fautes d'orthographe et les références sexuelles ou non sont minutieusement analysés.

La présentation de soi traduit la tension entre l'envie de se présenter sous son meilleur jour et le besoin de révéler sa véritable identité afin d'attirer l'âme soeur. Le caractère asynchrone des sites de rencontre permet à un utilisateur de se présenter en fonction de ce que les autres attendent. Toutefois, l'anticipation des interactions futures réduit la tromperie. Au final, les utilisateurs sont relativement honnêtes et réalistes sur les sites de rencontre.

Un paradoxe apparaît alors. Les gens sont prêts à divulguer plus d'informations sur eux-mêmes lorsqu'ils sont en confiance et qu'ils sont dans des relations d'intimité et d'amitié, mais c'est alors qu'il devient plus difficile de partager les sentiments intimes que l'on pourrait partager avec des inconnus.

En conclusion, le monde des rencontres virtuelles ne correspondrait pas à la métaphore du laboratoire de l'identité décrit dans la section précédente. Au fond, le monde en ligne ne serait pas un monde si différent, mais simplement un nouvel outil de communication qui n'est pas si déconnecté de la vie réelle des gens. L'utilisateur est alors incité à brouiller son identité : anonymat, pseudo, rétention d'informations, déclarations mensongères. L'internaute cultive son identité numérique. Sur ce point, la Fing préconise un droit à l'hétéronymat permettant à chacun de s'adjoindre des identités alternatives de longue durée.

Boucle interactive entre la construction et la projection de soi

Les deux approchés décrites dans les sections précédentes ne sont pas nécessairement contradictoires, car les internautes passent de l'une à l'autre dans un processus dynamique impliquant la construction et la projection de soi. Ainsi, selon Stenger et Coutant (2010), les individus construisent leurs identités à travers un processus à la fois individuel et collectif.

Cette boucle interactive est clairement établie sur les mondes virtuels comme Second life et sur les réseaux sociaux comme Facebook. Ainsi, McLeod et Lesehd (2011) étudient la manière dont les gens gèrent leurs identités dans second life : de l'anonymat à la divulgation complète de l'information sur leur vie réelle. Les membres de Second Life choisissent donc leur degré d'exposition. L'anonymat peut paradoxalement conduire à plus d'intimité qu'entre des personnes relativement proches. L'anonymat permet de réduire les aspects négatifs ou tabous de soi-même. Elle peut révéler le vrai soi. Par ailleurs, les gens définissent leur identité en fonction du groupe social auquel ils appartiennent. Cependant, la participation active aux groupes sociaux peut réduire l'anonymat, car il devient plus facile d'identifier une personne. Le monde physique peut être une contrainte sur ce que l'on est prêt à révéler sur soi. Les personnes se connectent à Second Life pour rencontrer de nouvelles personnes et pour exprimer leurs idées et opinions tout en gérant leur réputation. Ainsi, les internautes cherchent à s'exprimer; ils gèrent leur identité à travers les groupes auxquels ils appartiennent; il existe une tension entre la réputation et la confiance et entre l'anonymat et le désir de développer des relations sociales et de construire son identité par rapport aux autres; le développement du monde virtuel dépend de la tolérance et de l'expérimentation, même si les gens se présentent souvent selon leurs normes socio-culturelles.

VALEUR ECONOMIQUE DES DONNEES PERSONNELLES ET DE LA GESTION ACTIVE DES IDENTITES NUMERIQUES

Si les approches sociologiques ont poussé relativement loin l'analyse des identités numériques, la littérature économique n'aborde pas directement cette question. Je résume une partie de la littérature sur l'économie de la protection de la vie privée dans la section 3.1. Cette littérature aborde la notion de l'identité sous l'angle de la discrimination par les prix et de la customisation : les entreprises cherchent à collecter les informations personnelles liées à l'identité d'une personne pour trouver le meilleur prix auxquels ils peuvent vendre leur produit et service en fonction de la disponibilité à payer des clients potentiels. Ceux-ci bénéficient alors de services personnalisés (ce que certains comme Daniel Kaplan ou la Fing appellent la customisation). L'économie des

média et de l'internet justifie aussi la valeur économique liée à la collecte d'informations personnelles lorsque ces informations s'échangent sur un marché à plusieurs versants où un intermédiaire collecte des informations personnelles et le commercialisent à des annonceurs, comme Google ou Facebook. Cette littérature que je résume brièvement dans la section 3.2 explique pourquoi les internautes participent à ce type de marché : ils peuvent consommer gratuitement le service de l'intermédiaire qui se finance auprès des annonceurs. La section 3.3 aborde la littérature sur les asymétries d'information qui analyse la notion d'identité en ligne à travers la notion de réputation. Le développement d'une bonne réputation explique pourquoi il est important de gérer ses informations personnelles sur les plateformes de ventes telles qu'eBay ou Amazon Marketplace : il existe une prime à la réputation qui permet de facturer un produit ou service plus cher que celui de son concurrent.

Economie de la protection de la vie privée

Justification de la protection de la vie privée

Cette section reprend la présentation de Rochelandet (2010) pour justifier la protection de la vie privée : l'autonomie individuelle d'une part et l'intimité, la paix sociale et les réseaux relationnels d'autre part. De nombreuses études empiriques montrent que

- les individus surveillés en prison peuvent perdre l'autonomie de penser, d'agir de manière singulière, indépendante
- la vie privée permet de construire son identité
- la vie privée permet de développer la confiance en soi
- le partage d'informations personnelles permet de construire son cercle d'amis et son réseau social
- la vie privée permet d'établir un jugement moral indépendant.

La notion d'autonomie n'est cependant pas abordée en économie, sauf par Hirschleifer (1980). La littérature économique se concentre sur la notion de secret développée par l'école de Chicago, qui considère que les gens cherchent à dissimuler des informations sur eux, ce qui peut mener à des inefficiences dans la mesure où les entreprises doivent engager plus de moyen pour obtenir des informations supplémentaires. La protection de la vie privée ne serait alors qu'une manière de dissimuler un ou des vices cachés. Ceci engendre des coûts de collecte d'informations qui se répercuteraient sur les prix. Le manque d'informations peut également conduire à des décisions non optimales. Cet argument est clairement illustré par Turner et Vargese (2010) qui étudient le marché du crédit dans les pays en développement. Dans de tel marchés où il existe des asymétries d'informations entre les participants au marché, Akerlof a montré que lorsque seule la qualité moyenne peut être évaluée, les bons dossiers risquent d'être éliminés et il ne reste plus que les « lemons » (pour reprendre l'expression désignant les voitures d'occasion de très mauvaise qualité). Le mécanisme de prix ne suffit pas à lui seul pour empêcher le rejet des bons dossiers. Ainsi, lorsque les banques n'utilisent qu'une partie des informations sur un emprunteur potentiel (typiquement les informations négatives liées au nombre de fois où le candidat au prêt a payé ses mensualités en retard dans le passé), elles refusent plus souvent des bons dossiers que lorsqu'elles utilisent toutes les informations sur le candidat potentiel (y compris les informations positives comme le nombre de projets entrepreneuriaux réussis par le passé). Le manque d'informations peut donc conduire à des inefficiences qui pourraient être supprimées par plus d'informations.

Protection de la vie privée et externalités

Cependant, l'approche de l'école de Chicago ignore deux problèmes fondamentaux. D'une part, trop d'informations peuvent détruire certains marchés comme celui des assurances, car les compagnies d'assurance n'assureraient que les bons profils, qui n'ont pas forcément besoin d'assurance. D'autre part, l'argument de l'école de Chicago ignore l'existence d'externalités positives et négatives liées à l'exploitation des données personnelles par des entreprises tierces. Les externalités positives (bénéfiques pour les entreprises) peuvent conduire à une sous-production d'informations personnelles. Parmi les exemples d'externalités positives, on peut citer :

- une meilleure prévision de tendance de marché, de l'évolution des préférences individuelles;
- un meilleur ciblage des offres, le développement de produit de niches;
- de nouveaux business models tels que le big data et la commercialisation de données personnelles

Les externalités négatives (pour le consommateur) conduisent à une collecte d'informations personnelles trop abondante. Parmi les exemples d'externalités négatives, citons :

- le vol d'identité
- d'autres formes d'exploitation d'informations personnelles par des tiers à des fins douteuses comme le spam ou le marketing direct
- la perte de données personnelles, comme les numéros de cartes de crédits, liés à un manque de sécurisation des données stockées

Le risque lié à la perte d'anonymat est donc lié aux risques d'être exposé à des externalités négatives par un processus de ré-identification particulièrement mesurable avec le phénomène de big data qui permet de mettre en relations plusieurs bases de données a priori sans connexions.

Le marché peut-il suffire pour que les entreprises protègent elles-mêmes les données personnelles et la vie privée de leurs clients ? Acquisti (2010) argumente que l'on ne peut pas laisser le marché face à lui-même si les consommateurs ne sont pas parfaitement rationnels. Il développe des théories basées sur l'économie comportementale où la rationalité des consommateurs n'est pas supposée a priori.

Protection de la vie privée et discrimination par les prix

Lorsque les entreprises ont plus d'informations sur leurs clients, elles peuvent pratiquer la discrimination par les prix, à savoir tarifier un même produit ou service à des prix nets différents. Le prix net s'entend net de frais de livraisons ou du coût de production. Ainsi, pour les produits numériques, la forme de discrimination la plus répandue est celle où les entreprises peuvent identifier plusieurs groupes de consommateurs et leur proposent des versions différentes d'un même produit ou service. Par exemple, un fabricant de logiciel propose différentes versions de son produit avec des fonctionnalités différentes : une version professionnelle complète et une version plus simple (ou pour étudiants) où certaines fonctionnalités sont absentes. Chaque version permet d'étendre son marché à de nouveaux consommateurs. Cependant, chaque nouvelle version fait également basculer certains consommateurs d'un segment plus rentable vers un segment moins rentable pour l'entreprise. Le nombre de versions mises sur le marché dépend ainsi de l'équilibre entre ces deux forces. La collecte d'informations permet donc de customiser les offres aux clients ciblés, à des coûts souvent très faibles pour les entreprises fournissant des biens numériques facilement modulables. Quel est l'effet de cette pratique sur le profit des

entreprises et la satisfaction des consommateurs ? De manière surprenante, le surplus total ne diminue pas nécessairement lorsqu'une entreprise en monopole pratique la discrimination par les prix, en particulier lorsque les quantités consommées sur le marché augmentent. Ainsi, une augmentation de la collecte d'informations n'est pas nécessairement mauvaise pour les consommateurs qui se voient proposer des offres ciblées, même s'il existe un risque de clientélisme (un vendeur propose des prix à ses clients en fonction de critères non objectifs).

Marchés à plusieurs versants, valorisation des données et le gratuit

De nombreuses plateformes de réseaux sociaux, d'actualités et de recherche d'informations peuvent être caractérisées par ce que la littérature économique a appelé des marchés à deux versants. Ceux-ci sont caractérisés par des externalités croisées entre différents groupes d'agents. Par exemple, sur un site de recherche d'information tels que Google.com, les internautes accèdent à des contenus gratuitement, et le site se finance par publicité. Le site met donc en relation des consommateurs potentiels et des annonceurs. La satisfaction d'un consommateur potentiel lorsqu'il effectue une recherche sur un site dépend du nombre de publicités ciblées qu'il trouvera pertinentes par rapport à ses recherches. De même, pour un annonceur, le bénéfice d'une publicité dépend du nombre d'utilisateurs du moteur de recherche. Il y a donc une externalité croisée positive entre les internautes et les annonceurs. Ceci est également vrai pour tout site d'actualités (journaux en ligne par exemple) qui se finance par la publicité. La dynamique des marchés à deux versants où internautes et publicitaires/fournisseurs de contenu se rencontrent implique que les plateformes qui parviennent à gagner un petit avantage comparatif initial peuvent se retrouver dans une boucle positive alors que les autres subissent une spirale négative. Ainsi, les marchés des moteurs de recherche et des sites d'actualités où l'internaute est à même de découvrir de nouveaux produits risquent d'être très concentrés avec l'émergence d'un intermédiaire dominant. La théorie des marchés à plusieurs versants montre que l'intermédiaire a intérêt à subventionner le côté du marché qui est fondamental pour le développement de la plateforme. Ainsi les services de Google et de Facebook sont gratuits pour les utilisateurs, alors que les annonceurs paient pour accéder à leurs audiences. Cette pratique explique également pourquoi les utilisateurs ont du mal à donner une valeur à leurs données personnelles, puisqu'ils ne paient pas pour le service. Prises dans leur ensemble, ces données personnelles sont valorisées à des centaines de milliards de dollars par les investisseurs sur les marchés financiers.

Que ce passerait-il si le coût de la collecte d'informations augmentait. Tucker (2012) apporte un élément de réponse. Dans une étude comparative entre le choix des annonceurs aux Etats-Unis et en Europe, elle montre avec son co-auteur que dans un environnement plus régulé comme l'Europe, les annonceurs préfèrent financer des sites spécialisés plutôt que des sites généralistes.

Asymétrie d'information et réputation

La théorie économique de la réputation apporte également un éclairage sur les motivations des gens dans leur manière de gérer de manière active leurs identités. La réputation est définie comme un « goodwill », c'est un stock qui augmente avec les expériences positives. Cette "prime à la réputation" a une valeur économique, qui est cependant difficile à mesurer à partir de données collectées sur Internet. Premièrement, les internautes ont une probabilité plus élevée de donner une bonne évaluation si le rapport qualité/prix du produit ou du service acheté est bon. Dans ce cas, la note donnée dépend du prix : un acheteur ayant payé un prix jugé trop élevé donnera un moins bonne note qu'un acheteur ayant payé moins cher pour le même produit. Ainsi, il est fréquent de voir une transaction avec un prix élevé évaluée par une note faible voire négative. Si l'on cherche à utiliser cette note comme mesure de réputation, on risque d'estimer une prime à la réputation négative ! Deuxièmement, les vendeurs qui effectuent un gros volume de transactions

sont également ceux qui reçoivent le plus d'évaluations. Mais ce sont également eux qui bénéficient d'économies d'échelles et peuvent donc plus facilement proposer des rabais et des promotions face à leurs concurrents de plus petite taille. Si l'on utilise la variable "nombre d'évaluations" pour évaluer la réputation d'un vendeur, on risque à nouveau d'estimer une prime à la réputation négative. Troisièmement, la réputation peut être manipulée. Il existe même des marchés de la réputation où des vendeurs paient pour recevoir de bonnes évaluations (voir par exemple Dini et Spagnolo, 2007 et Lumeau, Masclet, Pénard, 2010). Il s'agit alors pour l'acheteur d'évaluer la réputation comme un signal imparfait de la véritable fiabilité d'un vendeur. Cabral (2012) propose un tour d'horizon succinct des modélisations économiques de la réputation. Quatrièmement, toutes les personnes qui effectuent une transaction sur Internet ne laissent pas nécessairement une évaluation. En effet, l'évaluation laissée contribue à la connaissance globale de la communauté en ligne (celle d'eBay ou d'Amazon par exemple) sans que la personne qui a rédigé cette évaluation ne soit directement rémunérée pour sa contribution. Ceci engendre une contribution sous optimale de chacun, un problème connu sous le terme de free-riding en économie. De plus, les internautes laissent des évaluations pour différentes raisons (réparer une injustice, développer une communauté, anticiper une évaluation réciproque, se faire plaisir, ...). S'il s'avère que les caractéristiques des personnes qui contribuent sont différentes de celles qui ne contribuent pas, on se retrouve devant un problème de biais de sélection liée à la nature non aléatoire de l'échantillon que l'on utilise pour estimer la prime à la réputation.

Malgré ces problèmes, un grand nombre d'auteurs étudient la plateforme eBay pour des raisons de disponibilité des données. Les études montrent qu'il existe une faible prime à la réputation. La réputation a cependant un effet sur la probabilité d'effectuer une transaction (pour un vendeur). Par exemple, Cabral et Hortacsu (2010) qu'une augmentation d'un pourcent du nombre d'évaluations négatives conduit à une baisse de 7,5% du prix de vente. Ils montrent également que lorsqu'un vendeur reçoit sa première évaluation négative, son volume de vente baisse de 13%. Un vendeur qui reçoit plusieurs évaluations négatives a plus de chance de quitter la plateforme de vente. De manière similaire, Bounie et al. (2012) montrent qu'il existe une prime à la réputation sur Amazon Marketplace. Ainsi, la gestion active de son profil de vendeur a une valeur économique et explique pourquoi les gens cherchent à se présenter sur leur meilleur angle.

PROSPECTIVES :

Les problèmes liés à la protection de la vie privée et à celle des données personnelles risquent de s'amplifier (même si les nombreux contentieux sur Facebook ou Google font déjà la une quotidienne des sites web technophiles) :

- reconnaissance faciale et orale,
- données biométriques
- Internet des objets
- réalité augmentée
- partage de données et Web collaboratif
- RFID
- paiements sans contact
- services de géo-localisation

Réalisation de valeurs données, propriété intellectuelle et marché des données

Frontière entre propriété intellectuelle et protection des données personnelles

Un internaute qui met en ligne des données, comme par exemple un commentaire laissé sur un forum Internet, est théoriquement protégé par le droit d'auteur/copyright. Cependant, l'utilisateur renonce à ce droit par la signature d'un contrat électronique qui transfère la propriété intellectuelle au propriétaire du site. La question de savoir ce qui se passerait si l'utilisateur pouvait récupérer son droit d'auteur soulève l'importante question du marché des données personnelles. L'utilisateur devient alors offreur/producteur de données personnelles, et les annonceurs, demandeurs de ces données. Même si les données personnelles semblent très hétérogènes, un marché pourrait se créer où s'établirait un prix de transaction payé par l'annonceur directement à l'utilisateur et non plus en passant par un intermédiaire comme Google et Facebook. Certains considèrent qu'il s'agit d'une "fausse bonne idée" (voir encadré 1).

Encadré 1 (D. Kaplan):

LA PROPRIÉTÉ PERSONNELLE DES DONNÉES, UNE PISTE ?

(PLUTÔT POUR UN ENCADRE...)

"Dans une économie de marché, tout incite les entreprises à collecter à notre propos des données plus nombreuses et plus précises. Dans une bureaucratie d'État, tout incite à étendre sans limite la surveillance", écrit l'historien des médias Siva Vaidhyanathan¹. La technique le permet, les acteurs le désirent : tout semble se combiner pour rendre de plus en plus illusoire l'idée que nous pourrions disposer d'un espace d'intimité, de secret, dans une société informatisée. Les défenses existantes sont submergées de toutes parts. Et les individus eux-mêmes semblent prendre peu d'intérêt à leur propre défense.

D'où l'idée récurrente de confier aux individus la propriété de leurs propres données et par là-même, de leur transférer certaines responsabilités et de leur conférer certains droits - dont, sans doute, celui d'en tirer une rémunération. À l'instar de la Cnil, les sénateurs Detraigne et Escoffier² considèrent ce droit de propriété sur les données personnelles comme une "fausse bonne idée" : "Ce régime juridique est inadapté aux enjeux d'Internet. Comment y faire valoir ses droits de propriété quand on a soi-même diffusé une donnée personnelle ? Quant aux informations relatives à sa vie publique, il ne peut être question d'un droit de propriété à moins de remettre en cause fondamentalement la liberté d'expression. Enfin, lorsque le rapport de force entre des contractants est inégal, quelle valeur donner au consentement à contracter ?"

De leur côté, plusieurs spécialistes américains de la vie privée ont tenté, sur la base des prix du marché, d'évaluer les sommes qu'un internaute pourrait tirer de ses propres données. Résultat : 15 à 25 \$ par an, une compensation bien maigre par rapport aux inconvénients prévisibles. D'autant que les conséquences les plus graves d'un usage abusif de ces données ne sont pas économiques : elles touchent à la dignité d'une personne, au degré de confiance ambiante qui peut exister dans une société. Les bénéfices et surtout les coûts de la circulation de données personnelles se calculent pour l'essentiel hors marché. En outre, comme les détenteurs de droits d'auteurs l'ont constaté, la propriété ne garantit en rien la protection des informations dans un monde numérique.

Enfin, la propriété des données personnelles ne résoudrait au mieux qu'une petite partie des problèmes relatifs à l'identité et la vie privée dans une société numérique. Le risque pour les individus ne réside pas dans les données, mais dans les décisions que prennent des humains ou des ordinateurs sur la base de ces données - généralement en les croisant avec d'autres données qui ne sont pas, ou pas directement, personnelles. "La propriété est une fausse piste. La question de la vie privée, c'est celle des droits que vous avez sur toutes les informations qui vous affectent"³, affirme ainsi Robin Wilton de la Kantara Initiative.

¹ "Naked in the Nonopticon": Surveillance and marketing combine to strip away our privacy", *The Chronicle of Higher Education*, février 2008 : <http://chronicle.com/article/Naked-in-the-Nonopticon/2300701>

² "La vie privée à l'heure des mémoires numériques", rapport de la commission des lois du Sénat, mai 2009 : <http://www.senat.fr/rap/08-44/rap08-441.html>

³ "What is the value of your personal identity?", 2010 : <http://www.futureidentity.eu/documents/RW-Maribor.pdf>

Enjeu : éthique et régulation

Avant de discuter la manière dont un équilibre et une structure de marché pourraient se construire sur un tel marché, il est important d'aborder les aspects éthiques de la question.

Peut-on vendre n'importe laquelle de ses données personnelles, quelle que soit sa nature ?

Enjeu : confiance, certification, labels

Le développement du marché des données personnelles ne se fera pas sans garantir la confiance des internautes envers les entreprises et les institutions qui les gouvernent. Face aux bafouements quotidiens des règles élémentaires de protection de la vie privée de Facebook, de nombreux

internauts ont déserté les réseaux socio-numériques ou sont à la recherche d'alternatives plus soucieuses de la protection de leurs données personnelles. La confiance mutuelle est la base des interactions sociales. Pour rétablir et garantir cette confiance, il sera sans doute nécessaire de passer par une étape de certification et de labellisation assurée par un tiers de confiance. Des pistes basées sur des licences d'utilisations de données personnelles ayant différents niveaux d'intimité ou d'extimité devront être explorées. Certains spécialistes comme D. Kaplan pensent qu'une auto-régulation basée sur un mécanisme de confiance pair à pair est suffisante.

Point de rupture/bifurcation: le contrôle du portefeuille d'identités numériques

Au mois d'octobre 2012, la CNIL a annoncé qu'elle allait demander à Google de revenir sur le changement de sa politique de protection de la vie privée mise en place en mars 2012. Ce changement permettait à Google de rassembler toutes les identités numériques liées à un utilisateur de ses différents services (Youtube, gmail, documents, picasa, ...) sous une identification unique. Au-delà des problèmes de protection de la vie privée, ce changement traduit, pour une entreprise du web, les enjeux stratégiques liés au contrôle du portefeuille d'identités numériques des internautes afin de devenir un portail d'accès au web. Les enjeux financiers pour les annonceurs et pour Google sont considérables.

Ce contrôle des identités numériques doit-il être laissé à des entreprises privées, ou ressort-il des autorités gouvernementales, comme pour la carte d'identité nationale ?

Big data: anticipation, prévention et responsabilisation

Les big data adoptent une approche macro-économique d'un problème en cherchant des structures statistiques basées sur des données massives et globales. Les big data permettront aux entreprises et aux gouvernements de mieux analyser les tendances de marché et les changements d'opinions.

Enjeu : gouvernementalité algorithmique

Dans un article du Wall Street Journal ("Meet the new boss: Big data", 20 septembre 2012), l'auteur montre comment les logiciels utilisant les données massives peuvent fournir des recommandations à un employeur d'un centre d'appel non pas basées sur l'expérience passée pour embaucher de nouvelles recrues, mais sur des tests de personnalités (attitude envers l'alcool, distance du domicile au travail, personnalité inquisitive, ...) qui garniraient qu'un nouvel employé reste au moins six mois.

Antoinette Rouvroy ("Face à la gouvernementalité algorithmique, repenser le sujet de droit comme puissance") soulève les risques liés à une utilisation systématique du data-mining et d'algorithmes statistiques utilisant des données massives. La création de catégories et la rationalité algorithmique éloignent les personnes de la véritable communication et remettent en question les notions de critique, de justice et de protection des minorités.

Comment réagir lorsqu'un algorithme vous a assigné une catégorie qui ne reflète pas forcément la complexité de votre personnalité ?

Enjeu : Big data et small data

Selon Alan Mitchell, l'un des principaux enjeux de productivité portera sur la logistique de l'information ou comment garantir que l'information pertinente arrive au bon endroit au bon moment. Il s'agit d'une approche micro-économique. L'efficacité des big data dépendra de celle de la combinaison entre les approches micro- et macro-économique d'un problème.

Enjeu : partage d'informations

Dans un article intitulé "Moving Beyond Untagging: Photo Privacy in a Tagged World", les auteurs montrent la difficulté de contrôler des informations personnelles sur une photo référencée par des moteurs de recherche. Lorsque plusieurs tags sont associés à une photo, la question de savoir qui possède le contrôle final sur les tags (règle de décision) est fondamentale pour que des internautes continuent de partager des photos sur des sites tels que Flickr ou Facebook (en particulier lorsqu'il s'agit de photos compromettantes).

Point de rupture/bifurcation : Contrôle des standards de communication

Pour le moment, le standard de communication de données non structurées utilisé par le big data est un standard ouvert appelé Hadoop et utilisé par des compagnies telles qu'Amazon, Yahoo ou Facebook. Le contrôle du protocole de communication déterminera quel standard s'imposera. Les éco-systèmes qui se formeront autour des différents protocoles de communication entreront en concurrence et définiront la structure de marché des big data.

Limites des big data

Biais de sélection et économie de la contribution

Ce chapitre soulève d'importantes limites à l'utilisation des données issues de collectes d'identités numériques et aux "big data" de manière générale. Premièrement, les individus peuvent se présenter sous de fausses identités ou des identités partielles ou "espérées" ou s'identifier au minimum pour rester le plus anonyme possible. Deuxièmement, ils peuvent contribuer beaucoup ou peu à des projets collectifs (open source, wikipedia, forums et autre communautés de savoir). Dès lors, certaines caractéristiques et opinions seront sur-représentées. Les solutions à ces problèmes de données se trouveront dans la modélisation explicite du choix des identités et de l'étendue de la contribution des individus.

Anticipations rationnelles

Les modèles de prévision et d'anticipation du comportement des individus ne sont opérationnels que si ces derniers sont relativement passifs. Au contraire, lorsque les individus anticipent eux-mêmes les règles qu'on leur applique, ils peuvent manipuler leur données pour rendre l'algorithme de prévision inefficace. Cette critique des modèles de prévision basée sur les anticipations rationnelles est due à Robert Lucas dans un contexte macro-économique.

Modélisation des comportements et des manières de penser

La compréhension des comportements humains complexes ne pourra reposer uniquement sur des approches purement statistiques. Il faudra impérativement développer et appliquer des modèles de raisonnement aux algorithmes existants en utilisant à la fois la théorie économique et l'intelligence artificielle.

Réaction forte du privacy

Guerre de protection

Dans un contexte de défiance envers les réseaux socio-numériques et les entreprises qui les ciblent, les internautes auront recours à des outils de plus en plus sophistiqués pour garantir le respect de leurs données personnelles. L'utilisation de VPN et d'outils cryptographiques déclenchera une réaction aussi forte des annonceurs qui se tourneront vers des outils de ciblage de plus en plus intrusifs, menant à une guerre de protection. Dans un scénario extrême, on peut

envisager que les internautes se déconnectent et n'échangent plus leurs données qu'à travers des disques durs portables.

Enjeu : innovations et équilibres juridico-économiques

Cette guerre de protection est néfaste pour l'innovation. En effet, les Etats-Unis qui sont plus laxistes sur l'utilisation des données personnelles ont connu une vague d'innovations marketing menées par Google, Facebook et Twitter.

Limites de l'anonymisation

Il existe un arbitrage fondamental entre l'utilité d'une donnée et son degré d'anonymat. Cet arbitrage est amplifié par le développement des big data et le web collaboratif qui permet de mettre en relation des bases de données auparavant séparées.

Relation avec le pouvoir : l'empowerment

L'individu comme point d'entrée

L'open data va permettre au citoyen de mieux évaluer les actions politiques. Il y gagnera en autonomie et le processus démocratique en sera renforcé.

Transparence réciproque

Combiné avec le big data, cette transparence réciproque devrait améliorer le dialogue entre gouvernant et gouverné et rendre la démocratie plus efficace.

Limites de l'empowerment : surveillance et sousveillance

L'historique des données de navigation et des traces volontairement ou involontairement laissées sur Internet facilitent la surveillance par une autorité centrale qui pourrait ainsi renforcer son pouvoir politique. Ce phénomène est déjà largement étudié sous le terme "Big Brother". Un autre phénomène plus insidieux appelé par Casteels "Little Sisters" renvoie à un phénomène de sousveillance où les actions, les traces et les posts d'un internaute sont sans cesse suivis par une multitude de petites soeurs qui rendent la frontière entre le monde physique et le monde en ligne de plus en plus poreuse.

Cette double évolution aura des conséquences sur l'évolution de l'autonomie des individus et sur la manière dont ils s'éloignent ou non d'une norme sociale imposée par un consensus.

BIBLIOGRAPHIE

- Acquisti A., 2010, "The Economics and Behavioral Economics of Privacy: A Note" , Proceedings of the Third International Conference on Ethics and Policy of Biometrics and International Data Sharing, Hong Kong, Springer LNCS Vol. 60059
- Akerlof G.A., Kranton R.E., 2000, "Economics and Identity", Quarterly Journal of Economics, Vol. 115 (3), p. 715-753
- Cabral L., 2012, "Reputation on the Internet", in: Oxford Handbook of the Digital Economy (Peitz et Waldfogel Eds.), Oxford University Press
- Cabral L., Hortascu A., 2010, "The Dynamics of Seller Reputation: Theory and Evidence from eBay", Journal of Industrial Economics, Vol. 58, p. 54-78
- Cardon D., 2008, "Le design de la visibilité : Un essai de cartographie du web 2.0", Réseaux, 152, p. 93-137
- Castronova E., 2003, "Theory of the Avatar", CES Ifo Working paper 863
- Chester, A., Bretherton, D., 2007, "Impression management and identity online", In: Oxford handbook of internet psychology (Joinson, McKenna, Postmes, Reips Eds.), pp. 223-236), Oxford: Oxford University Press
- Davis J.B., 2010, Individuals and Identity in Economics, Cambridge University Press
- Dini F., Spagnolo G., 2007, "Buying Reputation on eBay", Quaderno Consip VIII
- Higgins E.T., 1987, "Self-discrepancy: A Theory Relating Self and Affect", Psychological Review, 94, 319-340.
- Hirschleifer J., 1980, "Privacy: its Origin, Function, and Future", Journal of Legal Studies, Vol. 9 (4), p. 649-664
- Hui K.L., Png I.P.L., 2006, "The economics of privacy", Economics and Information Systems, Hendershott T. (ed.) Handbooks in Information Systems, vol. 1, chapitre 9, Elsevier
- Kafai Y.B., 2010, "World of Whyville: An Introduction to Tween Virtual Life", Games and Culture 2010; 5; 3
- Kaplan D., 2010, "Informatique, libertés, identités", FYP Editions, coll. La fabrique des possibles
- Kaufmann J-C., 2004, L'invention de soi. Une théorie de l'identité, Armand Colin
- Kirman A., 2004, "The Structure of Economic Interaction: Individual and Collective Rationality", in: Cognitive Economics: An Interdisciplinary Approach, Bourguine et Nadal (Eds.), Springer-Verlag, Chapitre 18
- Kirman A., Horst U., Teschl M., 2006, "Changing Identity : The emergence of social groups" Working paper GREQAM.
- Kirman A., Teschl M., 2006, "Searching for identity in the capability space," Journal of Economic Methodology, vol. 13(3), p. 299-325
- Lumeau M., Masclat D., Pénard T., 2010, "Les conséquences de la manipulation de la réputation dans les systèmes d'évaluation en ligne", Revue Economique, vol. 61 (6), 1123-1133.
- McLeod P. L., Leshed G., 2011, "As Long as they don't Know where I Live: Information Disclosure Strategies for Managing Identity in Second Life" in: Reinventing Ourselves: Contemporary Concepts of Identity in Virtual Worlds. Springer

- Merzeau L., 2009, "Du signe à la trace : l'information sur mesure", *Hermès*, n°59, Traçabilité et réseaux, p. 23-29.
- Rochelandet R., 2010, *Économie des données personnelles et de la vie privée*, Éd. La Découverte, coll. Repères, Paris.
- Coutant A., Stenger T., 2010, "Processus identitaire et ordre de l'interaction sur les réseaux sociaux numériques," *Mime*
- Sen A., 2002, *Rationality and Freedom*, Harvard, Harvard Belknap Press.
- Toma C., Hancock J., Ellison N., 2008, "Separating Fact from Fiction: An Examination of Deceptive Self-Presentation in Online Dating Profiles", *Personality and Social Psychology Bulletin* 34, 1023-1036.
- Turner M.A., Varghese R., 2010, "The Economic Consequences of Consumer Credit Information Sharing: Efficiency, Inclusion, and Privacy", *Mimeo*
- Tucker, C.E., 2012, "The Economics of Advertising and Privacy", *International Journal of Industrial Organization*, Vol. 30 (3) EARIE conference papers and proceedings, p. 326–329
- Turkle S., 2006, "Tethering". In *Sensorium: Embodied Experience, Technology, and Contemporary Art*, Caroline A. Jones (ed.). Cambridge, MA: List Visual Art Center and MIT Press.