



Rapport de stage

au sein du PEPS « Algorithmes et citoyenneté »
à l'Université Paris-Est Créteil

Marion VALARCHER

Sommaire

Sommaire	3
Cadre du stage	5
« Algorithmes et Citoyenneté »	5
<i>Présentation du groupe de recherche</i>	5
<i>Fonctionnement</i>	6
Un stage de sociologie dans un groupe de recherche interdisciplinaire	7
<i>Le cahier des charges du stage</i>	7
« Une sociologie de la place de la sociologie »	7
Réalisation d'un questionnaire et exploitation	9
Conception du questionnaire	9
<i>Un travail collectif</i>	9
<i>Justification des questions</i>	9
Diffusion	12
Un chapitre d'un livre collectif	15
A. <i>Diffusion du questionnaire</i>	16
B. <i>Pourquoi publier l'algorithme APB ? Retour sur une imposition de problématique</i>	18
C. « APB ? Peut mieux faire ! »	23
D. <i>Familiarités à l'informatique et stratégies</i>	27
Retour sur le travail sociologique	34
<i>Le stage : application des techniques d'enquête apprises en cours</i>	34
<i>Comment écrire et dire la sociologie ?</i>	34
<i>Un stage trop court</i>	34
Annexes	36
<i>Annexe 1 : Article publié dans le Monde du 12/12/2017</i>	36
<i>Annexe 2 : Projet de stage</i>	39
<i>Annexe 3 : Enregistrement auprès de la CNIL</i>	42
<i>Annexe 4 : Questionnaire : « Utilisations et représentations autour d'APB »</i>	43

Cadre du stage

« Algorithmes et Citoyenneté »

Présentation du groupe de recherche

L'Université Paris-Est Créteil est l'une des universités d'Île-de-France fondée en 1970 après la division de l'Université de Paris. Elle était initialement composée d'une faculté de médecine et d'un centre universitaire de droit et d'économie et elle se compose désormais et entre autres d'un UFR de Médecine, d'un UFR de Droit, d'un UFR de Sciences Economiques et de Gestion et d'un UFR de Sciences et Technologies. Il s'agit donc d'une université pluridisciplinaire et qui essaie de développer en son sein la recherche interdisciplinaire. Dans cette perspective, l'Université finance chaque année plusieurs projets interdisciplinaires « innovants et à risques » appelés des « Projets Exploratoires Premier Soutien » (PEPS). L'objectif de ce financement interne à l'université est de permettre aux projets de recherche de mûrir le temps de monter un dossier pour obtenir un financement par l'Agence Nationale pour la Recherche. En 2016, ce sont ainsi 200 000 € qui étaient destinés à financer ces PEPS de 18 mois.

C'est en réponse à cet appel à projets que s'est créé, en septembre 2015, le groupe de recherche « Algorithmes et Citoyenneté » (AlgoCit) porté par un laboratoire d'informatique et associant des chercheurs en droit, en économie, en gestion, en philosophie et en informatique. En effet, constatant que les algorithmes sont présents en très grand nombre et organisent certains pans de la vie des individus, le groupe AlgoCit interroge leur opacité aux yeux des individus qui les utilisent. Acquis à l'intérêt de la transparence des décisions administratives, les chercheurs du groupe souhaitent réfléchir sur une manière de rendre les algorithmes accessibles à tout un chacun.

Le groupe se composait donc initialement de onze enseignants-chercheurs de six disciplines différentes, rejoints en septembre 2016 par une nouvelle chercheuse :

- Mathias Bejean : gestion ;
- Patrick Cegielski : informatique ;
- Julien Cervelle : informatique ;
- Mikael Cozic : philosophie ;
- Maité Guillemain : droit privé ;
- Ronan Le Roux : communication ;
- Farida Temmak : informatique ;
- Julien Tesson : informatique ;
- Arnaud Thauvron : gestion ;
- Sylvie Thoron : économie ;
- Pierre Valarcher : informatique ;
- Noe Wagener : droit public.

Fonctionnement

Au cours des premiers mois, les membres du PEPS AlgoCit se sont réunis une fois par mois et ont tous proposé un exposé de leurs thématiques de recherche en lien avec les algorithmes. Les premiers exposés ont été présentés par les chercheurs en informatique qui présentaient le concept d'algorithme (et notamment son articulation avec celui de « programme ») de sorte que tous les chercheurs aient une définition commune de l'objet étudié. Les exposés étaient toujours suivis de questions qui permettaient aux autres membres de se familiariser avec le vocabulaire et la manière de penser de la discipline. Les interactions suivant les exposés étaient également l'occasion de proposer des ponts et des comparaisons entre les approches disciplinaires. C'était souvent à cette occasion que des idées apparaissaient ou que des résultats se profilaient. Il me revient notamment un exemple de cette collaboration : à la suite d'un exposé de droit public qui présentait notamment les lois qui régissent l'affectation à l'université, il était apparu que le droit évoquait quatre critères, mais que ces derniers n'étaient pas hiérarchisés par les législateurs. Alors que la plupart des non-familiers de l'informatique pensaient qu'il était impossible d'implémenter un programme intégrant plusieurs critères non hiérarchisés et que la discussion s'orientait sur l'impossible transfiguration de la loi en un programme, les chercheurs en informatique nous ont appris qu'il existait un langage qui permettait de ne pas hiérarchiser les critères.

Dans un deuxième temps, les exposés des membres du groupe ont été remplacés par des exposés de chercheurs extérieurs à l'université travaillant sur des thématiques proches (c'est d'ailleurs dans le cadre d'un exposé d'une personne extérieure que j'ai été présentée au groupe et qu'est apparue l'idée du stage). Dominique Cardon a par exemple proposé un exposé sur l'algorithme de Google et deux économistes sont venus présenter leur travail sur l'algorithme Affelnet qui organise l'affectation des lycéens parisiens dans un lycée de la ville. Ces exposés ont permis de mieux définir l'objet de la recherche et de proposer une nouvelle formule désignant les « algorithmes publics ». Dans le même temps, les membres du groupe de recherche ont convenu de travailler sur un algorithme public particulier, celui d'APB¹.

Par la suite, il a été question de publier un ouvrage collectif appelant à plus de transparence des algorithmes publics et proposant des pistes pour sa mise en œuvre. Les discussions se sont encore intensifiées et n'avaient plus seulement lieu lors des rencontres mensuelles, mais également par mail. De plus petits groupes se sont créés autour de thématiques précises et associaient désormais différentes disciplines — comme la philosophie et l'économie ou l'informatique et la gestion. Entre le 23 et le 26 octobre tout le groupe s'est retrouvé à Florence pour présenter de nouveaux exposés sur l'avancée du travail et décider de la suite du projet. Après ce voyage de recherche, un article collectif a été publié dans la version numérique du journal *Le Monde* en date du 12 décembre 2016 []. L'ouvrage collectif doit paraître courant 2017.

¹ APB, admission-postbac.fr, est un site Internet qui apparie les lycéens avec une formation supérieure qu'ils désirent suivre et qui les accepte. Cet appariement se fonde sur un algorithme.

Enfin, les réunions s'organisent désormais essentiellement autour de contacts avec des institutions ou des chercheurs intéressés par le thème de recherche dans le but de préparer la suite du projet qui arrive à échéance à la fin du mois de juin 2017. Le groupe a notamment été contacté par le directeur de la mission Etalab² dans le cadre d'un rapport gouvernemental sur les conditions d'ouverture d'Admission Post-Bac à qui une version provisoire du livre collectif a été transmise. D'autre part, deux membres du groupe sont invités à présenter l'ensemble des travaux à la Maison des Sciences Humaines de Paris-Saclay dans le cadre d'un travail de recherche intitulé « Le pouvoir des algorithmes ».

Un stage de sociologie dans un groupe de recherche interdisciplinaire

Le cahier des charges du stage

Je suis arrivée dans le groupe de recherche en mai 2016 à l'occasion de la présentation de Dominique Cardon à laquelle je suis allée assister. À l'issue de cet exposé et suite à des discussions avec des membres du groupe, il a été envisagé que je réalise un stage de recherche sur l'utilisation et les représentations autour de l'outil APB. J'ai donc rédigé un projet de recherche [Annexe 2] qui a été transmis au groupe accompagné de mon CV. Il a donc été décidé que je conçoive un questionnaire — dont la diffusion aux étudiants de l'université Paris-Est Créteil serait organisée par les membres du groupe — et que j'en propose une analyse. Le stage devait se dérouler, selon la convention, du 07 juillet au 30 septembre 2016 mais je suis restée associée au projet depuis et le travail d'analyse et de rédaction a eu lieu après ces dates.

Dans le même temps, insistant sur l'aspect interdisciplinaire du projet, il m'a été demandé de communiquer sur les méthodes de recherche mobilisées en sociologie afin d'en informer les chercheurs des autres disciplines. Certains chercheurs m'ont ainsi posé des questions techniques au sujet de la représentativité de mon échantillon et de la valeur d'un khi2 ainsi que sur la nature mathématique de l'opération qui permet de produire un tri croisé. Enfin, Il était également question que je présente un exposé de l'avancée de la recherche à la fin du mois d'octobre puis, plus tard, que je rédige un chapitre pour l'ouvrage collectif issu de ce travail de recherche. C'est d'ailleurs la version actuelle de ce chapitre qui constitue le corps de la partie sur le contenu du stage.

« *Une sociologie de la place de la sociologie* »³

² La mission Etalab relève de la Direction interministérielle du numérique et du système d'information et de communication de l'Etat (DINSIC) et coordonne la politique d'ouverture et de partage des données publiques (« open data »).

³ Expression reprise à Muriel DARMON dans son article « Le médecin, la sociologue et la boulangère : analyse d'un refus de terrain », 2005, Genèses

Ma position au sein du groupe était particulière pour deux raisons : j'étais à la fois la seule représentante de la sociologie et la seule étudiante. À ce titre, l'ensemble des membres du groupe a fait preuve de beaucoup de bienveillance à mon endroit, certainement du fait de leur statut d'enseignants-chercheurs habitués à aider des étudiants.

En tant qu'étudiante en sociologie, il a souvent été question de justifier des choix qui, me semble-t-il, n'auraient pas été questionnés par des sociologues. Ainsi, lors de la conception du questionnaire, il m'a fallu justifier le recours aux questions portant sur le métier et les diplômes détenus par les parents. En effet, certains membres du groupe de recherche considéraient que les catégories utilisées pour définir les groupes socioprofessionnels n'étaient pas pertinentes. Quelques minutes de discussions ont permis de comprendre la méprise : d'aucuns pensaient que les ouvriers représentaient moins de 5 % de la population active française. Le pendant de ma position dans le groupe de détentrice de la légitimité sociologique était que j'étais *in fine* libre de mes choix. De même, il m'a fallu justifier de la pertinence de l'utilisation du vocable associé aux classes sociales (« classes supérieures », « aisées », etc.) dont l'existence au XXIe a également été remise en cause.

La rédaction du chapitre du livre s'est accompagnée d'un grand nombre de relectures dans le but d'éviter les coquilles et d'homogénéiser le style du livre. Cette lecture par des personnes non familières de la sociologie m'a également poussée à être plus attentive aux termes employés. En effet, une remarque m'a été faite sur l'emploi de l'expression « capital culturel » que l'on voulait compléter par l'expression « classique » (ma version précisait pourtant « capital culturel sous la forme objectivée du diplôme »), et ma proposition « capital culturel légitime » a été refusée au motif que le terme « légitime » induisait la supériorité de ce capital culturel sur toute autre forme de culture. Cette remarque m'a donc conduit à interroger l'utilisation de ces termes par la sociologie. De même, dans le souci de ne pas froisser les enquêtés et, surtout, de ne pas tomber dans le misérabilisme, il m'a été demandé de modifier les expressions « les enquêtés ne comprennent pas » en « les enquêtés déclarent ne pas comprendre ».

Réalisation d'un questionnaire et exploitation

Conception du questionnaire

Un travail collectif

Au fur et à mesure des discussions avec les autres membres du groupe AlgoCit, quatre problématiques principales ont été définies pour structurer la conception du questionnaire [Annexe 4]. Tout d'abord, il s'agit d'appréhender la satisfaction des étudiants sur deux plans différents : d'une part la satisfaction qu'ils éprouvent à l'égard du résultat obtenu à l'issue de la procédure APB (satisfaction quant à la formation intégrée) et d'autre part celle qu'ils peuvent ressentir envers l'utilisation d'APB en tant qu'interface et que procédure de décision (satisfaction quant à l'outil APB). L'idée est ainsi de mettre à jour quelques déterminants de la satisfaction à l'usage d'un outil automatisé de décision. D'autre part, l'objectif est également de mettre en évidence, dans le cas où elles existent, quelles stratégies peuvent être mises en place lors de l'utilisation d'un tel outil et de voir ce qui explique les différences de stratégie dans l'utilisation d'APB. Enfin, il s'agit d'appréhender, pour la caractériser, la familiarité avec l'informatique — entendue comme un ensemble de connaissances algorithmiques, même élémentaires, approximatives ou intuitives.

Une fois ces problématiques définies collectivement j'ai dû trouver des indicateurs des pratiques et des manières d'appréhender les représentations des étudiants autour d'APB. Pour ce faire j'ai lu un certain nombre de questionnaires de sociologie afin de me familiariser avec les formulations des questions ainsi que les cours de master sur la conception d'un questionnaire ou encore un manuel sur la question⁴.

Justification des questions

La partie qui suit est une version adaptée d'un mail envoyé à l'ensemble du groupe. Dans ce mail je revenais sur la conception du questionnaire, la justification des questions et de leur formulation pour deux raisons : d'une part pour associer les autres membres et leur rendre des comptes quant à l'avancée mon travail, et d'autre part afin de répondre aux attentes de certains au sujet des méthodes de travail en sociologie.

L'objectif du questionnaire est de mettre à jour des relations qui pourraient exister entre la maîtrise de la notion d'algorithme et l'utilisation et les perceptions d'un algorithme en particulier : celui utilisé par l'outil admission post-bac (APB).

⁴ François de SINGLY, *Le questionnaire*, Armand Colin, coll. 128

Dans ce questionnaire il va être question de saisir principalement quatre éléments :

- la connaissance des algorithmes
- la satisfaction des lycéens qui utilisent APB
- la mise en place de stratégies dans la construction des vœux
- la confiance en APB

Dans un premier temps il s'agira donc d'interroger les néo-bacheliers sur leur utilisation d'APB.

La première question est une « question filtre » qui permet d'éliminer tous les répondants qui ne seraient pas passés par APB pour intégrer leur formation ou qui l'auraient utilisé en 2015 et non en 2016 : nous leur évitons ainsi de répondre à des questions qui ne les concerneraient pas. Nous choisissons d'éliminer les redoublants pour au moins deux raisons : la première est que leur utilisation d'APB est plus ancienne et qu'il est possible qu'ils ne se souviennent plus tout à fait de l'utilisation qu'ils ont pu en faire. La seconde raison relève d'un souci d'homogénéisation de l'échantillon : afin de pouvoir comparer toutes les réponses entre elles autant ne pas trop diversifier les caractéristiques des répondants, d'autant plus que l'utilisation d'APB pour la deuxième fois est très certainement influencée par la première utilisation.

Utilisation d'APB

Q2. Combien de vœu(x) avez-vous validé(s) sur le site APB ?

La deuxième question peut permettre de mettre en évidence différents « types » d'utilisateurs (les « inquiets » qui mettent le nombre maximal de vœux ou les « sûrs d'eux » qui n'en mettent qu'un ou deux par exemple). En créant une variable synthétique indiquant la position relative du vœu auquel le lycéen a été accepté (Q7) par rapport au nombre total de ses vœux nous obtiendrons une variable qu'il semble intéressant de croiser avec la satisfaction (en effet, ce ne semble pas être la même chose d'être accepté à son 2e vœu si l'on en a demandé 24 et si l'on en a demandé deux). D'autre part, si le but de ce questionnaire est de voir si la connaissance de la notion d'algorithme peut avoir un impact sur les perceptions d'un algorithme particulier il peut être judicieux de récolter d'autres variables pouvant expliquer les variations de cette perception dans le but de les neutraliser lors de la phase d'analyse des données.

Q3. À quelle fréquence vous connectiez-vous sur le site *admission-postbac.fr* ?

Q4. Quand avez-vous effectué les dernières modifications dans vos vœux (ajout, suppression, modification de l'ordre, etc.) ?

Q5. Quand avez-vous validé vos vœux ?

Q6. Comment avez-vous conservé votre liste de vœux ?

De la même manière que pour la deuxième question et en synthétisant les réponses à toutes ces questions (2 à 6) il semble possible de mettre à jour des « types » d'utilisateurs. Une fois ces catégories construites (à la suite d'un arbitrage qui sera opéré à partir des réponses lors de la phase d'analyse des données) nous pourrions voir si la familiarité aux algorithmes a un impact sur l'utilisation d'un outil reposant sur un algorithme particulier.

La question 6 porte sur un indicateur original et assez peu étudié en sociologie qui est le rapport à la matérialité des utilisateurs du numérique (une enquête menée dans le cadre du master sociologie d'enquête en 2016 et portant sur la lecture « papier » et la lecture « numérique » a permis de mettre en évidence l'importance du rapport à la matérialité dans l'utilisation des nouvelles technologies). Dans le cadre de ce questionnaire, la question peut être un indicateur de la confiance dans les outils numériques.

Résultats et satisfaction

Pour toutes les questions proposant une échelle (de satisfaction ou autre), quatre réponses sont proposées aux enquêtés : il s'agit d'un parti pris. Le but est de saisir la diversité des réponses possibles donc une réponse telle que « oui » ou « non » serait trop restrictive. Ensuite, concernant le nombre de propositions, il y a deux écoles. L'une d'elles voudrait que l'on propose une réponse « neutre » (ici cela pourrait être « ni satisfait, ni insatisfait »), en considérant que les répondants ont le droit de ne pas avoir d'avis sur la question. Mais cette réponse neutre a un inconvénient : c'est une réponse refuge, beaucoup de personnes peuvent choisir cette modalité pour s'épargner la peine de se poser la question. Ici d'une part nous pouvons supposer que les étudiants ont un avis concernant leur satisfaction et, d'autre part, il faut savoir que — exceptée la toute première question — aucune des questions suivantes n'est obligatoire. Si le répondant ne veut/peut pas répondre à la question il le peut tout à fait sans que cela ait de conséquences sur les suivantes. Ainsi, nous avons plutôt choisi la deuxième école et aucune réponse neutre ne sera proposée.

Classement des vœux

Lors de la procédure APB, le site rappelle régulièrement aux futurs bacheliers de ne prendre en compte dans l'ordonnancement des vœux que leurs goûts personnels. Ainsi, dans le mail adressé aux lycéens récapitulant la liste ordonnée de leurs vœux il leur est précisé : « Lors des phases d'admission vous ne pouvez avoir, au mieux, qu'une seule proposition d'admission : la meilleure possible dans votre liste préférentielle. Il est donc très important de hiérarchiser ses vœux, non pas en fonction de ses chances d'être classé, mais en fonction de ses préférences. » L'objectif de la deuxième partie du questionnaire est ainsi de comprendre si la seule logique d'ordonnancement des vœux est celle recommandée par APB ou si les lycéens mettent en place des stratégies, et si oui, quelles sont-elles.

Connaissance des algorithmes

L'objectif de cette partie est de tester la familiarité avec l'informatique en général et la notion d'algorithme plus particulièrement. La familiarité avec les algorithmes recouvre deux réalités : la connaissance théorique de ce qu'est un algorithme et des connaissances plus pratiques au sujet des algorithmes que l'on rencontre dans la vie de tous les jours. Cette familiarité est appréhendée à travers divers indicateurs, dont des indicateurs académiques qui permettent de voir si les élèves ont bénéficié d'enseignements de l'informatique en classe (bac obtenu, option suivie, etc.). Ces indicateurs soulèvent cependant des difficultés : 1. on ne sait pas ce que les élèves

ont retenu des cours et on se doute que les cours ne sont ni assimilés de la même manière par les lycéens, ni enseignés de la même manière partout, etc. 2. on se retrouve dans une situation qui va nécessairement conduire à une critique de l'enseignement de l'informatique au lycée alors que ce n'est pas l'objectif.

Une solution alternative aurait pu être de tester les étudiants, leur proposer des exercices permettant de déterminer leurs connaissances théoriques autour des algorithmes et des exercices d'applications afin de voir s'ils ont compris le fonctionnement concret de certains algorithmes. Mais il nous a semblé assez problématique de proposer un contrôle de connaissances dans un questionnaire, d'autant plus qu'il est diffusé par des professeurs, sur une adresse mail universitaire : cela pourrait aisément être pris pour un examen. La solution envisagée se fonde sur des déclarations (le principe même du questionnaire d'ailleurs) ; on demande aux étudiants de s'autoévaluer et de dire s'ils ont conscience et s'ils ont compris le fonctionnement de certains algorithmes rencontrés dans leur vie quotidienne.

Équité d'APB

Il y a une autre possibilité pour avoir l'avis d'une personne sur un sujet, en lui proposant une échelle de réponses (par exemple de 1 à 5 ou de 1 à 10), mais c'est relativement impersonnel et à partir du moment où il est possible de proposer des réponses clairement identifiables il vaut mieux le faire (non seulement parce que c'est vécu comme plus « humain » par les répondants, mais également parce que rien ne nous assure que l'échelle des valeurs est la même pour tous). Lorsque nous choisissons de proposer une échelle numérique, la question est souvent couplée avec d'autres questions plus facilement objectivables dans le but de comparer le ressenti des personnes avec des caractéristiques plus objectives. Dans ce questionnaire, la seule échelle numérique qui est proposée porte sur l'équité d'APB et elle fonctionne de pair avec les questions suivantes qui portent sur des critères tels que le fait d'avoir déjà été confronté à des situations inéquitables.

Caractéristiques sociales des enquêtés

Cette dernière partie peut permettre de retrouver des résultats précédemment mis en évidence par des recherches sur la familiarité avec les outils numériques. En effet, des études sur les pratiques numériques des adolescents ont montré que la « fracture numérique » entre les générations que décrivent régulièrement les médias peut se retrouver entre des adolescents issus de groupes sociaux différents. De même, les groupes sociaux ont des rapports différents à la fois à l'État et aux institutions qui l'incarnent. On retrouvera certainement des distinctions dans la partie portant sur l'équité et la confiance en APB ainsi que dans la mise en place de stratégies lors de la saisie des vœux sur APB.

Diffusion

Le questionnaire a d'abord été transmis aux étudiants de première année de l'IUT de Paris Descartes dont j'avais contacté la scolarité par mail. Dans les autres filières de première année, je n'ai pas eu de réponse des responsables pédagogiques.

La décision a donc été prise de contacter les étudiants en première année sur *Facebook*, en s'inscrivant dans des groupes Facebook de promotion. Voici un tableau récapitulatif des groupes Facebook dans lesquels le questionnaire a été publié.

	Université/École	Nombre d'étudiants sur le groupe	Réactions
Maths info	Paris 5	30	
Informatique	Paris 8 et Paris 13	83	
Eco-gestion	Rennes 1	343	Vidéo « M'en bat les couilles, rien à foutre »
Info-com	Nancy	315	
AES	Bordeaux	520	
AES	Strasbourg	457	
Sciences sociales	Lyon 2	246	4 mentions « j'aime »
Droit	Rouen	883	
LEA	Rouen	382	
Psycho	Nancy	805	1 mention « j'aime »
Psycho	Aix-Marseille	1277	
Sciences de la vie	Nancy	337	1 mention « j'aime »
Sciences de la vie	Strasbourg	291	
Biologie / Biochimie	Orléans	249	
Génie civil	La Réunion	96	
STAPS	Amiens	968	
PACES	Bordeaux	3397	3 mentions « j'aime »
PACES	Paris Sud	535	

Notons tout d'abord que la diffusion d'un questionnaire sur des groupes *Facebook* induit nécessairement des difficultés quant au suivi des caractéristiques des enquêtés. En effet, il n'est notamment pas possible de savoir si tous les membres du groupe Facebook ont vu la publication et il est donc encore moins aisé de savoir quel type de personnes a répondu au questionnaire.

Nous avons pu constater que les réactions à la publication n'étaient pas les mêmes selon la discipline du groupe. C'est dans le groupe de la L1 de Sciences Sociales à Lyon 2 que nous avons obtenu le plus grand nombre de mentions « j'aime ». Cela peut être interprété à la fois comme une

incarnation de la solidarité des étudiants en sociologie et comme l'expression d'un intérêt plus grand des étudiants en SHS pour ce type de questionnement. D'autre part, nous avons obtenu 3 mentions « j'aime » dans le groupe de la PACES de l'université de Bordeaux. Nous proposons de l'interpréter comme l'expression d'une insatisfaction des étudiants en PACES qui sont notamment tirés au sort par APB pour leur affectation dans une université. Enfin, sur le groupe de la L1 d'économie-gestion de Rennes 1 un étudiant a commenté la publication en y mettant une vidéo intitulée « M'en bat les couilles, rien à foutre ! », cette réaction nous éclaire sur l'effet d'imposition de problématique que peut constituer le questionnaire pour une partie des personnes contactées. C'est d'autant plus flagrant que nous avons constaté une absence de réaction de la part des étudiants sur la majorité des groupes *Facebook* contactés. Les étudiants ne semblent pas, à première vue, préoccupés par cette thématique.



Enfin, le questionnaire a été diffusé aux étudiants de première année à l'université Paris-Est Créteil sur leurs adresses mail universitaires après un enregistrement auprès de la CNIL [Annexe 3], à la suite d'échanges avec le correspondant Informatique et Libertés de l'université Paris-Est Créteil. Cela concerne tous les étudiants de L1 et d'IUT de l'université excepté les étudiants en PACES qui n'ont pas pu être contactés.

Un chapitre d'un livre collectif

Les pages qui suivent sont consacrées à la reproduction intégrale du chapitre du livre collectif à paraître. Le ton peut parfois sembler assez éloigné de la sociologie (on remarquera par exemple qu'une CAH a été réalisée sur les axes d'une ACM qui n'est pas présentée ici parce que cela nous a semblé trop ardu pour une publication plus ou moins grand public). En effet, le public visé par l'ouvrage est un public non constitué de spécialistes et il a fallu adapter le ton et le contenu autant que possible sans toutefois recourir à des usages illégitimes de la sociologie. D'autre part, le chapitre est le treizième du livre à paraître, il se fonde notamment sur un présupposé développé dans les chapitres précédents selon lequel il est nécessaire de rendre publique la procédure de décision administrative qui conduit à l'affectation des lycéens dans le supérieur. Enfin, la première partie du chapitre revient sur la diffusion du questionnaire, et peut donc paraître redondante avec la partie précédente même si elle s'attache davantage à décrire les caractéristiques de l'échantillon.

S'il nous apparaît nécessaire de rendre lisibles les décisions administratives prises automatiquement par des algorithmes — à nous, chercheurs et étudiants acquis à la nécessité de la transparence des algorithmes administratifs et familiers des procédures automatisées de décision — il est essentiel de se demander ce qu'en pensent les principaux intéressés, à savoir les utilisateurs de l'outil pris en exemple. Dans cette perspective, un questionnaire (présent en annexes) a été diffusé auprès d'étudiants en première année d'études supérieures, bacheliers depuis la session 2016. Après une diffusion en ligne par l'intermédiaire des relations des personnes engagées dans le projet et des institutions auxquelles ils se rattachent, le questionnaire a rassemblé près de 800 réponses complètes et analysables.

Notons tout d'abord que dans cet ouvrage nous avons jusqu'alors pris le parti de défendre la lisibilité des procédures qui conduisent à une prise de décision automatisée. Mais qu'en pensent les citoyens ? Sont-ils tous demandeurs d'une telle transparence ? Et quelles sont les conditions de possibilité d'une demande de transparence administrative ? La réponse à ces questions occupera une première partie de ce chapitre et, une fois quelques réponses proposées, il sera question d'interroger la satisfaction des étudiants à l'égard de l'outil APB sur le plan de l'équité de la procédure. Qui sont les étudiants qui trouvent APB équitable ? Qu'est-ce qui motive ce jugement ? Au contraire, pourquoi certains étudiants jugent-ils APB inéquitable ? Enfin, dans un dernier temps, nous nous efforcerons de questionner la familiarité aux algorithmes d'une génération « connectée ». À partir de l'enquête menée auprès des étudiants, nous tenterons de proposer une définition de la « familiarité à l'informatique » (sous-entendue dans la familiarité au « numérique », mais rarement étudiée indépendamment) au regard des pratiques concrètes des étudiants d'aujourd'hui. Une fois cette définition posée, il s'agira de comprendre dans quelle mesure cette familiarité peut influencer sur les comportements dans l'utilisation d'algorithmes

institutionnels. En bref, nous nous demanderons si une plus grande familiarité avec l'informatique favorise la mise en place de stratégies dans l'utilisation d'APB.

A. Diffusion du questionnaire

Quatre problématiques principales structurent le questionnaire tel qu'il a été diffusé. Tout d'abord, il s'agit d'appréhender la satisfaction des étudiants sur deux plans différents : d'une part la satisfaction qu'ils éprouvent à l'égard du résultat obtenu à l'issue de la procédure APB (satisfaction quant à la formation intégrée) et d'autre part celle qu'ils peuvent ressentir envers l'utilisation d'APB en tant qu'interface et que procédure de décision (satisfaction quant à l'outil APB). L'idée est ainsi de mettre à jour quelques déterminants de la satisfaction à l'usage d'un outil automatisé de décision. D'autre part, l'objectif est également de mettre en évidence, dans le cas où elles existent, quelles stratégies peuvent être mises en place lors de l'utilisation d'un tel outil et de voir ce qui explique les différences de stratégie dans l'utilisation d'APB. Enfin, il s'agit d'appréhender, pour la caractériser, la familiarité avec l'informatique — entendue comme un ensemble de connaissances algorithmiques, même élémentaires, approximatives ou intuitives.

Le questionnaire a été conçu entre le mois de juin et le mois d'août 2016 à la suite d'échanges entre les membres du groupe AlgoCit, avec comme objectif de commencer à le diffuser dès le début du mois de septembre, une fois les étudiants inscrits dans leurs formations post-bac respectives. Dès le 2 septembre, il a donc été diffusé auprès d'étudiants en IUT, en CPGE ou dans des écoles post-bac, par l'intermédiaire de leurs enseignants ou en partageant le questionnaire sur des groupes de promotion trouvés sur *Facebook*. Avant de pouvoir le diffuser massivement à l'ensemble des étudiants en première année de l'UPEC, un enregistrement auprès de la CNIL a été effectué (cet enregistrement est disponible en annexes), dans lequel les chercheurs s'engageaient à ne pas recueillir de données sensibles et à ne pas conserver les données sur une durée trop longue. Le 17 octobre les étudiants de l'UPEC ont été contactés via leur adresse courriel étudiante et, parallèlement, le questionnaire a été à nouveau utilisé sur des groupes de classe sur *Facebook* dans des universités et des écoles non parisiennes.

Soulignons que la diffusion du questionnaire induit des spécificités quant à la structure de l'échantillon puisque nous n'avons pas, à notre connaissance, réussi à le faire parvenir à des étudiants en BTS et que, d'autre part, les étudiants en IUT qui ont répondu sont, pour une grande partie d'entre eux, des étudiants en DUT Informatique. Cette particularité n'est pas sans conséquence pour la suite de l'analyse, puisque nous cherchons à analyser les effets de la familiarité avec l'informatique sur l'utilisation d'APB ; il faudra donc tenir compte de cette spécificité des étudiants en DUT, familiers de l'informatique, au moins dans la mesure où ils se sont orientés dans une formation dispensant des enseignements d'informatique.

Le 9 novembre 2016, nous disposons de 799 réponses « complètes » (c'est-à-dire que les répondants sont allés au bout du questionnaire, même s'ils n'ont pas toujours répondu à

l'intégralité des questions). Aussi avons-nous pris la décision de clôturer la passation du questionnaire, pour commencer à analyser les données ainsi recueillies.

L'échantillon retenu pour l'analyse comprend ainsi 799 individus recrutés essentiellement par contact électronique, et par « effet boule de neige ». Cet échantillon est « spontané », car composé de volontaires pour répondre au questionnaire et il n'est donc pas conçu comme représentatif au sens de la méthode des quotas. Cependant, il agit comme un « prisme »⁵ qui permet d'apercevoir une population particulière dont nous pouvons produire une description a posteriori. C'est pourquoi il est primordial de décrire de la façon la plus précise possible la structure de l'échantillon ainsi construit.

Structure de l'échantillon

Signalons tout d'abord que le nombre d'étudiantes ayant répondu à cette enquête est très nettement supérieur au nombre d'hommes, puisque l'échantillon est constitué à 62 % de femmes contre seulement 36 % d'hommes (et 2 % de non-réponses à la question). Cependant, les analyses ne semblent pas montrer de variations significatives des comportements selon le genre ainsi nous pouvons supposer que le déséquilibre de l'échantillon n'entraîne pas de distorsion majeure des résultats.

D'autre part, nous remarquons que les enquêtés sont principalement issus de milieux intermédiaires ou favorisés, mais pas davantage que l'ensemble de la population des étudiants puisque les proportions sont sensiblement les mêmes que celles que l'on retrouve dans les chiffres fournis par le ministère de l'Enseignement supérieur en 2007⁶. Soulignons cependant que la diffusion du questionnaire induit des spécificités quant à la structure de l'échantillon puisque nous n'avons pas, à notre connaissance, réussi à le faire parvenir à des étudiants en STS. Cela explique la très faible part d'étudiants titulaires d'un baccalauréat professionnel ou technologique dans notre échantillon (15 %) puisque 67 % des bacheliers professionnels qui poursuivent leurs études sont inscrits dans une STS (la proportion est de 50 % pour les bacs technologiques)⁷. D'autre part, les étudiants en IUT qui ont répondu sont, pour une grande partie d'entre eux, des étudiants en DUT Production Informatique (les enseignants de cette formation, impliqués dans le projet de recherche, ont dégagé quelques minutes au début de leurs cours en salle informatique afin que les étudiants répondent au questionnaire). Cette particularité n'est pas sans conséquence pour la suite de l'analyse ; il faudra donc tenir compte de la spécificité des étudiants en IUT, familiers de

⁵ Martin, O., 2012, *L'analyse quantitative des données*, Paris, Armand Colin, Coll. 128.

⁶ EDUCATION NATIONALE, 2007, *Repères et références statistiques*, « Les étudiants » à partir des systèmes d'information SISE et SCOLARITE. Il s'agit de statistiques réalisées à partir des données recueillies lors des inscriptions administratives des étudiants dans un établissement du supérieur. Ces données sont ensuite envoyées au ministère de l'éducation nationale qui se charge de produire des statistiques notamment sur l'origine sociale des étudiants

⁷ INSPECTION GÉNÉRALE DE L'ÉDUCATION NATIONALE, 2012, « Analyse des poursuites d'études des lycéens à partir de la procédure d'admission post-bac », Rapport n°2012-123 remis à la ministre de l'enseignement supérieur et de la recherche.

l'informatique au moins dans la mesure où ils se sont orientés dans une formation dispensant des enseignements d'informatique.

Toujours concernant la familiarité aux algorithmes, la répartition des réponses à la question portant sur l'écriture d'un programme informatique nous a particulièrement étonnés. En effet, 85 % des répondants ont déclaré en avoir déjà écrit un programme et 33 % ont déclaré en écrire souvent. L'importance de ces proportions peut s'expliquer par l'inscription de l'algorithmique au programme de seconde générale et par la surreprésentation des bacheliers scientifiques — censés être équipés d'une calculatrice programmable pour leurs cours de mathématiques et ayant reçu un enseignement plus avancé sur les algorithmes dans le cadre de leur formation scolaire. Ainsi, nous pouvons supposer que cet échantillon va nous donner « une image stylisée [de la population étudiante générale] par l'accentuation des traits pertinents »⁸ puisque nous nous retrouvons en présence d'un grand nombre d'étudiants à première vue familiers des algorithmes. Il sera ainsi plus aisé de distinguer des degrés et des types de familiarité aux algorithmes dans une population qui semble plus familière que la population de référence.

Précisons, s'il était encore nécessaire, que cette étude, du fait même de sa nature exploratoire et de son mode de passation, ne permet pas de donner des résultats valables pour l'ensemble de la population étudiante. Cependant, elle nous suggère quelques hypothèses vérifiées dans notre échantillon qu'il serait pertinent de tester dans un travail systématique à plus grande échelle. Ainsi, quand nous parlerons des étudiants dans ce chapitre nous demandons au lecteur de garder à l'esprit qu'il s'agira uniquement des étudiants de l'échantillon (qui ont des caractéristiques bien particulières que nous venons d'évoquer).

B. Pourquoi publier l'algorithme APB ? Retour sur une imposition de problématique

La mise en place et la diffusion du questionnaire ont nécessairement soulevé quelques questions : n'imposons-nous pas aux répondants de se poser une question qu'ils n'envisageaient pas initialement ? Et, dès lors, que valent les réponses aux questions qu'ils ont pu donner ? Se demander si les utilisateurs sont favorables à la publication dudit algorithme va impliquer de prendre en considération leur familiarité aux algorithmes.

Effectifs et fréquence des réponses à la question portant sur la nécessité de la publication de l'algorithme d'APB

	Effectifs	Fréquence
Non réponse	81	10,1 %

⁸ BOLTANSKI L. et MALDIDIER P., 1977, *La vulgarisation scientifique et son public*, Paris, Centre de Sociologie européenne

Non, pas du tout	32	4,0 %
Non, pas vraiment	106	13,3 %
Oui, un peu	206	25,8 %
Oui, tout à fait	374	46,8 %
Total	799	100 %

L'une des questions posées dans le questionnaire était ainsi formulée ([Q26]) : « Une association de lycéens a demandé la publication de l'algorithme d'APB, jugez-vous cette publication nécessaire ? » Quatre modalités de réponses étaient alors proposées : « Oui, tout à fait », « Oui un peu », « Non, pas vraiment » et « Non, pas du tout ». À cette question, 10 % des enquêtés n'ont pas répondu — il s'agit d'une proportion particulièrement élevée au regard des non-réponses aux autres questions (à titre d'exemple seuls 2 % des enquêtés n'ont pas répondu à la question « Avez-vous déjà entendu parler d'algorithme ? »). Nous faisons l'hypothèse qu'il s'agissait d'étudiants n'ayant pas d'avis sur la question, ou qui ne comprenaient pas bien de quoi il relevait. Mais qu'ont voulu dire les étudiants déclarant ne pas être d'accord avec l'idée selon laquelle il faut publier l'algorithme d'APB ? Nous proposons comme hypothèses au moins six réalités différentes auxquelles peut renvoyer le fait de ne pas être d'accord avec la publication de l'algorithme APB.

- Tout d'abord, le répondant peut considérer que certaines décisions administratives doivent rester secrètes. Les défenseurs des secrets d'État, convaincus au nom d'une raison supérieure que tout ne doit pas être rendu public, ont ainsi pu se déclarer contre la publication de l'algorithme d'APB.
- Il peut également s'agir d'un positionnement lié à la non-compréhension de la question ou de ses implications. En effet, au cours d'entretiens réalisés avec des élèves de Terminale il est apparu qu'un certain nombre de jeunes ne s'étaient pas particulièrement intéressés à la question de la transparence autour de l'algorithme d'APB. Ainsi, une lycéenne de Terminale S a pu dire « Oui j'en ai entendu parler au journal, je crois, mais qu'est-ce que c'est ? ». Des élèves soucieux de répondre aux questions qui leur étaient adressées, mais qui ne comprenaient pas la teneur de la question ont pu préférer dire qu'ils n'étaient pas d'accord avec une affirmation qu'ils n'entendaient pas complètement.
- Nous pouvons également avoir affaire à des opposants convaincus de l'utilisation d'outils numériques pour prendre des décisions. Ils ne sont pas d'accord avec le fait qu'il faudrait publier l'algorithme d'APB puisqu'il faudrait, selon eux, s'abstenir d'utiliser APB ou tout autre outil se fondant sur un traitement automatisé des données.

- D'autre part, la publication de l'algorithme d'APB peut ne pas constituer une solution idéale dans un souci de transparence des décisions publiques ; un étudiant un peu informé peut donc avoir conscience que cette publication n'est pas suffisante et qu'elle devrait s'accompagner d'une aide à la compréhension de l'algorithme.
- Enfin, la question évoquait la requête d'une association de lycéens, ainsi, une partie des répondants a pu se positionner, à travers sa réponse, contre les associations de lycéens. Une mauvaise expérience avec des associations lycéennes ou des convictions politiques opposées aux associations ou syndicats étudiants peut éventuellement conduire une partie des enquêtés à rejeter en bloc toute proposition de publication de l'algorithme du fait même de l'origine de cette demande.
- Enfin, et il nous semble que cette explication doit plus probablement concerner les lycéens interrogés, les étudiants ont, pendant leur année de Terminale, été formés à l'utilisation d'APB au cours de séances d'information au lycée et ont pu lire le guide du candidat qui semble expliquer relativement clairement le fonctionnement de l'outil. Ainsi, considérant qu'ils avaient compris le fonctionnement d'APB, ils ne remettent pas en cause l'implémentation du programme, accordant une grande confiance à la communication du ministère. Ils n'estiment pas nécessaire de vérifier la justesse de l'algorithme.

Régression linéaire appliquée au fait de ne pas vouloir la publication de

	Ne pas être du tout d'accord avec l'idée selon laquelle il faudrait publier l'algorithme APB	Coefficient
Constante		0,035
Profession du père	Cadre supérieur	-0,029
	Cadre moyen	-0,007
	Employé	-0,054
	Ouvrier	-0,04
	Sans profession	<i>réf.</i>
Profession de la mère	Cadre supérieur	-0,004
	Cadre moyen	0,03
	Employée ou ouvrière	-0,002
	Sans profession	<i>réf.</i>
	Diplôme de 2ème ou 3ème cycle	-0,02

	Ne pas être du tout d'accord avec l'idée selon laquelle il faudrait publier l'algorithme APB	Coefficient
Diplôme du père	Bac ou diplôme de 1er cycle	-0,027
	Diplôme inférieur au bac	0,028
	Non réponse ou ne sait pas	-0,012
	Pas de diplôme	réf.
Diplôme de la mère	Diplôme de 2ème ou 3ème cycle	0,007
	Bac ou diplôme de 1er cycle	-0,001
	Diplôme inférieur au bac	-0,001
	Non réponse ou ne sait pas	0,001
	Pas de diplôme	réf.
Genre	Homme	-0,022
	Femme	réf.
Satisfaction quant à la formation finalement obtenue	Pas du tout satisfait	0,096**
	Pas vraiment satisfait	-0,005
	Assez satisfait	0,011
	Très satisfait	réf.
Avoir déjà constaté qu'on peut obtenir des résultats différents sur un moteur de recherche depuis des ordinateurs différents	Non et je ne comprends pas pourquoi	0,047*
	Non mais cela ne m'étonne pas	0,047*
	Oui et je comprends pourquoi	0,02
	Oui mais je ne comprends pas pourquoi	réf.
Mode de conservation de la liste de vœux APB	Imprimée	0,066**
	Sur son ordinateur ou dans ses mails	0,019
	En ligne, sur le site APB	réf.
<p>Le R2 de cette régression est de 0,009, il est ainsi très proche de 0. Les résidus sont donc grands par rapport à la variance ; de ce fait, il est très difficile d'analyser la régression dans un registre prédictif puisque l'erreur est relativement importante.</p> <p>*** : $p < 0,001$ ** : $p < 0,01$ * : $p < 0,05$</p>		

Une analyse par régression linéaire nous indique qu'une fois neutralisés les effets des métiers et des diplômes des parents ainsi que celui du genre et de la satisfaction du répondant, le fait de n'avoir jamais constaté que la même recherche depuis deux ordinateurs différents ne

donnait pas nécessairement les mêmes résultats est corrélé positivement avec le fait d'être tout à fait d'accord avec le fait que l'algorithme soit publié. N'avoir jamais fait ce constat peut être appréhendé comme une absence d'expérience de la présence des algorithmes dans la vie quotidienne. En effet, ne pas faire l'expérience des algorithmes peut impliquer de ne pas avoir conscience des implications que ces algorithmes peuvent avoir dans la vie quotidienne, autant pour la recherche en ligne que pour le logiciel d'affectation des formations post-bac. D'autre part, on observe une corrélation entre le fait de se positionner contre la publication et le fait de se déclarer très insatisfait de la formation finalement obtenue sur APB, toutes les variables prises en compte étant égales par ailleurs. Ce résultat est assez perturbant : les moins satisfaits ne questionneraient pas la procédure qui les affecte dans une formation insatisfaisante ? Ainsi, les étudiants qui se positionnent contre la publication de l'algorithme semblent donc être des déçus d'APB n'ayant pas d'intuition de la présence des algorithmes dans la vie de tous les jours. Nous supposons qu'il s'agit d'étudiants qui ne savent pas que la publication d'un algorithme réclamée par les associations de lycéens était une demande de transparence, notamment parce qu'ils n'ont peut-être pas connaissance de la présence d'un algorithme à l'origine de la décision d'affectation qu'ils ont reçue ou parce qu'ils n'ont pas conscience qu'il existe des critères de sélection et des informations sur le fonctionnement d'APB qui peuvent être explicitées. En effet, certains étudiants ne sont peut-être pas au fait de l'existence de règles explicites qui conduisent à l'attribution d'une formation à chaque étudiant.

D'autre part, toutes les variables précitées étant égales par ailleurs, on observe également une corrélation positive entre le fait d'avoir conservé une version imprimée de sa liste de vœux et le fait de ne pas vouloir publier l'algorithme. Pour une génération que l'on considère aisément comme très familière des outils numériques, la conservation d'une version papier de la liste de vœu peut être interprétée comme l'indicateur d'une moindre aisance avec le numérique et les données dématérialisées. Il semble donc que cette moindre aisance, qui aurait pu être corrélée à une plus grande méfiance vis-à-vis des algorithmes, aille de pair avec une moindre conscience des enjeux liés à la présence d'algorithmes dans le quotidien. Ainsi, les étudiants qui ne sont pas du tout d'accord avec la demande de publication de l'algorithme sont principalement des étudiants relativement peu familiers des algorithmes dans le sens où ils n'ont pas fait l'expérience de la présence des algorithmes dans leur vie quotidienne et où ils ne sont pas à l'aise avec la dématérialisation des données.

Cette régression nous permet ainsi de mettre en évidence les corrélations de quelques variables avec le fait de s'opposer à la publication de l'algorithme d'APB. Les variables ainsi identifiées nous montrent qu'une familiarité moindre avec les algorithmes ou le numérique est liée à ce que nous pourrions *a priori* appeler une plus grande confiance en APB. Nous avons pu formuler quelques hypothèses expliquant le refus de la publication de l'algorithme : cette requête implique une certaine aisance avec le numérique et une familiarité avec les algorithmes. Comment

peut-on vouloir la publication d'un algorithme si l'on n'a aucune conscience de l'existence d'un tel algorithme et de sa manière de fonctionner ?

Précisons cependant que l'objectif d'une telle étude n'était pas de montrer qu'un certain nombre des administrés était contre la transparence des algorithmes ; c'est même le contraire puisque 72 % des répondants considèrent à l'inverse qu'il est nécessaire de publier cet algorithme. Mais cela signifie-t-il pour autant que les trois quarts des répondants ne sont pas satisfaits du fonctionnement d'APB ?

C. « APB ? Peut mieux faire ! »

Nous avons demandé aux étudiants de noter l'équité d'APB sur une échelle de 0 à 10 : s'ils jugent qu'APB est très équitable, ils peuvent lui attribuer une note maximale de 10/10 et, dans le cas contraire, ils peuvent descendre jusqu'à 0/10. Seuls 3,75 % des répondants au questionnaire se sont abstenus de répondre à cette question, et parmi ceux qui ont répondu 8 % ont mis 0/10 et 2 % ont mis 10/10.

Fréquences cumulées des notes de l'équité d'APB

Note à l'équité d'APB	Fréquence	Cumul	Cumul inversé
0	8,3 %	8,3 %	100,0 %
1	3,0 %	11,3 %	91,7 %
2	6,1 %	17,4 %	88,7 %
3	9,9 %	27,2 %	82,6 %
4	10,8 %	38,0 %	72,8 %
5	18,4 %	56,4 %	62,0 %
6	11,4 %	67,9 %	43,6 %
7	16,7 %	84,6 %	32,1 %
8	11,4 %	96,1 %	15,4 %
9	2,4 %	98,4 %	3,9 %
10	1,6 %	100,0 %	1,6 %
Total	100,0 %		
	<i>Champ</i> : les 760 étudiants ayant répondu à la question portant sur l'équité d'APB		

Il est apparu que plus d'un tiers des étudiants juge que, sur le plan de l'équité, APB ne mérite pas une note supérieure à 4/10.⁹ Que signifie cette proportion ? Qu'est-ce qui n'est pas équitable dans APB selon ces étudiants ? Nous avons procédé à une régression linéaire afin de mettre en évidence les liens entre certaines caractéristiques des répondants et le fait de mettre une note inférieure ou égale à 4.

Régression linéaire appliquée à une note inférieure ou égale à 4/10 sur l'équité attribuée à APB

	Note inférieure ou égale à 4/10 pour l'équité	Coefficient
Constante		0,106
Profession du père	Cadre supérieur	-0,077
	Cadre moyen	-0,053
	Employé	-0,111
	Ouvrier	-0,084
	Sans profession	<i>réf.</i>
Profession de la mère	Cadre supérieur	0,011
	Cadre moyen	0,019
	Employée, ouvrière	0,040
	Sans profession	<i>réf.</i>
Diplôme du père	Diplôme de 2ème ou 3ème cycle	0,144*
	Bac ou diplôme de 1er cycle	-0,005
	Diplôme inférieur au bac	0,062
	Non réponse ou ne sait pas	0,199
	Pas de diplôme	<i>réf.</i>
Diplôme de la mère	Diplôme de 2ème ou 3ème cycle	-0,119
	Bac ou diplôme de 1er cycle	-0,022
	Diplôme inférieur au bac	-0,052
	Non réponse ou ne sait pas	-0,100
	Pas de diplôme	<i>réf.</i>
Satisfaction quant à la formation finalement obtenue	Non réponse ou pas satisfait	0,178***
	Assez satisfait	0,039
	Très satisfait	<i>réf.</i>

⁹ Nous sommes bien conscients des limites de cette question et c'est pourquoi elle était complétée par des questions portant sur des pratiques concrètes des étudiants. Pourtant, il nous semble important de prendre en compte les représentations des étudiants sur l'équité d'APB, tout en proposant d'expliquer ces représentations par des caractéristiques et des pratiques plus « objectivables » des étudiants.

	Note inférieure ou égale à 4/10 pour l'équité	Coefficient
Être étonné du résultat obtenu	Étonné	0,219***
	Pas étonné	-0,035
	A intégré son premier vœu	réf.
Avoir entendu parler de personnes ayant bénéficié de procédures particulières pour intégrer la formation	Non réponse	-0,037
	Oui et c'est courant	0,217***
	Oui mais ce n'est pas courant	0,060
	Non mais c'est courant	0,101*
	Non et ce n'est pas courant	réf.
Utilisation d'un logiciel de programmation de type algobox	Oui	-0,075*
	Non, jamais	réf.
Formation actuelle	École postbac	0,027
	Licence non sélective	0,046
	Licence sous tension	0,052
	PACES, droit, psychologie ou STAPS	0,119*
	IUT, CPGE ou STS	réf.
	<p>Le R2 de cette régression est de 0,196, il est ainsi plus proche de 0 que de 1. Les résidus sont relativement grands par rapport à la variance ; de ce fait, il très difficile d'analyser la régression dans un registre prédictif puisque l'erreur est relativement importante.</p> <p>*** : $p < 0,001$ ** : $p < 0,01$ * : $p < 0,05$</p>	

Il est possible d'identifier deux logiques principales pour rendre compte des corrélations ainsi observées. D'une part, mettre une mauvaise note à l'équité d'APB est fortement corrélé à des situations particulières d'iniquité rencontrées par les étudiants : s'ils jugent APB inéquitable, il semble que c'est parce qu'ils ont été confrontés à des situations qu'ils ont jugé inéquitables. Mais une seconde logique, moins intuitive que la précédente, nous indique qu'une mauvaise note à l'équité est corrélée à une moindre proximité avec les algorithmes. En effet, un certain degré de proximité aux algorithmes, appréhendé à partir de l'utilisation d'un logiciel d'aide à l'élaboration et à l'exécution d'algorithme (comme le logiciel libre AlgoBox, le plus souvent utilisé en Terminale) diminue la probabilité de mettre une note inférieure ou égale à 4/10 à l'équité d'APB.

De manière relativement intuitive, nous pouvons constater que le fait d'avoir été confronté à certaines situations qui peuvent être jugées inéquitables au regard de la loi est fortement corrélé avec le fait de juger APB inéquitable. De toute évidence, un étudiant en PACES, droit, psychologie ou STAPS a entendu parlé du recours au tirage au sort pour départager des étudiants lors de l'affectation dans une formation étant donné que ces quatre formations sont susceptibles de recourir à ce procédé pour répartir les étudiants entre des universités où il n'y a pas assez de places. Ainsi, des étudiants conscients du recours au tirage au sort et concernés par ce tirage au sort sont enclins à juger APB inéquitable parce qu'ils jugent inéquitable l'une des modalités d'affectation des lycéens et parce qu'ils ont été concernés par cette modalité. On ne peut expliquer ce sentiment d'iniquité par une insatisfaction de leur part — éventuellement liée au tirage au sort d'ailleurs — puisque nous avons pris le parti d'annuler l'effet propre de la satisfaction dans ce raisonnement par régression linéaire. D'autre part, toutes les variables prises en compte étant égales par ailleurs, le fait de penser qu'il est courant que certaines personnes bénéficient de « procédures particulières » (par l'intermédiaire d'un proche par exemple) lors de leur affectation est également corrélé positivement avec le fait de juger APB inéquitable. On remarque une corrélation encore plus grande lorsque l'étudiant a lui-même entendu parler de personnes qui ont bénéficié de telles procédures. Il s'agit d'une corrélation entre une situation objective et un sentiment d'iniquité. D'autre part, les personnes ayant répondu qu'elles n'avaient pas entendu parler de telles pratiques, mais qui pensent tout de même que c'est un phénomène courant font preuve d'une suspicion de principe à l'égard d'APB qui n'est pas liée à une situation objective d'inégalité qu'ils auraient pu rencontrer. Dès lors, on peut se demander ce qui conduit l'étudiant à considérer ce type de situations comme courant. S'il n'a pas entendu parler directement de personnes à qui s'est arrivé, ce sentiment peut faire écho aux discours des médias eux-mêmes suspicieux à l'égard d'APB, à ceux des enseignants qui peuvent parfois être très critiques vis-à-vis d'un logiciel sur lequel ils n'ont pas particulièrement de maîtrise ou encore à ceux des parents qui jugent le processus compliqué et peu lisible...

Il est également possible d'identifier un biais de représentativité dans le fait de trouver APB inéquitable. En effet, on observe une corrélation positive entre l'insatisfaction quant à la formation obtenue et le fait de juger APB inéquitable en l'absence de variation des autres variables. On suppose que les étudiants jugent l'équité d'APB à l'aune de situations qu'ils connaissent et donc *a fortiori* à partir de leur propre situation à la suite d'un raisonnement par similarité. S'ils ne sont pas satisfaits, ils peuvent se focaliser davantage sur leurs connaissances qui ne sont pas non plus satisfaites de leur situation et, de ce fait, considérer qu'APB est inéquitable parce qu'il fait un grand nombre d'insatisfaits.

La corrélation est encore plus importante avec le fait d'être étonné du résultat obtenu à l'issue de la procédure APB. L'étonnement vis-à-vis de la réponse peut être pensé comme un indicateur des stratégies d'utilisation d'APB. En effet, les personnes les plus étonnées étaient sensées être celles qui ont été les moins informées des écarts entre leurs prétentions et leurs chances d'intégrer. Ainsi, on peut considérer que les étudiants qui ont le moins de ressources pour

anticiper leurs résultats sur APB jugent APB inéquitable puisqu'ils pensaient obtenir un vœu mieux classé.

Enfin, le fait d'avoir déjà utilisé un logiciel d'aide à l'élaboration et à l'exécution d'algorithmes est corrélé négativement avec le fait de juger APB inéquitable. Il semble que le fait de comprendre comment fonctionne un algorithme diminue les chances de juger l'algorithme APB inéquitable. Les étudiants qui ont été familiarisés aux algorithmes — puisqu'ils en ont conçu — ont peut-être une conscience plus grande qu'il y a un algorithme derrière la décision APB et que, étant donné qu'il suit des règles explicitées *a priori*, il est forcément équitable à l'aune des règles qu'il implémente. Ce résultat est particulièrement éclairant puisqu'il nous indique l'existence d'une corrélation entre un premier indicateur de la familiarité aux algorithmes et les représentations des utilisateurs d'un algorithme particulier. On peut ainsi deviner dans quelle mesure les utilisateurs d'un objet comme APB ont conscience et connaissance des logiques qui sous-tendent la décision et comment cette connaissance influe sur les représentations de cet objet. Par la suite, il sera question de mettre à jour des modalités de la familiarité à l'informatique afin de mettre en évidence leur influence cette fois sur l'utilisation d'APB.

D. Familiarités à l'informatique et stratégies

La principale difficulté dans l'élaboration du questionnaire était de mettre à jour quelques questions afin de différencier les étudiants selon leur familiarité à l'informatique non pas au sens courant de l'utilisation quotidienne des ordinateurs, mais dans un sens plus scientifique renvoyant à la discipline universitaire de traitement automatique et relationnel de l'information. La porte d'entrée dans cette proximité à l'informatique qui a été choisie est celle de l'*algorithme*, objet complexe par sa technicité et pourtant appréhendable par une population étudiante du fait de sa confrontation au quotidien avec un tel objet. Les algorithmes sont présents au quotidien et les étudiants interrogés ont été aux prises avec un algorithme particulier lors de leur orientation post-bac. Commençons par reprendre avec Cédric Fluckiger¹⁰ le concept de « capital informatique ». Ce dernier, à l'instar du capital culturel bourdieusien, existe sous trois formes : objectivée ou matérielle (que constitue l'accès à un matériel informatique comme des ordinateurs ou des connexions internet, etc.), sous la forme institutionnalisée (à travers les certifications ou les diplômes, ici B2i, C2i, DUT d'informatique, etc.) et, enfin sous une forme incorporée. Cette dernière forme du capital informatique est la plus polymorphe puisqu'elle renvoie à des « savoir-faire pratiques », des « habiletés techniques » ou encore à des « schèmes d'utilisation ». Ces trois formes de capital impliquent essentiellement une définition quotidienne de l'informatique comme une manière de se *débrouiller* quand on est face à un ordinateur. Qu'en est-il de la compréhension

¹⁰ FLUCKIGER C., 2007, « Les collégiens et la transmission familiale d'un capital informatique », *Agora débats/jeunesses*, 4/2007 (N° 46), p. 32-42.

des logiques *derrière* les opérations quotidiennes ? Un parallèle peut être fait avec des individus à l'aise avec l'informatique du quotidien, la bureautique ou les outils de recherche, mais qui ne maîtrisent pas les logiques à l'origine des résultats obtenus lors d'une recherche sur Google par exemple. Une autre illustration : je peux savoir qu'en cliquant sur un bouton en particulier sur le logiciel Excel je peux calculer une moyenne sans pour autant connaître la formule du calcul d'une moyenne. Si le capital informatique tel qu'il est présenté par Cédric Flickiger semble se référer essentiellement à cette informatique du quotidien, nous proposons, à travers le terme d'algorithme, d'appréhender l'informatique sous un jour un peu plus théorique. En cliquant sur les boutons d'APB pour candidater et valider leurs inscriptions dans le supérieur, les étudiants ont-ils pleinement conscience de la présence d'un algorithme de décision ? Une conscience d'un tel algorithme modifie-t-elle les usages qui sont faits d'APB ? C'est à ces questions que nous allons essayer de répondre ici.

Le capital informatique incorporé, s'il fait référence à des savoir-faire et des habiletés pratiques, recouvre également la familiarité aux algorithmes que nous allons tenter de mettre en évidence. Cédric Fluckiger a montré que certains collégiens rencontrés avaient « une disposition à dépasser le simple engagement pragmatique qui caractérise les usages de presque tous les collégiens rencontrés et à considérer l'ordinateur dans sa dimension d'objet technique ». C'est cette considération d'APB comme un objet technique et les modalités de celle-ci que nous allons tenter de définir. Ce travail permet de dégager au moins quatre types de proximité aux algorithmes. On retrouve chacune de ces modalités dans, au moins, une des questions du questionnaire qui sert alors d'indicateur de cette proximité. En premier lieu, nous avons pu identifier une « proximité terminologique aux algorithmes », appréhendée par le fait de déclarer savoir ce qu'est un algorithme (89 % des étudiants sont concernés par la réponse), cette proximité se caractérise par le fait d'avoir déjà entendu le terme « algorithme » et de comprendre à quel domaine il fait référence, même sans être en mesure d'en produire une définition précise. D'autre part, il était possible de discerner une proximité cette fois plus « expérimentale » que l'on saisit à travers le fait d'avoir déjà écrit un programme, y compris sur une calculatrice programmable (cela concerne 85 % des étudiants ayant répondu). Cette « familiarité expérimentale » est l'indicateur d'un rapport empirique aux algorithmes même si cette familiarité a pu être très contextuelle. En effet, rien ne dit que le grand nombre de programmes écrits n'est pas essentiellement dû à l'évolution des programmes scolaires puisqu'ils incluent désormais une initiation à l'algorithmique dans les cours de mathématiques dès la classe de seconde et que cette initiation passe notamment par la programmation sur des logiciels spécifiques. C'est pourquoi nous dissocierons les étudiants n'ayant jamais écrit de programmes de ceux qui l'ont déjà fait et de ceux qui le font souvent, introduisant ainsi une échelle dans la proximité expérimentale. En outre, nous avons également pu mettre en évidence l'existence d'une « proximité quotidienne aux algorithmes », entendue comme le fait d'avoir fait l'expérience de l'incidence d'algorithmes dans la vie quotidienne sans nécessairement avoir conscience que ce sont des algorithmes qui sont à l'origine des situations observées (par exemple, avoir déjà constaté qu'un même moteur de recherche pouvait renvoyer des réponses

différentes en fonction du poste utilisé pour effectuer la recherche, ou encore avoir déjà constaté que les publicités en ligne évoluaient au gré de l'historique de navigation). Enfin, la proximité quotidienne est complétée par une « proximité réflexive » qui se caractérise par le fait d'avoir conscience de la présence d'algorithmes derrière des outils utilisés au quotidien.

Ces quatre modalités de la proximité aux algorithmes ne sont pas indépendantes les unes des autres et nous nous trouvons en fait en présence d'une échelle de Guttman¹¹. Cela signifie que les différentes proximités s'échelonnent de telle sorte que la présence d'un certain type de proximité chez le répondant suppose qu'il est familier des algorithmes selon toutes les modalités qui lui sont préalables. On peut ainsi constater non seulement que ces familiarités sont hiérarchisées, mais également qu'il y a un effet cumulatif des différentes modalités de la proximité aux algorithmes. Ces différentes proximités s'échelonnent dans l'ordre de leur évocation, de la plus répandue dans la population (la proximité terminologique) à la moins courante (la proximité réflexive). À titre d'illustration, cela se traduit par le fait que certains étudiants font preuve d'une familiarité terminologique avec les algorithmes, que d'autres font aussi preuve d'une familiarité expérimentale, mais qu'aucun de ceux qui ne font pas preuve d'une familiarité terminologique ne fait preuve d'une familiarité expérimentale.

En plus de mettre en évidence des natures de proximité différentes, nous avons pu distinguer les étudiants selon l'intensité de leur proximité aux algorithmes. Celle-ci a été appréhendée en attribuant des points en fonction des réponses données par les étudiants : les étudiants qui ont le plus souvent répondu aux questions par des modalités que l'on jugeait révélatrices de la proximité à l'informatique ont obtenu un nombre de points plus important que ceux qui répondaient davantage par des modalités indiquant une certaine distance aux algorithmes. À titre d'exemple, un étudiant qui savait ce qu'était un algorithme, mais qui n'avait jamais écrit de programme, y compris sur une calculatrice programmable, obtenait 1 point sur un maximum de 4 (maximum qui pouvait être atteint si l'étudiant déclarait écrire des programmes « souvent »)¹². À partir de ce score, nous avons pu constituer six groupes d'étudiants de tailles similaires selon l'intensité de leur proximité aux algorithmes.

¹¹ Techniquement, nous nous trouvons en présence d'une échelle de Guttman avec un coefficient de reproductibilité de Goodenough de 0,958, le critère de Matalon est donc rempli. Près de 80% de l'échantillon est compris dans le patron théorique.

¹² Cette variable score d'intensité de la proximité était conçue comme suit :

- 1 point pour savoir ce qu'est un algorithme, 0 sinon
- 1 point pour avoir déjà écrit un programme, 3 points pour en écrire souvent, 0 sinon
- 1 point pour avoir déjà utilisé un logiciel d'aide à l'élaboration et l'exécution d'algorithmes, 0 sinon
- 2 points pour avoir déjà constaté que les moteurs de recherche pouvaient donner des résultats différents et déclarer comprendre pourquoi, 1 point si on ne l'avait pas constaté mais qu'on n'était pas étonné, 0 dans les autres cas,
- 1 point si l'on avait déjà constaté que les publicités en ligne étaient ciblées et si l'on comprenait pourquoi ou si l'on ne l'avait pas constaté mais qu'on en était pas étonné, 0 dans les autres cas

Ainsi, il était possible d'obtenir un score compris entre 0 et 8 points au maximum.

Score		Proportions	Cumul croissant	Cumul décroissant
Inférieur ou égal à 2	Très éloignés	14,1 %	14,1 %	100 %
Égal à 3	Éloignés	14,8 %	28,9 %	85,9 %
Égal à 4	Assez éloignés	17,5 %	46,4 %	71,1 %
Égal à 5	Assez proches	18,8 %	65,2 %	53,6 %
Égal à 6	Proches	16,8 %	82 %	34,8 %
Supérieur ou égal à 7	Très proches	18,0 %	100 %	18,0 %

En fait, ces questions portant sur la familiarité à l'informatique nous ont permis d'identifier sept groupes parmi les étudiants¹³ qui s'échelonnent à la fois selon l'intensité et la nature de la proximité aux algorithmes.

- Un premier groupe rassemble les étudiants qui semblent les moins familiers de l'informatique puisqu'on y retrouve les 11 % d'étudiants qui n'ont jamais entendu le terme « algorithme » et qui ne savent pas à quelle réalité il fait référence. Dans ce groupe la moyenne obtenue au score d'intensité de la proximité est de 2,01 et 68 % des membres du groupe ont obtenu un score inférieur ou égal à 2. Il s'agit d'un groupe très distant des algorithmes sur tous les plans évoqués plus tôt : tant sur le plan de la compréhension terminologique de l'objet que sur le plan d'une proximité plus expérimentale puisque la classe se caractérise par une surreprésentation des étudiants n'ayant jamais écrit un programme. Mais également sur les plans quotidiens et intellectuels puisque les étudiants n'ayant pas constaté de variation des résultats lors d'une recherche en ligne depuis des postes différents et qui disent ne pas comprendre pourquoi il en serait ainsi sont également surreprésentés dans ce groupe par rapport à l'ensemble de la population étudiée. Cette classe se caractérise également par une surreprésentation des étudiants dont le père n'a aucun diplôme et 60 % des étudiants de ce groupe ont un père sans profession, ouvrier ou employé. Il s'agit d'une classe regroupant des étudiants ayant une origine sociale relativement faible et dont les deux parents disposent de peu de capital culturel sous la forme objectivée du diplôme. La classe est également marquée par une forte surreprésentation des étudiants titulaires d'un bac technologique ou professionnel. Appelons cette classe, par abus de langage, classe de « l'absence de familiarité aux algorithmes » et notons qu'elle est aussi caractérisée par une origine sociale et scolaire plus défavorisée que le reste de l'échantillon.

¹³ Nous avons procédé à une Classification Ascendante Hiérarchique selon la méthode de Ward sur le carré des distances sur les axes d'une Analyse des Composantes Multiples sur cinq questions :

- la connaissance du terme « algorithme »,
- l'utilisation d'un logiciel d'aide à l'élaboration et l'exécution d'algorithmes,
- l'exécution d'algorithmes notamment sur une calculatrice programmable,
- la conscience de la présence d'un algorithme dans les moteurs de recherche en ligne,
- la conscience de la présence d'un algorithme dans les publicités ciblées en ligne.

- La deuxième classe que nous pouvons identifier se caractérise par une très nette surreprésentation des étudiants n'ayant jamais écrit de programme. Ce groupe rassemble 9 % des répondants et les étudiants ont obtenu, en moyenne, un score de 2,86 à l'intensité de leur proximité aux algorithmes. Ce groupe se caractérise également par une surreprésentation des titulaires d'un bac S ou ES, et des étudiants en licence non sélective. D'autre part, il faut constater une légère surreprésentation des étudiants dont le père n'occupe pas de profession au moment de l'enquête. Nous nommerons cette classe « absence de familiarité expérimentale aux algorithmes » et nous remarquons que les étudiants qui la composent ont moins souvent suivi des études secondaires scientifiques. L'origine sociale et scolaire des étudiants qui composent la classe est légèrement plus élevée que pour la classe précédente, mais toujours moins que l'ensemble des étudiants interrogés dans le cadre de l'enquête.
- Un troisième groupe se dégage également et il est cette fois-ci essentiellement caractérisé par une surreprésentation des étudiants n'ayant pas remarqué que les publicités en ligne peuvent être ciblées. A fortiori, les étudiants n'ont pas non plus remarqué que les résultats des recherches en ligne dépendent de l'historique de recherche et de navigation. Il s'agit d'une classe de l'« absence de familiarité quotidienne aux algorithmes » qui regroupe 12 % de l'échantillon. En moyenne, les étudiants de cette classe ont obtenu un score de 3,57 pour l'intensité de leur proximité. Cette classe ne semble pas se caractériser par une sur ou une sous-représentation majeure des étudiants en fonction de leurs caractéristiques sociales et scolaires (bac, études suivies, genre, métiers et diplômes des parents).
- La quatrième classe qui apparaît ici peu être nommée la classe d'un « début de familiarité quotidienne aux algorithmes ». En effet, elle se caractérise essentiellement par une surreprésentation des étudiants ayant remarqué que les publicités en ligne étaient ciblées, mais également par une sous-représentation de ceux qui n'ont pas constaté que les résultats des recherches en ligne dépendaient des historiques de recherche et de navigation. Environ 10 % des étudiants de l'échantillon se retrouvent dans ce groupe qui a une moyenne de 4,19 au score de l'intensité de la proximité aux algorithmes. Sur le plan des caractéristiques sociales et scolaires des étudiants on remarque une surreprésentation des titulaires d'un bac ES ou L, bac général non scientifique, et des étudiants en licence non sélective. D'autre part, le groupe se caractérise également par une surreprésentation des étudiants dont le père n'exerce pas de profession.
- Le cinquième groupe que l'on peut identifier regroupe 14 % des étudiants. Ce sont des étudiants qui font preuve de la familiarité quotidienne aux algorithmes dont nous venons de parler tant du point de vue des publicités ciblées que de la recherche en ligne. Cependant, le groupe se caractérise par une surreprésentation des étudiants qui ont bien fait l'expérience des algorithmes dans leur vie quotidienne, mais qui ne parviennent pas à se l'expliquer, à en donner une cause formelle. Ils semblent ne pas identifier la présence d'algorithmes. Appelons cette classe « familiarité quotidienne, mais pas réflexive aux algorithmes » et remarquons que les

étudiants de cette classe 5 ont obtenu en moyenne un score très proche de la classe précédente : 4,27. D'autre part, nous pouvons noter une surreprésentation des étudiants dont la mère est employée et de ceux dont le père appartient aux classes moyennes.

- La sixième classe a des caractéristiques particulières si l'on se fie à nos attentes. En effet, elle est essentiellement caractérisée par une surreprésentation des étudiants n'ayant pas remarqué que les moteurs de recherche donnaient des résultats différents selon l'historique de recherche, mais qui ne sont pas étonnés de cela. Ils semblent faire preuve d'une familiarité réflexive, mais pas d'une familiarité quotidienne. La classe regroupe tout de même 19 % des étudiants interrogés et ceux-ci ont obtenu, en moyenne, un score de 5,49 à l'intensité de leur familiarité. Ce groupe sera celui de la « familiarité réflexive, mais pas quotidienne aux algorithmes ». On remarque une surreprésentation des étudiants issus de milieux aisés (dont l'un au moins des parents appartient aux classes supérieures) et qui réussissent scolairement : ils sont plus souvent titulaires d'un bac S et étudiants dans une école post-bac que le reste de la population de l'échantillon.
- Enfin, la dernière classe que nous avons pu mettre à jour regroupe 26 % de l'échantillon et pourrait être nommée classe du cumul des « familiarités quotidienne et réflexive aux algorithmes ». En moyenne, les étudiants qui composent cette classe ont obtenu un score de 6,63 à l'intensité de la proximité. Cette classe se caractérise par une surreprésentation des étudiants en CPGE ou IUT Informatique et des étudiants titulaires d'un bac S. Contrairement aux classes précédentes qui semblaient s'échelonner le long de l'axe de la dotation en capitaux (appréhendée à partir de la profession des parents), cette classe se caractérise par une surreprésentation des étudiants dont le père appartient aux classes moyennes et non supérieures. Cela nous semble lié à la structure de l'échantillon : cette classe regroupe un grand nombre d'étudiants en IUT d'informatique, assez familiers des algorithmes, ce qui débouche par construction sur une surreprésentation des classes moyennes étant donné les caractéristiques de la population inscrite en IUT en France.

Certes nous avons mis à jour une échelle de la familiarité aux algorithmes. *Quid ?* Commençons par ne pas sous-estimer ce que nous venons de mettre en évidence : savoir qu'il existe différents types de familiarité et que ceux-ci ne sont pas indépendants les uns des autres n'est pas sans intérêt. D'autant plus que certaines de ces familiarités semblent avoir des effets sur les stratégies adoptées face à APB.

Faire preuve de familiarité terminologique vis-à-vis des algorithmes est ainsi corrélé avec le fait de mettre les formations sélectives demandées au début de la liste de vœux. Cette question avait été pensée dans l'idée de distinguer les étudiants qui avaient bien compris le fonctionnement d'APB des autres. En effet, il était par exemple inutile de classer une formation sélective après une licence qui n'était pas sous-tension puisqu'on était sûr de l'avoir et il apparaît que savoir ce qu'est

un algorithme est très lié avec le fait d'avoir conscience qu'il est inutile de mettre une formation non sélective après une formation que l'on est certain d'intégrer.

Mais il semble que c'est la familiarité quotidienne qui est la plus corrélée à des stratégies. En effet, les étudiants ne faisant pas du tout preuve de familiarité quotidienne sont plus souvent étonnés du résultat qu'ils obtiennent, et obtiennent d'ailleurs moins souvent leur premier vœu que ceux qui en font preuve. Sans pour autant tomber dans un causalisme trop simple, nous pouvons nous demander si la familiarité quotidienne ne permet pas aux étudiants de mieux appréhender les critères de sélection appliqués dans le supérieur et d'évaluer leur dossier au regard de ces critères. Nous venons de mettre à jour des inégalités sociales dans la familiarité aux algorithmes et cette familiarité semble permettre d'appréhender plus sereinement les algorithmes présents dans la vie de tous les jours, comme celui d'APB.

Retour sur le travail sociologique

En guise de conclusion et d'ouverture, je propose de revenir brièvement sur le travail sociologique réalisé, sur les limites de ce travail et les améliorations qui pourraient y être apportées.

Le stage : application des techniques d'enquête apprises en cours

Dans un premier temps, l'un des avantages du stage a été de me permettre de mettre en pratique des connaissances acquises dans le cadre du master qu'il s'agisse de la méthodologie de la conception ou de la diffusion d'un questionnaire ou des techniques d'analyse statistique. Le stage s'étant déroulé sur une relativement longue période, la mobilisation de certaines techniques était souvent concomitante avec leur apprentissage dans le cadre des cours. Ainsi, le stage a souvent constitué pour moi l'occasion de réaliser un exercice supplémentaire sur chaque technique et de lier l'apprentissage d'une technique avec une réflexion sociologique. Cela m'a également permis de comprendre l'intérêt de travailler sur R. En effet, j'ai d'abord essayé de réaliser des régressions sur *Modalisa*, mais je n'arrivais pas à choisir mes modalités de référence.

Comment écrire et dire la sociologie ?

Deux difficultés très liées se sont présentées à moi lors de cette année de stage. En effet, le cadre universitaire des études nous conduit souvent à écrire des devoirs de sociologie pour des sociologues. Dans ce cadre précis, il m'a été demandé de communiquer sur des travaux de sociologie soit à des universitaires dans d'autres disciplines soit à un public ciblé pour le livre également constitué de non-sociologues. Cela m'a donc semblé relativement complexe d'expliquer ma démarche à des personnes qui n'y étaient pas familières, voire parfois plutôt opposées au mode de raisonnement sociologique. L'autre difficulté résidait dans la rédaction du chapitre qui me semblait compliquée parce qu'il fallait articuler entre elles honnêteté sociologique et clarté de l'explication pour un non-sociologue.

Un stage trop court

Si quelques critiques peuvent être formulées, elles portent essentiellement sur la durée du stage. S'il a commencé en juin 2016 et n'est toujours pas vraiment terminé, ce n'était pas un travail à temps plein et il était initialement pensé pour durer deux mois. Ces deux mois m'ont semblé trop court pour concevoir un questionnaire, en analyser des résultats et rédiger un document. Aussi, avec du recul il me semble que la formulation de certaines questions n'est pas idéale et je regrette de ne pas avoir davantage d'informations sur la filière d'étude des étudiants par exemple. D'autre part, le temps de la passation du questionnaire étant limité dans le temps, j'ai essayé au maximum

de diversifier l'échantillon ce qui peut être problématique dans le sens où il se retrouve finalement constitué de deux manières sans qu'il ne soit possible de différencier les deux échantillons. : d'un côté nous retrouvons les étudiants de l'université Paris-Est Créteil et, de l'autre, tous les étudiants contactés par mail. Enfin, le temps d'analyse était également limité et il reste une certaine frustration de ne pas pouvoir analyser l'intégralité des questions.

Annexes

Annexe 1 : Article publié dans le Monde du 12/12/2017

Civic Tech : « De l’algorithme administratif à l’algorithme public »

LE MONDE | 12.12.2016 à 13h18 • Mis à jour le 13.12.2016 à 09h12 |

À partir du 20 janvier, quelque 800 000 lycéens de terminale et étudiants en réorientation pourront se préinscrire dans les formations de l’enseignement supérieur sur la plate-forme Admission postbac (APB). Il y a fort à parier qu’APB 2017 ne dérogera pas aux éditions précédentes et que les débats resurgiront à l’approche de l’été, en particulier pour ce qui concerne les formations dites « *en tension* ».

Rappelons les derniers épisodes. Au début de 2016, l’association Droit des lycéens sollicite de la ministre de l’éducation nationale, de l’enseignement supérieur et de la recherche la divulgation du code source de l’algorithme APB. Faute de réponse, l’association saisit la Commission d’accès aux documents administratifs (CADA), qui délivre un avis favorable à sa divulgation.

Lire aussi : « Civic Tech » : vers une boîte à outils de la démocratie numérique

Cette décision conduit le ministère à rendre public, dans un premier temps, un résumé des critères utilisés pour départager les candidats à des formations « *en tension* », puis une partie du code source. Depuis, les membres de cette association se sont attelés à un véritable travail que l’on appelle en informatique « *rétro-ingénierie* », afin de reconstruire l’algorithme initial.

Suspicion

Reconnaissons que la présentation d’APB comme une boîte noire pose un certain nombre de problèmes. Du côté du public, elle alimente la suspicion à l’égard de l’administration et de la procédure mise en place. Elle fait aussi de l’individu qui a recours au service administratif numérisé un simple utilisateur focalisé sur ses préférences plutôt qu’un citoyen conscient des contraintes collectives.

La méthode adoptée pose également problème du côté de l’administration : elle n’encourage pas à faire l’effort d’explicitier et de justifier les objectifs que celle-ci s’est fixés. Elle n’encourage pas non plus à vérifier que l’algorithme effectivement utilisé est fidèle à ces objectifs. Ces problèmes ne sont pas propres à APB. Ils se posent dans les mêmes termes pour tous les algorithmes auxquels recourt l’administration pour prendre ses décisions, que ce soit Affelnet, pour l’affectation des collégiens dans les lycées, ou, demain peut-être, le fameux algorithme « *antichômage* », que développe le jeune Paul Duan avec Pôle emploi.

« Cahier des charges »

Au contraire de cette méthode de la boîte noire, une méthode plus transparente permettrait de renvoyer chacun à ses responsabilités. Encore doit-on réfléchir aux moyens de cette transparence. On peut sans doute tirer parti du fait que les algorithmes sont constitués de règles totalement explicites, mais la rétro-ingénierie à partir du code source ne suffira pas.

Concrètement, il faut effectivement que plus d'information soit disponible pour les utilisateurs et pour des experts indépendants. Quelle information ? Il faut commencer par informer le citoyen des objectifs recherchés par l'administration. Ils apparaissent dans un « *cahier des charges* » que l'administration a dû transmettre aux informaticiens qui ont conçu APB, comme l'aurait fait n'importe quel autre maître d'ouvrage pour n'importe quel algorithme.

Lire aussi : [La « civic tech » veut favoriser la participation des citoyens](#)

Seul ce cahier des charges permet de connaître les choix sur la base desquels l'algorithme a été développé : choix d'interprétation des dispositions du code de l'éducation, choix sociaux, choix politiques. Il faut aussi rendre public le descriptif des propriétés centrales de l'algorithme que réalise le programme conçu par les informaticiens ; enfin, parce que la construction d'un système technique qui puisse répondre au cahier des charges n'a rien d'évident, il faut donner accès au programme lui-même.

Processus ouvert au débat public

Ainsi il sera possible pour des experts de vérifier qu'il satisfait bien à la description qu'en donne l'administration. On peut aussi vouloir aller plus loin et, dans l'esprit de ce qui a été fait lors de la préparation de la loi pour une République numérique, avec la mise en place d'une plate-forme de concertation en ligne, imaginer que le processus même de construction de l'algorithme puisse être ouvert au débat public.

Les contours d'une telle méthode demandent encore à être affinés, mais il ne fait guère de doute, à nos yeux, qu'il faut œuvrer pour que les algorithmes administratifs dont la place et le rôle croissent constamment, deviennent de véritables « algorithmes publics ».

Ces propositions, loin d'être utopiques, s'inscrivent parfaitement dans le Plan d'action national de la France (2015-2017) pour une action publique transparente et collaborative et satisfont son engagement à l'« *ouverture des modèles de calculs et des simulateurs de l'Etat* ».

Lire aussi :

La France qui vient de prendre la présidence de l'Open Government Partnership (Partenariat pour un gouvernement ouvert) doit réfléchir à la mise en œuvre concrète de la loi du 7 octobre 2016 pour une République numérique, qui rend obligatoire la publication en ligne, par les administrations, « *des règles définissant les principaux traitements algorithmiques utilisés dans l'accomplissement de leurs missions lorsqu'ils fondent des décisions individuelles* ».

Les signataires du collectif : Mathias Bejean (maître de conférences en gestion), **Patrick Cegielski** (professeur d'informatique), **Julien Cervelle** (professeur d'informatique), **Mikael Cozic** (maître de conférences en philosophie), **Maïté Guillemain** (maître de conférence en droit privé), **Ronan Le Roux** (maître de conférences en sciences de l'information et de la communication), **Farida Semmak** (maître de conférences d'informatique), **Arnaud Thauvron** (maître de

conférences en gestion), **Sylvie Thoron** (professeur en économie), **Pierre Valarcher** (professeur d'informatique, responsable du projet), **Noe Wagener** (maître de conférences en droit public).

Stage de sociologie quantitative au sein du PEPS AlgoCit

Présentation de l'organisme d'accueil

Le groupe de recherche AlgoCit (Algorithme et Citoyenneté) est composé de chercheurs de l'université Paris-Est Créteil dans les domaines de l'informatique, l'algorithmique, les sciences politiques, la philosophie, le droit et la gestion. Ce groupe, financé par un Projet Exploratoire Premier Soutien, s'intéresse aux relations qu'entretiennent les algorithmes avec la cité, avec la vie citoyenne.

Constatant que les algorithmes sont présents en très grand nombre et organisent certains pans de la vie des individus, le groupe AlgoCit interroge leur opacité aux yeux des individus qui les utilisent. L'algorithme est un outil qui fait des choix en suivant les règles qui le construisent et il faut s'interroger sur la moralité de cet outil c'est-à-dire sur la justesse de ses critères de choix. À partir de la définition de critères de validité, la rétroingénierie permet de tester cette justesse. Se pose alors la question de la définition de ces critères de justesse — qui sont des enjeux politiques —, mais aussi celle, dans un second temps, du moyen de vérifier que l'algorithme remplit la fonction pour laquelle il a été construit et donc de vérifier qu'il remplit bien les critères susnommés.

En tant qu'ils sont opaques, les algorithmes ne sont pas compréhensibles par le plus grand nombre, y compris par les individus qui les utilisent pourtant. Le fait qu'ils soient opaques alors même qu'ils influent sur de larges segments de nos vies (depuis notre orientation scolaire jusqu'au règlement des impôts en ligne) peut s'avérer être un problème pour la démocratie, attendu que les principes démocratiques impliquent sinon une participation des citoyens à la construction de la cité à tout le moins un contrôle de leur part. Les algorithmes — au moins ceux que l'on pourrait qualifier de « publics », c'est-à-dire créés à la demande d'un État — constituent donc un outil politique ; ils sont le résultat de l'implémentation de décisions politiques et ont des conséquences concrètes pour tout un chacun. Cependant, il est relativement compliqué de mettre en évidence l'impact des différents algorithmes sur les individus étant donné qu'ils modifient principalement les espaces des possibles qui leur sont ouverts. Il faut se demander s'il modifient les comportements humains, et si oui, dans quelle mesure.

Enfin, les chercheurs d'AlgoCit se demandent dans quelle mesure les utilisateurs d'un algorithme connaissent les critères de ce dernier. En effet, on peut se demander si les utilisateurs d'un algorithme ont une bonne intuition de son fonctionnement. C'est-à-dire qu'il faut se demander dans quelle mesure les citoyens connaissent les conséquences que leurs actions peuvent avoir sur le résultat final de l'algorithme. Les individus maîtrisent-ils l'algorithme de telle sorte à élaborer des stratégies de comportement face à celui-ci ?

Objectifs du stage

L'apport du travail sociologique dans le cadre de ce stage pourrait être de donner un éclairage sur les pratiques et les représentations des individus liées à ces algorithmes. Il s'agirait donc d'interroger la connaissance d'un algorithme en particulier et de mettre à jour des comportements face à celui-ci. Étant actuellement en train de réaliser un mémoire sur l'orientation scolaire en classe de Terminale je propose de réaliser et d'analyser les résultats d'un questionnaire destiné aux utilisateurs du portail admission post bac.

En supposant que le fonctionnement de l'algorithme d'admission postbac n'est pas compréhensible par la grande majorité des lycéens utilisateurs, il semble pertinent d'interroger la connaissance qu'ils ont du fonctionnement d'APB et donc de se demander s'ils ont conscience des conséquences de leurs actions sur l'appariement final. En fait, il paraît important de voir si les lycéens qui utilisent APB pour candidater dans des formations de l'enseignement supérieur élaborent des stratégies dans le choix et le classement de leurs vœux de telle sorte à optimiser leurs résultats ; et, dans le même temps, il faudrait mettre à jour ces éventuelles stratégies en les expliquant par exemple par les différents états de connaissance de l'outil APB des utilisateurs. L'intérêt est donc double : l'étude permettrait à la fois de comprendre les différences qui pourraient exister dans la compréhension du fonctionnement d'APB entre les utilisateurs (le fait de suivre l'option ISN en Terminale S aide-t-il ? La familiarité avec certains outils informatiques est-elle explicative ? Utiliser APB pour la seconde fois facilite-t-il la compréhension de son fonctionnement ?), mais elle permettrait aussi de voir si une meilleure connaissance du fonctionnement a des conséquences sur l'utilisation de l'outil.

L'autre question qui intéresse les chercheurs du groupe AlgoCit et qui pourrait être étudiée dans ce cadre relève davantage de la notion de « justesse » d'un algorithme. Il pourrait alors s'agir d'interroger les utilisateurs d'APB au sujet de leurs propres critères de justesse concernant cet outil : il faudrait leur demander quels rôles l'outil devrait remplir selon eux ainsi que s'il les remplit effectivement à leurs yeux. En effet, les utilisateurs peuvent considérer qu'il est plus important que le plus grand monde obtienne finalement son premier vœu ou bien ils peuvent privilégier un système qui s'efforcerait d'associer le moins souvent possible un élève à son dernier vœu. Considèrent-ils qu'il vaut mieux un système qui maximise la satisfaction du plus grand nombre ou un système qui minimise l'insatisfaction de la majorité ?

Il s'agirait donc d'un questionnaire interrogeant à la fois les pratiques des utilisateurs du portail APB et leurs représentations à son sujet de telle sorte à saisir ensemble les mécanismes sociaux à l'œuvre dans l'utilisation d'un algorithme, les différentes expériences individuelles lors de la manipulation de cet outil et, éventuellement, les stratégies de comportement qui peuvent être élaborées au cours de l'usage d'APB. L'analyse s'efforcerait ensuite de mettre à jour des liens — s'ils existent — entre les différentes pratiques et les représentations des individus.

Déroulement du stage

Je propose de réaliser ce stage en trois phases. Une première, avant l'été 2016, au cours de laquelle il faudrait définir les questionnements qui intéressent davantage le groupe de chercheurs

et construire le questionnaire à partir des problématiques définies. Il s'agirait donc de trouver des indicateurs permettant d'appréhender les pratiques et les représentations des utilisateurs d'APB et de les transformer en questions auxquelles pourraient répondre les enquêtés sous la forme d'un questionnaire autoadministré.

Dans une deuxième phase qui mobiliserait l'ensemble du groupe de recherche, il s'agirait de diffuser au maximum le questionnaire aux étudiants en bac+1, plus accessibles à l'équipe composée d'enseignants chercheurs à l'université. Pour cela le réseau de chacun des membres serait mobilisé afin que le questionnaire atteigne le plus grand nombre d'étudiants. Cette phase pourrait avoir lieu au moment de la rentrée universitaire pour l'année 2016-2017.

Enfin, dans un dernier moment de cette étude, je m'efforcerai d'analyser les réponses au questionnaire et de mettre à jour des logiques d'utilisation d'APB de manière à répondre aux problématiques décidées dans la première phase. Un dossier sera remis aux membres de l'équipe d'AlgoCit une fois ce travail d'analyse réalisé et les résultats obtenus pourront leur être présentés sous la forme d'un exposé s'ils le souhaitent.

Annexe 3 : Enregistrement auprès de la CNIL



FILIERE JURIDIQUE RECHERCHE EN DROIT
DE DONNÉES À CARACTÈRE PERSONNEL

Responsable du traitement :
Université Paris-Est Créteil
83 avenue du Général de Gaulle
93017 Créteil Cedex France
upec@univ-est-creteil.fr

Co-responsable informatique et libertés
Chargés des affaires juridiques à la DPAJ
83 avenue du Général de Gaulle
93017 Créteil Cedex France
upec@univ-est-creteil.fr

"BUTIMAP" N° de consultation à la CNIL	
Titre du traitement : - Nom d'usage d'un site ou traitement (exemple : Agenda, L'agenda, Annuaire, etc.)	Questionnaire "ALDOCO"
Base de données ou source : - date à laquelle le traitement a commencé à être utilisé par l'université (ou est envisagé d'être utilisé). S'il s'agit d'un traitement utilisé depuis longtemps, l'indiquer.	01/09/2016
Finalité principale : - but du traitement : - voir exemple pour l'agenda : (noter administrative de la pédagogie et de la sécurité des étudiants, par exemple pour les budgets - gestion des horaires en (ré)tempo de 24h/24)	Opérations relatives à l'utilisation d'APP par les tiers invités
Détails des finalités du traitement :	Etude sociologique sur la question des algorithmes dans la société (confiance, comportement) Exemple avec AFB.
Services chargés de la mise en œuvre : - il s'agit du nom des services utilisant le traitement	Projet de recherche "ALDOCO" (PEPS financé par l'UPEC, Le laboratoire "AEL (Faculté des Sciences et Technologies)
Fonction de la personne ou du service exerçant le contrôle de la durée d'accès : - il s'agit des fonctionnaires de la personne à contacter pour demander la communication de données ou exiger leur destruction, si l'utilisation de ces données est indépendante	Enseignants/Chercheurs
Catégorie des personnes concernées par le traitement : - il s'agit des personnes dont les données personnelles sont traitées. Par exemple pour l'agenda : enseignants, etc.	Etudiants
Base(s) de données : - il s'agit des données relatives aux personnes concernées par le traitement	Base(s) de données traitées : - Nom(s) la (les) personne(s) (par catégorie) ; formation (niveau, formation, type, etc.) ; "Fonction exercée"

Catégories de destinataires (= il s'agit des tiers à l'université auxquels les données sont communiquées, par exemple au Ministère, au Rectorat, à la Comue, à un cocontractant). Cette case peut rester vide si les données ne sont pas communiquées à des tiers)	Catégories de destinataires Université Paris 5 (le stagiaire est étudiant à Paris 5 et les données sont stockées sur les machines du master « Sociologie de l'enquête »)	Données concernées Toutes
Moyens de sécurité (exemple : accès limité à certains agents au sein du service, mot de passe pour accéder au logiciel...)		
Régime du traitement dans le cadre de la réglementation informatique et libertés	<i>(à remplir par le CIL)</i>	
Durée de conservation (Si une « durée d'utilité administrative » DUA existe, il s'agit de cette durée (cf. documentation archive sur l'intranet). A défaut de DUA, il conviendrait de renseigner la durée de conservation suivie en pratique).	L'UPEC conserve les données à caractère personnel pendant ... 6 mois	
Dernière mise à jour de la fiche-registre	<i>(à remplir par le CIL)</i>	

Annexe 4 : Questionnaire : « Utilisations et représentations autour d'APB »

Bonjour,

Voici un questionnaire portant sur votre utilisation d'admission post-bac (APB) au cours de l'année 2015/2016. Nous vous invitons à répondre à des questions portant sur votre utilisation concrète d'APB et ce que vous avez pu en penser. Ce questionnaire est donc adressé aux bacheliers de 2016 qui poursuivent actuellement leurs études dans une formation à laquelle ils ont candidaté par l'intermédiaire d'APB.

Toutes les réponses seront totalement **anonymes**, ni vos noms ni vos coordonnées ne vous seront demandés, et les réponses ne seront conservées que dans un objectif scientifique. Ce questionnaire a été conçu dans le cadre d'un groupe interdisciplinaire de recherche qui s'intéresse à l'articulation entre les algorithmes et la citoyenneté.

Merci de répondre le plus précisément possible. Choisissez la réponse qui vous convient le mieux parmi celles qui vous sont proposées. Si aucune des réponses ne semble vous convenir, ne répondez pas à la question et poursuivez le questionnaire.

Une dizaine de minutes suffisent pour remplir ce questionnaire. Nous vous remercions vivement par avance pour votre participation.

Les chercheurs et stagiaires du groupe AlgoCit

Utilisation d'APB

Commençons par quelques questions sur votre utilisation d'APB l'année dernière, en 2015-2016.

Q1. Au cours de quelle phase APB avez-vous été accepté dans la formation que vous suivez actuellement ?

- La première phase (le 8 juin)
 - La deuxième phase (le 23 juin)
 - La troisième phase (le 14 juillet)
 - J'ai été accepté en 2015 et je redouble mon année
 - Je n'ai pas été accepté dans cette formation sur APB
- si la réponse était l'une des deux dernières propositions, le questionnaire s'arrêterait ici.

Q2. Combien de vœu(x) avez-vous validé(s) sur le site APB ?

Nombre entre 1 et 24

Q3. À quelle fréquence vous connectiez-vous sur le site admission-postbac.fr ?

- Tous les jours
- Plusieurs fois par semaine
- Plusieurs fois par mois
- Seulement juste avant les échéances de saisie et de validation des vœux

Q4. Quand avez-vous effectué les dernières modifications dans vos vœux (ajout, suppression, modification de l'ordre, etc.) ?

- Au moins un mois avant la date limite
- Quelques semaines avant la date limite
- Quelques jours avant la date limite

- Le jour de la date limite

Q5. Quand avez-vous validé vos vœux ?

- Au moins un mois avant la date limite
- Quelques semaines avant la date limite
- Quelques jours avant la date limite
- Le jour de la date limite

Q6. Comment avez-vous conservé votre liste de vœux ?

- J'allais la consulter en ligne sur le site APB
- Je l'ai reçue par mail
- Je l'ai enregistrée sur mon ordinateur
- Je l'ai imprimée

Résultats et satisfaction

Voici quelques questions qui vont porter sur les résultats que vous avez obtenus lors des différentes phases d'APB et sur votre satisfaction quant à ces résultats.

Q7. Quel était le numéro du vœu auquel vous avez été finalement accepté à l'issue des trois phases d'APB ?

Vœu entre 0 et 24

Il s'agit donc probablement du vœu dans lequel vous poursuivez actuellement vos études.

Q8. En étiez-vous satisfait ?

- Oui, très satisfait
- Oui, assez satisfait
- Non, pas vraiment satisfait
- Non, pas du tout satisfait

Q9. Étiez-vous étonné de ne pas être pris dans un vœu mieux classé ?

- Oui, très étonné
- Oui, assez étonné
- Non, pas vraiment étonné
- Non, pas du tout étonné

Q10. Dans quel type de formation poursuivez-vous vos études actuellement ?

- Classe prépa, IUT, BTS
- PACES, droit, psychologie, STAPS
- École sur concours
- École sur dossier
- Licence à capacité limitée dite « sous tension »
- Licence non sélective

Q11. Envisagez-vous de vous réorienter dans une autre formation au cours de l'année scolaire ou à la fin de l'année ?

- Oui
- Non

Classement des vœux

Nous nous intéressons ici à la manière dont vous avez pu ordonner vos vœux.

Q12. Vous êtes-vous renseigné sur les critères de sélection des formations ?

- Oui, en cherchant sur Internet
- Oui, j'ai demandé à des étudiants qui suivaient la formation
- Oui, j'ai demandé à des enseignants ou personnels administratifs de la formation
- Oui, j'ai demandé à d'autres personnes qui me semblaient mieux informées
- Non, je ne me suis pas renseigné
- Non, je n'ai candidaté qu'à des licences non sélectives sans pression de sélection

Q13. Avez-vous commencé votre liste de vœux par des formations sélectives ?

- Les formations sélectives que j'ai demandées étaient au début de ma liste de vœux
- Les formations sélectives que j'ai demandées étaient réparties tout au long de ma liste de vœux
- Les formations sélectives que j'ai demandées étaient à la fin de ma liste de vœux
- Je n'ai pas demandé de formations sélectives

On entend par « formation sélective » les formations telles que les IUT, les BTS, les classes préparatoires et les écoles sur concours ou sur dossier.

Q14. Auriez-vous été satisfait quel que soit le vœu que vous auriez intégré ?

- Oui, tous les vœux que j'ai mis me plaisaient autant
- Oui, mais certains de mes vœux me plaisaient plus que d'autres
- Non, certains de mes vœux ne me plaisaient pas vraiment
- Non, certains de mes vœux ne me plaisaient pas du tout

Q15. Votre premier vœu était-il votre préféré ?

- Oui, tous mes vœux étaient classés par ordre de préférence
- Oui, le premier était mon préféré
- Non, ils étaient classés selon un autre critère

Q15bis. Quel était cet autre critère ?

Veuillez écrire votre réponse ici.

Q16. Pour choisir vos vœux et leur ordre avez-vous pris en compte la sélectivité des formations et vos chances d'intégrer ?

- Oui, beaucoup
- Oui, un peu
- Non, pas vraiment
- Non, pas du tout

Q17. Comment vous êtes-vous principalement renseigné sur les formations que vous demandiez ?

- Sur le site APB
- Sur les sites Internet des formations
- Sur des sites spécialisés (ONISEP, l'Étudiant.fr, etc.)
- En discutant avec des personnes mieux informées (camarades, professeurs, conseillers d'orientation, lors de portes ouvertes, lors de salons, etc.)
- Je ne me suis pas renseigné sur les formations

Connaissance des algorithmes

Dans cette partie nous essayons de voir quelle familiarité vous entretenez avec les algorithmes et l'informatique en général.

Q18. Disposez-vous personnellement des outils numériques suivants ?

- Un ordinateur fixe
- Un ordinateur portable
- Un smartphone
- Une tablette
- Une console de jeux vidéos

Q19. Avez-vous déjà constaté qu'en recherchant la même chose sur un moteur de recherches depuis deux ordinateurs différents vous n'obteniez pas les mêmes résultats ?

- Oui et je comprends comment cela fonctionne
- Oui, mais j'ignore comment cela fonctionne
- Non, mais cela ne m'étonne pas
- Non et je ne comprends pas pourquoi

Q20. Avez-vous déjà constaté que les publicités sur les sites Internet évoluaient en fonction de votre historique de recherche et de navigation ?

- Oui et je comprends comment cela fonctionne
- Oui, mais j'ignore comment cela fonctionne
- Non, mais cela ne m'étonne pas
- Non et je ne comprends pas pourquoi

Q21. Quel baccalauréat avez-vous obtenu ?

- Un bac S
- Un bac STI2D
- Un bac ES ou L
- Un bac Technologique autre que STI2D
- Un bac Professionnel

Q21bis. Avez-vous suivi l'option ISN (Informatique et Sciences du Numérique) en Terminale ?

- Oui
- Non

Q22. Avez-vous déjà utilisé un logiciel de type AlgoBox lors de vos cours de maths au lycée ?

- Oui
- Non

Q23. Avez-vous déjà écrit des programmes informatiques (par exemple sur une calculatrice scientifique, en cours) ?

- Oui, j'en fais souvent
- Oui, j'ai déjà essayé
- Non, je n'en ai jamais fait
- Non, et je ne vois pas de quoi il s'agit

Q24. Avez-vous déjà entendu parler d'« algorithme » ?

- Oui et je sais ce que c'est
- Oui, mais je ne sais pas de quoi il s'agit
- Non et je ne sais pas de quoi il s'agit

Q25. Une association de lycéens a demandé la publication de l'algorithme d'APB, jugez-vous cette publication nécessaire ?

- Oui, tout à fait
- Oui, un peu
- Non, pas vraiment
- Non, pas du tout

Équité d'APB

Les questions suivantes ont pour thème votre représentation de l'équité d'APB afin de voir si vous trouvez équitable la sélection dans l'enseignement supérieur.

Q26. Considérez-vous que l'orientation dans l'enseignement supérieur par l'intermédiaire d'APB se fait de manière équitable, sur une échelle de 0 à 10 ?

0 signifie que vous trouvez le fonctionnement d'APB totalement inéquitable et 10, au contraire, très équitable.

Q27. Avez-vous entendu parler de personnes ayant bénéficié de procédures particulières (par l'intermédiaire de leurs connaissances personnelles par exemple) pour s'inscrire dans la formation qu'elles souhaitaient ?

- Oui et je pense que c'est courant
- Oui, mais je ne pense pas que cela soit courant
- Non, mais je pense que c'est courant
- Non et je ne pense pas que cela soit courant

Q28. Avez-vous entendu parler du recours au tirage aléatoire de certaines formations pour sélectionner leurs étudiants ?

- Oui et je pense que c'est courant
- Oui, mais je ne pense pas que cela soit courant
- Non, mais je pense que c'est courant
- Non et je ne pense pas que cela soit courant

Q29. Connaissez-vous quelqu'un qui n'aurait pas dû d'intégrer une formation selon vous et qui l'a tout de même obtenue ?

- Oui et je pense que c'est courant
- Oui, mais je ne pense pas que cela soit courant
- Non, mais je pense que c'est courant
- Non et je ne pense pas que cela soit courant

Q30. Saviez-vous qu'il était possible d'effectuer un recours en justice pour contester les résultats APB ?

- Oui
- Non

Q30bis. Avez-vous pensé à faire un recours en justice ?

- Oui, j'en ai fait un
- Oui, mais je n'en ai pas fait un
- Non, j'étais satisfait de mes résultats

Ce questionnaire est presque terminé, pour finir voici quelques questions sur vous.

Q31. Êtes-vous ?

- Une femme
- Un homme

Q32. Quel est le plus haut niveau de diplôme de votre mère ?

- Vous ne connaissez pas votre mère
- Vous ne savez pas
- Aucun diplôme
- Diplôme inférieur au Bac (CEP, BEPC, CAP, BEP...)
- Baccalauréat général (philo, sciences, A, B, C, D, E,...)
- Baccalauréat technologique, professionnel (F, G, H, etc.), brevet professionnel ou de technicien, autre brevet (BEA, BEC, BEI, etc.), capacité en droit
- Diplôme universitaire de 1er cycle, BTS, DUT, diplôme des professions sociales ou de la santé
- Diplôme universitaire de 2e cycle (licence ou maîtrise)
- Diplôme de 3e cycle (y compris médecine, pharmacie, dentaire), diplôme d'ingénieur, d'une grande école

Q33. Quelle est (ou était) la catégorie socioprofessionnelle de votre mère ?

- Vous ne savez pas
- Elle n'a jamais exercé de profession
- Ouvrier
- Employé
- Cadre moyen, technicien, instituteur, infirmier
- Cadre supérieur, ingénieur, professeur, chef d'entreprise

Q34. Quel est le plus haut niveau de diplôme de votre père ?

- Vous ne connaissez pas votre père
- Vous ne savez pas
- Aucun diplôme
- Diplôme inférieur au Bac (CEP, BEPC, CAP, BEP...)
- Baccalauréat général (philo, sciences, A, B, C, D, E,...)
- Baccalauréat technologique, professionnel (F, G, H, etc.), brevet professionnel ou de technicien, autre brevet (BEA, BEC, BEI, etc.), capacité en droit
- Diplôme universitaire de 1er cycle, BTS, DUT, diplôme des professions sociales ou de la santé
- Diplôme universitaire de 2e cycle (licence ou maîtrise)
- Diplôme de 3e cycle (y compris médecine, pharmacie, dentaire), diplôme d'ingénieur, d'une grande école

Q35. Quelle est (ou était) la catégorie socioprofessionnelle de votre père ?

- Vous ne savez pas
- Il n'a jamais exercé de profession
- Ouvrier
- Employé
- Cadre moyen, technicien, instituteur, infirmier
- Cadre supérieur, ingénieur, professeur, chef d'entreprise

Ce questionnaire est fini, nous vous remercions vivement pour votre participation.

Vos réponses permettront aux chercheurs de mieux saisir comment les bacheliers utilisent APB et ce qu'ils pensent de cet outil.

Avec nos sincères remerciements.

L'équipe de chercheurs d'AlgoCit