

numéro

11

1024

B U L L E T I N

de la société informatique
de France

septembre

2017

COMITÉ DE RÉDACTION

SYLVIE ALAYRANGUES
Université de Poitiers

OLIVIER BAUDON
Université de Bordeaux

COLIN DE LA HIGUERA
Université de Nantes

JEAN-PAUL DELAHAYE
Université Lille 1

OLIVIA DUMAS
Académie de Caen

CHRISTINE FROIDEVAUX
Université Paris-Sud

THIERRY GARCIA
*Université de Versailles Saint-Quentin-
en-Yvelines*

MARIE-CLAUDE GAUDEL
Université Paris-Sud

CLAUDIA MARINICA
Université de Cergy-Pontoise

PHILIPPE MARQUET
Université Lille 1

EUNIKA MERCIER-LAURENT
Innovation 3D et Université Lyon 3

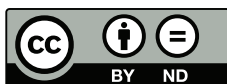
DENIS PALLEZ
Université Nice Sophia Antipolis

JEAN-MARC PETIT
INSA Lyon

VINCENT RIBAUD
Université de Bretagne Occidentale

ÉRIC SOPENA, *Université de Bordeaux*, rédacteur en chef

Contact : 1024@societe-informatique-de-france.fr



*Cette œuvre est mise à disposition sous licence Attribution - Pas de Modification 4.0 France.
Pour voir une copie de cette licence, visitez*

<http://creativecommons.org/licenses/by-nd/4.0/deed.fr>

*ou écrivez à Creative Commons, 444 Castro Street, Suite 900, Mountain View, California,
94041, USA.*

SOCIÉTÉ INFORMATIQUE DE FRANCE
Institut Henri Poincaré, 11 rue Pierre et Marie Curie, 75231 Paris Cedex 05

Prix public : 32 € (adhérents SIF : -30%)

Directeur de la publication : Jean-Marc Petit

ISSN : 2270-1419

Couverture : d'après une maquette réalisée par Lollygraph.com.

SOMMAIRE DU N° 11



SIF

Mot du président, <i>Jean-Marc Petit</i>	3
Actualité de la SIF, <i>Sylvie Alayrangues</i>	7

DOSSIER : CYBERSÉCURITÉ

CyberEdu, ou l'ambition de former à la sécurité tous les acteurs du numérique, <i>Olivier Levillain</i>	19
Cybersécurité, <i>Gildas Avoine et Marc-Olivier Killijian</i>	25

ENTRETIENS

Le logiciel libre, de l'usage à la recherche, <i>Entretien avec Roberto Di Cosmo</i>	39
--	----

ENSEIGNEMENT

Défendre une éducation qui permet de décoder le numérique, <i>Collectif</i> ..	53
Un enseignement de l'informatique de culture générale pour tous les élèves – Pourquoi et comment, <i>Jean-Pierre Archambault</i>	61
L'informatique en CPGE, <i>Judicaël Courant</i>	77

HOMMAGE

Décès d'Alain Colmerauer, pionnier de l'intelligence artificielle, <i>Frédéric Benhamou</i>	97
Daniel Kayser, un enseignant-chercheur au cœur de l'IA, <i>François Lévy et Adeline Nazarenko</i>	101
In memoriam Roger Mohr, <i>K. Tombre, L. Quan, R. Horaud, P. Gros, C. Schmid, P. Sturm</i>	107

HISTOIRE

Le séminaire de Sèvres (mars 1970), *Jacques Baudé* 115

PRIX ET DISTINCTIONS

Pour tout le monde : Tim Berners-Lee, lauréat du prix Turing 2016 pour avoir inventé... le Web, *Fabien Gandon* 129

RÉCRÉATION

Les chapeaux de l'icosaèdre, *Jean-Paul Delahaye* 155



Mot du président

Jean-Marc Petit¹

Alors que ce numéro de 1024 était quasiment bouclé, nous avons appris le décès de Maurice Nivat. Maurice était un supporter inconditionnel de la SiF. Père de l'informatique théorique en France, en Europe et dans le monde, il a été un acteur et un témoin actif de l'émergence de cette discipline. En tant que président de la SiF, j'ai eu l'occasion de le côtoyer ces dernières années et d'échan-



ger sur l'enseignement de l'informatique, un sujet qui le passionnait. Il était particulièrement conscient et convaincu que l'informatique ne pouvait se contenter de rester théorique et qu'elle devait *se mettre en musique* en construisant des logiciels, ces artefacts immatériels qui n'en finissent pas de changer notre monde. Cette dimension pratique était tout aussi importante à ses yeux, lui le théoricien par excellence. Il ne laissait personne indifférent, lui qui savait mieux que quiconque avoir le verbe

1. Président de la Société informatique de France, professeur à l'INSA de Lyon, LIRIS (UMR 5205 CNRS), jean-marc.petit@insa-lyon.fr, twitter : @calixtepetit.

haut. Sa plume était incisive, aiguisée comme une lame de rasoir. Que l'on soit d'accord ou non avec ses mots, ils suscitaient des questionnements et ne restaient jamais neutres. Sa fougue, son énergie, sa passion, ses fulgurances, son intelligence étaient restées intactes, simplement remarquables. La cohésion de la SiF pour la promotion de l'informatique était l'argument ultime lorsque des divergences surgissaient. Enfin, Maurice était profondément humain et aimait la vie. Simplement. Il va manquer à notre association qui perd un de ses plus fidèles amis. Nos pensées vont à sa famille, son épouse Paule et ses enfants, petits-enfants et proches à qui nous adressons toutes nos condoléances.

o O o

L'informatique comme ascenseur social ?

Alors que l'offre de formations en informatique dans l'enseignement supérieur public atteint une certaine stabilité, nous assistons à une explosion de l'offre de formations issues du monde industriel et/ou initiées par la puissance publique, en dehors des ministères de l'enseignement scolaire ou supérieur. Sont visés les néo-bacheliers pour les écoles privées d'ingénieurs mais aussi les décrocheurs du système scolaire ainsi que les personnes sans emploi pour les initiatives locales soutenues par l'État, par exemple avec la *grande école du numérique* qui a permis l'ouverture de nombreuses actions de formation en région. D'autres initiatives privées ne sont pas en reste non plus, dans la lignée de l'École 42 qui essaime, en France et dans le monde, de nouveaux types de formation.

Il faut dire que les entreprises ont besoin de personnel qualifié pour les nouveaux métiers de l'économie du numérique que les formations classiques en informatique ne suffisent plus à former.

Pour les néo-bacheliers, ces écoles privées offrent des débouchés via des modalités pédagogiques centrées sur les projets et un corps professoral réduit à sa portion congrue. C'est le prix de leur vélocité et de leur capacité d'adaptation au « marché ». Pour les décrocheurs, les initiatives pour apprendre le code leur permettent souvent de décrocher un premier emploi. Ce n'est pas rien. L'informatique n'y est plus vue comme une sous-discipline de telle ou telle discipline de notre système scolaire, mais comme la discipline qui ouvre des opportunités d'emploi.

Certain·e·s estiment que ces formations, parfois très courtes, rendront ces personnes inopérantes sur le marché du travail quelques années après leur premier emploi, en raison de l'évolution rapide des technologies et du manque de recul sur les fondements de l'informatique. Sans nier cet effet de bord, on peut constater que de nombreuses personnes formées par ces canaux montrent des capacités d'évolution et d'adaptation dans leur carrière, années après années. Les exemples autour de nous ne manquent pas. Le domaine du numérique est suffisamment vaste et actif pour leur

permettre de trouver leur place. Que ce ne soit pas « académique » n'est au fond pas bien grave.

Nous touchons là un des paradoxes de notre discipline : elle vise à la fois des profils passés par les processus très sélectifs de notre système éducatif et des profils plus atypiques, allant des autodidactes à celles et ceux qui ne s'épanouissent pas ou mal dans notre système scolaire, voire qui en sont sorti-e-s prématurément. En filigrane, on serait tenté de dire que l'influence du milieu social des élèves serait moins forte pour étudier l'informatique que pour les autres disciplines. L'informatique permettrait donc de mobiliser des compétences restées muettes au sein des disciplines classiques du collège et du lycée. Intégrer l'erreur comme une façon normale de progresser, ne pas être évalué directement par un-e professeur-e mais par la médiation d'une machine, expérimenter le triptyque *essayer, se tromper, et recommencer* sans faire face à un jugement inquisiteur, pouvoir travailler à la maison, avec des ami-e-s, faire des projets... tout cela serait au fond bénéfique pour une partie des élèves !

Au fond, *l'informatique rendrait possible l'ascenseur social* en bouleversant les codes de notre système scolaire. Chacun-e à sa vitesse, chacun-e à son niveau, être utile dans notre société et y apporter sa contribution serait possible. Là encore, ce n'est pas rien.

o O o

Des professeur-e-s d'informatique correctement formé-e-s pour les élèves.

Sur l'enseignement scolaire maintenant, après les avancées effectives sur les programmes pour l'enseignement de l'informatique, notre *angoisse collective*² concerne les professeur-e-s en charge de ces enseignements, qui risquent de se retrouver rapidement devant des élèves (du collège aux classes préparatoires) qui, pour un certain nombre, seront plus à l'aise qu'eux-mêmes.

Faire entrer la France dans l'ère du numérique sans professeur-e-s spécifiquement formé-e-s pour éduquer nos enfants à l'informatique est un non-sens. En effet, comme nous le répétons depuis longtemps, la question de la formation des professeur-e-s est restée béante, sans traitement de fond par les gouvernements précédents. Le CAPES de mathématiques option informatique est une initiative intéressante mais qui aura du mal à répondre aux besoins de la France. Les échos sur les formations à ce CAPES qui avaient ouvert l'année dernière font état de tout petits chiffres pour cette nouvelle année, avec le risque que certaines n'ouvrent pas. Il est vrai que demander aux candidat-e-s un niveau M1 en mathématiques est trop exigeant et ne peut être attractif pour la plupart des étudiant-e-s en informatique. Si

2. Naturellement, l'actualité regorge de sujets plus anxiogènes encore.

une telle barre en mathématiques était demandée pour les autres CAPES (physique-chimie par exemple), les mêmes causes produiraient les mêmes effets. Clairement, ce CAPES est réservé aux mathématicien-ne-s qui ont fait aussi de l'informatique, alors que l'on sait que les flux d'étudiant-e-s en Licence de mathématiques sont faibles.

Pourtant des solutions existent. Elles doivent remplir au moins deux conditions pour être réalistes : la première est de s'assurer du vivier en analysant par exemple les flux d'étudiant-e-s en Licence. La seconde est d'identifier dans ce vivier les profils sociaux des futur-e-s candidat-e-s qui pourraient être attiré-e-s par le métier d'enseignant-e. Une solution simple que la nouvelle équipe gouvernementale pourrait mettre en œuvre serait de créer un CAPES et une Agrégation d'informatique. C'est une hypothèse de travail simple et frappée au coin du bon sens. C'est une idée ni de droite, ni de gauche qui remplirait les deux conditions ci-dessus pour le CAPES : les licences d'informatique des universités françaises sont attractives et elles accueillent de nombreux étudiant.e.s aux profils variés qui pourraient trouver dans les métiers de l'enseignement une forme de promotion sociale.

Nul doute que les départements d'informatique des universités de France se mobiliseraient pour former des candidat-e-s à ce concours.

Aux femmes et aux hommes politiques de notre pays de prendre leurs responsabilités.

o O o

Rejoignez nous !

En ce début d'automne, si vous souhaitez rejoindre le CA de la SIF en 2018 pour un mandat de trois ans afin d'agir à nos côtés pour renforcer les actions de promotion de l'informatique, c'est le moment ! Nous vous invitons à vous rapprocher de nous et à déposer votre candidature avant le 15 novembre 2017.

o O o

Bonne lecture de ce nouveau numéro de 1024 qui devrait, une fois encore, permettre de mesurer la profondeur et la largeur de notre discipline.



Actualité de la SIF

Informations collectées par Sylvie Alayrangues

Une rubrique pour vous tenir informé-e de la vie de la SIF et vous inviter à y participer. Vous trouverez notamment ici des annonces de manifestations organisées ou soutenues par la SIF, un point sur les dossiers en cours, des appels à participation...

La SIF vous parle d'elle

Assemblée générale extraordinaire du 19 juin

Le 19 juin dernier, la SIF convoquait une assemblée générale extraordinaire pour voter une modification des statuts de notre association dont le vote n'avait finalement pas pu être validé lors de la précédente AG faute du quorum requis. Ce dernier portait en effet sur les personnes physiquement présentes et n'incluait pas les votes par procuration.

Les quatre questions posées ont reçu une réponse positive à l'unanimité des votants.

Pour mémoire, les questions mises au vote étaient les suivantes :

- Approuvez-vous le passage de la SIF en ARUP ?
- Approuvez-vous les modifications apportées aux statuts ?
- Approuvez-vous les modifications apportées au règlement intérieur ?
- Donnez-vous mandat à Élisabeth Muriasco et Jean-Marc Petit pour modifier les statuts et règlement intérieur à la demande du ministère de l'intérieur pour l'éligibilité au statut d'ARUP ?

Campagne d'adhésion

Cette année, la SIF a lancé une vaste campagne d'adhésion auprès des *adhérents institutionnels* (ou personnes morales) en ciblant tout d'abord ceux qui ont été adhérents et qui n'ont pas renouvelé, puis de nouveaux prospects. Pour cette campagne, notre cœur d'action est resté centré sur le monde académique – laboratoire de recherche, structure d'enseignement (IUT, UFR, école d'ingénieur) – partenaires historiques de la SIF.

Bien sûr, nous nous tournons également vers le monde industriel, avec notamment l'adhésion de Criteo¹ comme personne morale.

À court terme, notre objectif est de continuer notre croissance, asseoir notre organisation côté académique et industriel pour peser davantage en terme de nombre d'adhérent-e-s et atteindre un budget récurrent qui donne des marges de manœuvre. Cela nous permettrait de renforcer nos actions mais aussi d'en engager de nouvelles. Par exemple, nous n'avons toujours ni locaux ni personnels pour le soutien administratif et la communication.

Nous comptons naturellement sur votre soutien pour nous aider à promouvoir concrètement la SIF via l'adhésion des personnes morales auxquelles vous appartenez. Le préalable est de ne pas se laisser enfermer dans une vision étriquée de l'informatique. L'informatique que nous défendons va des systèmes d'information à l'intelligence artificielle en passant par la robotique ou la bio-informatique, elle est à la fois une science et une technique, théorique et pratique. Harmonieusement.

Du côté du Conseil d'administration

Comme chaque année, le CA de la SIF sera partiellement renouvelé lors de l'assemblée générale de l'association qui aura lieu pendant son congrès annuel.

Le calendrier est le suivant :

- contact au fil de l'eau avec Jean-Marc Petit, président de la SIF
- envoi des candidatures avant le 15 novembre à Olivier Baudon, secrétaire de la SIF, secretariat@societe-informatique-de-france.fr,
- vote en AG et proclamation des résultats le 1^{er} février 2018.

Toute candidature doit comprendre une profession de foi qui donne un sens à l'engagement que vous souhaitez prendre. Vous pourrez y indiquer sur quel(s) sujet(s) vous souhaitez vous investir. Et les sujets ne manquent pas : de l'enseignement de la science informatique en primaire/collège/lycée/université à la promotion de la recherche en informatique, en passant par la médiation scientifique, les liens avec l'industrie, l'innovation, l'international. Vous pourrez éventuellement aussi préciser quel rôle vous seriez prêt à endosser pour aider au bon fonctionnement de l'association : secrétariat, communication, etc.

1. www.criteo.fr

Les relevés de décisions des dernières réunions du CA sont toujours accessibles en ligne².

Comme chaque année la réunion du mois de juillet s'est déroulée « au vert » sur deux jours. Elle a notamment été l'occasion d'accueillir des invité.e-s et d'échanger avec eux-elles sur des sujets clés pour la SIF : *politique internationale de la SIF* avec Valérie Issarny (Inria), membre nouvellement élue à ACM Europe, *enseignement des sciences* avec François Moussavou (APMEP³ et IREM⁴ de Marseille), Mickaël Prost (UPS⁵) et Bernard Julia (SFP⁶), *Recherche sur l'enseignement de l'informatique* avec Yves Bertrand (président de la CDUS⁷), Stéphane Vinatier et Fabrice Vandebrouck (respectivement président et vice-président de l'ADIREM⁸).

Du côté du Conseil scientifique

Les comptes-rendus sont accessibles en ligne⁹.

La dernière rencontre du conseil scientifique a eu lieu le 9 juin dernier. Elle accueillait trois invités, Jean-Pierre Archambault, membre du conseil d'administration de la SIF et de l'EPI¹⁰, Marc de Falco, professeur agrégé de mathématiques, docteur en informatique, enseignant en mathématiques et informatique en CPGE MPSI¹¹ à Valbonne, représentant l'UPS et Pierre Saurel, agrégé de mathématiques, docteur en sciences cognitives, maître de conférences à l'université Paris-Sorbonne, habilité à diriger des recherches en intelligence artificielle, enseignant en informatique et mathématiques à l'ESPE¹² de Paris, spécialiste en droit de l'informatique.

Pierre Saurel était présent pour présenter son point de vue¹³ et discuter autour de deux sujets chers à la SIF : le CAPES bivalent mathématiques option informatique et l'introduction de l'informatique dans les programmes en école élémentaire.

2. <http://www.societe-informatique-de-france.fr/la-sif/ca-2017/>

3. Association des professeurs de mathématiques de l'enseignement public.

4. Institut de recherche sur l'enseignement des mathématiques.

5. Union des professeurs de classes préparatoires.

6. Société française de physique.

7. Conférence des doyens et directeurs des UFR scientifiques.

8. Assemblée des directeurs d'IREM.

9. <http://www.societe-informatique-de-france.fr/la-sif/conseil-scientifique-de-la-societe-informatique-de-france/>

10. Association enseignement public & informatique.

11. Classes préparatoires aux grandes écoles, mathématiques, physique et sciences de l'ingénieur.

12. École supérieure du professorat et de l'éducation.

13. Les transparents sont accessibles sur http://www.societe-informatique-de-france.fr/wp-content/uploads/2017/07/SAUREL_PIERRE.pdf

La place grandissante que prennent les outils développés par les GAFAM¹⁴ dans l'éducation de nos enfants a été largement débattue puisqu'après un accord controversé en 2015 entre le ministère et Microsoft¹⁵, c'est maintenant la Direction du numérique pour l'éducation (DNE) qui pousse les établissements à utiliser les services des GAFAM¹⁶. Au-delà de cette question des GAFAM, le conseil scientifique insiste sur les conditions qui doivent, a minima, être respectées pour tout logiciel utilisé dans l'Éducation nationale et distingue le développement des outils de l'exploitation des données. Un communiqué présentant sa position et ses propositions est accessible en ligne¹⁷.

Du côté du Conseil des associations

La prochaine réunion du conseil des associations se déroulera sur Paris le 11 octobre prochain. Outre un bilan de l'année écoulée, et une discussion des projets à venir, elle sera l'occasion de créer du lien entre ce conseil et le conseil scientifique. Il s'agira aussi de voir comment travailler ensemble autour d'un enjeu commun : la médiation scientifique en informatique.

Il est possible de s'inscrire en ligne¹⁸.

À l'international

Le comité de pilotage d'IFIP-France s'est réuni le 28 juin dernier. Florence Sèdes a été choisie pour succéder à Bernard Cornu en tant que représentante de la France à l'IFIP, quand celui-ci quittera ses fonctions fin 2018. Elle est d'ores et déjà associée au comité de pilotage.

Le même jour deux autres réunions étaient organisées, la première rassemblait les représentants français dans les différents TC (*Technical Committees*). Elle a permis de discuter du rôle de ces représentants et d'envisager comment améliorer la communication auprès de la communauté informatique et les échanges autour des travaux menés dans le cadre de l'IFIP. Une réunion générale des membres français clôturait cette journée. Après un bref historique relatant la mise en place du consortium IFIP France, quelques points d'actualités autour de l'IFIP ont été évoqués. Les informations relatives à l'IFIP sont disponibles sur le site <http://www.ifip.org/>. Il est à noter que l'IFIP met en place une « bibliothèque électronique ouverte », accessible via le portail <https://hal.inria.fr/IFIP>. Elle propose notamment les actes des événements IFIP ainsi que des ouvrages publiés par l'IFIP. Les différents TC ont

14. Google, Apple, Facebook, Amazon, Microsoft.

15. Voir la réaction de la SIF à ce sujet : <http://www.societe-informatique-de-france.fr/2015/12/communique-partenariat-microsoft-education-nationale/>

16. Incitation à laquelle a réagi l'EPI : <http://www.enseignerlinformatique.org/2017/05/25/communique-de-lassociation-enseignement-public-et-informatique-epi/>

17. <http://www.societe-informatique-de-france.fr/wp-content/uploads/2017/07/donnees-pedagogiques.pdf>

18. <https://framadata.org/eHN7z91ji7aIVN9w>

ensuite été présentés et des pistes ont été envisagées pour renforcer la participation de la France à l'IFIP et accroître la visibilité de l'IFIP en France. Le compte-rendu des trois réunions est accessible sur le site de la SIF¹⁹.

La SIF vous propose

Congrès 2018 : l'informatique au carrefour des sciences

Le congrès 2018, « l'informatique au carrefour des sciences », se construit en partenariat avec le Palais de la découverte. Il se déroulera les 31 janvier et 1^{er} février 2018 à Paris.

L'informatique, science au cœur du numérique, bouleverse tous les pans de la société et notamment les sciences dites « traditionnelles ». Comment la science informatique et les autres sciences s'enrichissent-elles mutuellement et contribuent-elles chacune à l'émergence de nouvelles idées, de nouveaux concepts ? Comment engendrent-elles, ensemble, de nouvelles disciplines : bioinformatique, astroinformatique... ? Comment travaillent-elles de concert pour modeler l'avenir de l'humain (agronomie, biologie, médecine, robotique...) ? Comment, enfin, la science informatique renouvelle-t-elle également le visage des sciences humaines ?

Voici quelques-unes des questions que nous vous proposons d'explorer lors de ce congrès, en compagnie de chercheur·euse·s dont les travaux s'effectuent à l'interface de plusieurs sciences. Vous y rencontrerez des informaticien·ne·s ainsi que les collègues des autres disciplines avec lesquelles il·le·s travaillent au quotidien.

Temps fort de ce congrès, dédié à la mémoire de Maurice Nivat, une intervention de Michel Serres aura lieu le 31 janvier au soir.

Journée « Casser les codes – Femmes, genre et informatique »

Cette journée, organisée le 16 octobre à Paris, explorera les enjeux passés, présents et futurs de la participation des femmes à l'informatique. Cette proposition vous est faite conjointement par la Société informatique de France, Inria et l'Institut des sciences de la communication (CNRS/Paris-Sorbonne/UPMC). Ensemble nous croiserons approches académiques et retours d'expériences professionnelles.

La journée s'ouvrira par une allocution de Marlène Schiappa, Secrétaire d'État chargée de l'Égalité entre les femmes et les hommes. Puis après un mot d'accueil des différentes partenaires, nous entrerons dans le vif du sujet avec un point de vue historique puis prospectif sur l'évolution de la présence féminine dans l'informatique. Une table ronde permettra ensuite d'échanger autour de trajectoires d'informaticiennes dans l'enseignement supérieur.

19. <http://www.societe-informatique-de-france.fr/wp-content/uploads/2017/09/170628-IFIPFranceRe%CC%81unions-CR.pdf>

Après la pause méridienne, nous nous intéresserons à la place des femmes dans la sphère plus large du « numérique ». Puis nous discuterons des moyens à mettre en œuvre pour permettre aux jeunes de se construire une image de la science informatique plus conforme à la réalité.

L'inscription est gratuite mais obligatoire avant le 12 octobre. Le programme détaillé est accessible en ligne²⁰.

Jam Session : « Fausses informations, post vérité : allons aux faits ! »

Cette Jam Session se déroulera le 25 octobre 2017 de 18h à 20h dans la salle de conférence du Musée des arts et métiers.

Elle est co-organisée par la SIF et Inria Alumni et s'inscrit dans la suite du Colloque de l'Académie des sciences « Les enjeux scientifiques de l'éthique du numérique » du 23 mai et de la Journée organisée par la SiF et la CERNA sur l'éthique des algorithmes, à Telecom ParisTech le 19 juin.

« Fausses informations, post vérité : allons aux faits ! » : ces questions de société, réactivées par l'actualité politique, amplifiées par le développement du numérique, sont aussi au cœur de recherches scientifiques et techniques de pointe, où éthique et numérique s'interpellent.

Les intervenants – scientifique, journalistes, historien – apporteront différents points de vue pour tenter d'éclairer le débat : Ioana Manolescu (Inria Saclay), Nicolas Kayser-Bril (datajournalisme, agence Journalism ++), Benjamin Thierry (université Paris-Sorbonne) et Nicolas Vanderbiest (université catholique de Louvain et Reputatio Lab).

Toutes les informations sont accessibles sur notre site²¹.

Concours de vidéos : #ScienceInfoStream

La SIF lance son premier concours vidéo : « L'informatique en stream », #ScienceInfoStream.

Vous êtes passionné-e-s de science informatique ? Vous aimez les défis ? Vous vous sentez l'âme d'un-e vidéaste ? N'attendez plus ! Prenez votre smartphone, webcam, ou caméra : vous avez moins de 2⁸ secondes pour partager ces pépites de science qui vous fascinent.

Plus précisément, votre vidéo, tournée en français, aura une durée comprise entre 2⁷ (128) et 2⁸ (256) secondes. Elle sera au format de votre choix (les formats libres, comme ogv, sont encouragés) et mise à disposition sur le web (via votre page personnelle ou via un site de streaming ou...). Une déclaration d'intention se fera par courrier électronique envoyé à l'adresse

20. <http://www.societe-informatique-de-france.fr/2017/07/casser-les-codes-femmes-genre-et-informatique/>

21. <http://www.societe-informatique-de-france.fr/les-journees-sif/session-fausses-informations-post-verite-allons-aux-faits-paris-25-octobre-2017/>

scienceinfostream@societe-informatique-de-france.fr, avant le 1^{er} décembre 2017. La soumission finale est attendue au plus tard le 8 janvier. Le message contiendra, outre le lien, le titre de la vidéo et les noms et prénoms de son ou de ses auteur-e-s avec une présentation succincte (âge, rôles respectifs dans le projet...).

Un jury composé notamment d’informaticien-ne-s et de spécialistes de la médiation scientifique réalisera une première sélection de vidéos dont la diffusion et la promotion seront assurées par la SIF et ses partenaires. Trois d’entre elles seront ensuite choisies et recevront des récompenses allant de 500 à 1500 euros. Les vidéos seront également mises en ligne par la SIF, de manière ouverte avec une licence Creative Commons choisie par l’auteur-e. La remise des prix se déroulera au Palais de la découverte le 31 janvier 2018 lors du congrès de la SIF en présence de nombreux intervenants prestigieux.

Le jury appréciera particulièrement les vidéos qui satisfont un ou plusieurs des critères suivants :

- originalité du sujet ou du point de vue utilisé pour le traiter,
- aptitude à susciter une vocation pour la science informatique,
- sensibilisation aux enjeux sociétaux de la science informatique,
- capacité à expliquer en langage simple un point de science informatique.

Qui peut participer ?

Enfants, lycéen-ne-s, étudiant-e-s, doctorant-e-s, chercheur-euse-s juniors ou seniors, enseignant-e-s, ingénieur-e-s, programmeur-euse-s, ou découvreur-euse-s de l’informatique, etc. Bref, toute personne qui le souhaite, en individuel ou en équipe, sans limite d’âge.

N’hésitez pas à nous contacter si vous avez besoin de précisions. Une FAQ est disponible sur le site de la société informatique de France, que nous enrichirons au fur et à mesure de vos questions²².

Journée Sciences & Médias : « Comment lutter contre la désinformation scientifique »

Pour la deuxième fois, la SIF participe à l’organisation d’une journée Sciences & Médias²³, avec la société chimique de France, la société française de physique, la société française de statistique, la société de mathématiques appliquées et industrielles et la société mathématique de France. Pour cette édition, sont également partenaires l’association des journalistes scientifiques de la presse d’information et la bibliothèque nationale de France.

Cette édition se déroulera le 11 janvier 2018 dans le Grand auditorium de la Bibliothèque nationale de France, site François-Mitterrand. La précédente édition, en 2016, s’est interrogée sur les moyens à mettre en place pour parler de science aux

22. www.societe-informatique-de-france.fr/2017/07/concours-de-vidéos-scienceinfosatream/

23. <http://sciencesetmedia.org/>

jeunes. Dans une certaine continuité, l'édition 2018 s'oriente sur la question de l'information scientifique dans les médias et son utilisation. En effet, ces dernières années, l'essor des nouvelles technologies d'information et de communication a fortement changé notre façon de nous informer et de nous approprier l'information. Elle est devenue instantanée et prend des formes très diverses. Au sein de cette variété, de nombreux travers laissent place à la surinterprétation voire à la diffusion de « fausses vérités ». Lors de cette journée, il s'agira de mettre en perspective les points de vue de différent-e-s professionnel-le-s tel-le-s que des sociologues des sciences, des scientifiques, des journalistes, des blogueur-euse-s... afin de comprendre les enjeux de la diffusion de l'information, les problèmes rencontrés et d'imaginer les solutions qui pourraient être apportées.

Qui choisit les expert-e-s scientifiques et comment ? Quelles informations sont choisies et comment ? Comment rectifier une mauvaise information ? Comment préparer les scientifiques à l'échange avec les médias ? Voici les principales questions auxquelles nous tenterons d'apporter une réponse.

École d'été autour de la médiation scientifique

Les 6, 7 et 8 juin 2018, nous vous proposons de vous initier à la médiation scientifique en informatique lors d'une école d'été qui se déroulera dans la ville rose.

Si vous êtes chercheur-euse / enseignant-e-chercheur-euse, si vous participez ou souhaitez participer à la fête de la science, à des forums ou à des journées portes ouvertes où vous parlez d'informatique à des non initiés, venez apprendre à construire vos actions de médiation. Cette école sur trois jours vous propose de découvrir, imaginer et tester des actions de médiation en vue de monter un événement pour votre laboratoire (Fête de la science, portes ouvertes...). Un parcours progressif vous amènera de la découverte à la mise en pratique devant un « public test ».

Retour sur la journée sur l'éthique des algorithmes – 19 juin 2017

La SIF et la CERNA (Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistene) ont conjointement organisé une journée sur l'éthique des algorithmes. Intitulée « Une éthique des algorithmes : une exigence morale et un avantage concurrentiel », elle a eu lieu le 19 juin dernier dans les locaux de Telecom ParisTech.

Une centaine de personnes étaient présentes lors de cette journée qui rentrait dans le cadre de la consultation nationale coordonnée par la CNIL « *Éthique et numérique : les algorithmes en débat*²⁴ ».

Les questionnements éthiques autour du numérique diffusent dans toutes les composantes de notre société. Lorsqu'il s'agit de choisir entre plusieurs algorithmes ou lorsqu'il s'agit de développer de tels algorithmes, les qualités éthiques d'un système peuvent devenir des arguments importants. En particulier pour la confiance qu'elles

24. <https://www.cnil.fr/fr/ethique-et-numerique-les-algorithmes-en-debat-0>

suscitent mais aussi les coûts directs ou indirects qu'elles évitent – comme la fuite de données ou d'informations.

L'objectif de cette journée était de débattre des questionnements éthiques autour du numérique, questionnements qui émergent dans toutes les composantes de la société. À travers quelques exposés et trois tables rondes, elle a ainsi permis d'aborder l'impact sur les usagers et le point de vue des industriels, les problématiques nouvelles introduites par l'apprentissage machine, ainsi que les algorithmes éthiques dans l'éducation.

Class'Code

Il est difficile de résumer en quelques lignes les multiples avancées de Class'Code, formation hybride mêlant cours en ligne et rencontres sur le terrain pour « se former pour initier les jeunes à la pensée informatique ».

Les cinq modules de formation sont en ligne et utilisés. Il existe maintenant deux livres « 1,2,3, codez », l'un pour le primaire, l'autre pour le collège, proposant des séquences pédagogiques directement utilisables. Un troisième pour le lycée est en préparation. L'intégration dans des formations existantes est amorcée (notamment dans la plateforme de formation continue de l'éducation nationale M@gistère).

Une nouvelle version de la plateforme <http://classcode.fr/> sera bientôt en ligne : elle tient compte des retours utilisateurs et présentera une meilleure ergonomie avec des ressources bien modularisées pour faciliter leur intégration dans d'autres parcours et les rendre plus aisément accessibles.

Le déploiement sur le territoire s'est intensifié. Au-delà des deux régions initialement partenaires du projet, sept autres régions se sont maintenant également dotées d'une équipe de coordination et dans cinq autres, des actions Class'Code ont commencé à émerger. L'offre de formation a également été intégrée au plan académique de formation de cinq académies et devrait prochainement être insérée dans six autres. Elle commence aussi à être utilisée dans des ESPEs, et est déployée nationalement par le réseau Canopé. Sa diffusion s'appuie également sur des associations d'éducation populaire. Depuis peu, des vidéos Class'Code sont aussi disponibles sur francetvéducation²⁵.

Class'Code prend également une dimension internationale via l'attribution d'une chaire UNESCO à l'université de Nantes, chaire portée par Colin de la Higuera, initiateur du projet Class'Code. Cette chaire se place dans la continuité de Class'Code, et va travailler autour de modèles de formation hybride s'appuyant sur des ressources éducatives libres.

Pour que Class'Code franchisse un nouveau pallier, il s'agit maintenant qu'au-delà de rencontres ponctuelles d'échanges en petits groupes, se forment des communautés d'entraides plus larges. En outre, Class'Code utilise le code comme un point d'entrée à la science informatique, et non pas comme une fin en soi. Convaincre de

25. <http://education.francetv.fr/recherche?q=Class'Code>

l'importance pour tout-e citoyen-ne de se doter d'une culture scientifique et technique plus large est également un des axes de travail du projet.

La SIF soutient

Le concours Castor

La prochaine édition du concours Castor informatique²⁶, qui s'adresse aux classes du CM1 à la terminale, se déroulera de 12 au 25 novembre 2017.

Les inscriptions des enseignant-e-s sont d'ores et déjà ouvertes.

La SIF prend position et agit

Pour mémoire, la SIF vous propose deux revues de presse, l'une sur « l'enseignement de l'informatique : primaire et secondaire »²⁷, l'autre sur le thème « femmes et informatique »²⁸.

En enseignement

Le communiqué²⁹ de la SIF sur les données pédagogiques porté par le conseil scientifique publié courant juin a déjà été évoqué.

Suite à la mise en place du gouvernement d'Édouard Philippe sous l'autorité du président de la République Emmanuel Macron, la SIF a contacté plusieurs membres du gouvernement en insistant sur l'urgence du dossier informatique. Concernant l'enseignement, la SIF a été reçue au ministère de l'Éducation nationale par Isabelle Bourhis et David Knecht, membres du cabinet de Michel Blanquer. La DGESCO (Direction générale de l'enseignement scolaire) était également représentée à cette rencontre en la personne de Jean-Yves Capul, directeur du Service du développement du numérique éducatif. La SIF a, à nouveau, insisté sur la nécessité de formation des enseignants à l'informatique pour qu'ils soient en mesure de l'enseigner.

La SIF poursuit également des discussions avec les inspecteurs généraux de différentes disciplines pour discuter des articulations possibles avec l'enseignement de l'informatique.

Le groupe de travail pluridisciplinaire sur l'enseignement des sciences au lycée continue également de travailler. Les travaux de ce groupe nourrissent également ceux du groupe de réflexion sur l'informatique au lycée qui réfléchit au bagage informatique nécessaire pour tout-e bachelier-lière.

26. <http://castor-informatique.fr/>

27. <http://www.societe-informatique-de-france.fr/revue-presse-ens-info/>

28. <http://www.societe-informatique-de-france.fr/revue-de-presse-femmes-info/>

29. <http://www.societe-informatique-de-france.fr/wp-content/uploads/2017/07/donnees-pedagogiques.pdf>

Enfin la SIF se concentre aussi sur la question de la recherche en enseignement de l'informatique et se rapproche pour ce faire des instituts de recherche en mathématiques qui ont une forte expérience du côté des mathématiques mais accueillent déjà également des groupes de travail autour d'autres sciences, dont l'informatique.

En recherche et innovation

Dans le cadre des entretiens annuels avec l'INS2I et Inria, une délégation de la SIF a rencontré le 22 mai dernier Michel Bidoit, directeur du département INS2I du CNRS, et Jean Mairesse qui en était à l'époque directeur adjoint. Le même jour, une délégation de la SIF rencontrait également Antoine Petit, PDG d'Inria.

Une rencontre avec le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation est prévue dans les prochaines semaines.

En médiation scientifique

La fondation Blaise Pascal vise à financer des actions de médiation à l'informatique et aux mathématiques. Elle a accordé une première série de subventions à des projets d'origines diverses suite à l'appel à projets paru au mois de juin. D'autres appels auront lieu dans les prochains mois. Si vous avez envie de porter des actions qui entrent dans le cadre des missions de la fondation³⁰, n'hésitez pas à y répondre.

Sur les enjeux de société

La SIF a été reçue au secrétariat d'état au numérique par Côme Berbain, conseiller transformation numérique de l'État et sécurité numérique. Des projets de collaboration, notamment autour de la création d'une maison du numérique, sont envisagés.

La SIF poursuit sa participation au collectif CNIL/EduNum³¹. Dans ce cadre, il a été récemment décidé de mettre en place des groupes de travail reprenant les thématiques développées dans le courrier adressé aux candidat-e-s à l'élection présidentielle « Pour une éducation populaire au numérique », afin d'avancer concrètement sur chaque sujet.

Quatre groupes ont ainsi été créés par le collectif :

- GT 1 : « Numérique et personnes vulnérables »,
- GT 2 : « Accompagner les parents »,
- GT 3 : « Promouvoir une culture citoyenne du numérique auprès des jeunes »,
- GT 4 : « Développer le droit à la formation au numérique tout au long de la vie ».

La SIF émerge concrètement aux groupes de travail 3 et 4.

30. <http://fondationblaisepascal.strikingly.com/>

31. <https://www.educnum.fr/>



CyberEdu, ou l'ambition de former à la sécurité tous les acteurs du numérique

Olivier Levillain¹

Depuis plusieurs années, on entend régulièrement des annonces, en France ou à l'étranger, affirmant que les besoins en sécurité du numérique sont immenses et qu'il faut former plus d'experts en cybersécurité. Bien que ce constat soit avéré, il est important de comprendre que la sécurité du numérique ne doit pas reposer uniquement sur des experts. Chaque acteur de la chaîne des systèmes d'information (administrateur, développeur, chef de projet, etc.) doit se sentir concerné et être impliqué, afin que la sécurité soit prise en compte tout au long des projets et de la vie des systèmes. On parle parfois de « *security by design* ».

Ainsi, tout administrateur réseau devrait savoir que les adresses réseau source d'un paquet peuvent être trivialement modifiées. Cela permettra d'éviter que le filtrage par adresse MAC soit vu comme un mécanisme de défense absolu. De manière similaire, tout développeur devrait avoir quelques notions concernant les vulnérabilités logicielles classiques, telles qu'un débordement de tampon (*buffer overflow* en anglais), ainsi que les moyens de s'en protéger (en effet, demander aux utilisateurs d'un site Internet de ne pas utiliser certains caractères dans leur mot de passe n'est pas une solution acceptable).

Développer chez les étudiants cet état d'esprit, cultiver ce réflexe du questionnement, susciter la curiosité face au monde du numérique, c'est la base de la sécurité.

1. Secrétaire adjoint de l'association CyberEdu, responsable du centre de formation de l'ANSSI.

CyberEdu, une démarche initiée par l'ANSSI

La démarche CyberEdu, initiée en 2013 par l'agence nationale de la sécurité des systèmes d'information (ANSSI), a pour objectif d'introduire des notions de sécurité dans l'ensemble des formations du numérique en France. Deux grandes actions ont été lancées par l'ANSSI pour CyberEdu : d'une part la réalisation et la mise à disposition de supports de cours destinés aux enseignants du supérieur en informatique souhaitant intégrer des éléments de sécurité dans leurs interventions et d'autre part la tenue de colloques réunissant des enseignants du supérieur en informatique et des experts de l'ANSSI.

Ces travaux ont été plus particulièrement menés par le centre de formation à la sécurité des systèmes d'information (CFSSI), avec l'appui des équipes techniques de l'ANSSI. Le CFSSI propose aux agents des fonctions publiques et aux militaires français des formations en SSI sur des sujets allant du panorama de la SSI à la cryptologie, en passant par l'analyse de risques ou l'audit. Un des objectifs de CyberEdu pour le CFSSI était de toucher un public plus large.

Les supports de cours CyberEdu

L'université européenne de Bretagne (qui regroupe 28 établissements d'enseignement supérieur et de recherche) et Orange ont remporté l'appel d'offre de l'ANSSI pour la fourniture des supports pédagogiques. Ces supports ont été mis à la disposition de tous les enseignants souhaitant utiliser ces ressources pour leurs cours² et comportent :

- un guide pédagogique, qui présente la démarche et propose d'expliquer quelques grands principes. Parmi ceux-ci, citons l'importance de ne pas traiter la sécurité comme un module à part, mais d'intégrer ce sujet, de le tisser autant que faire se peut au sein des cours existants ;
- un ensemble de planches de présentation, qui est un cours prêt à l'emploi d'une durée de 20 heures environ. Il est à destination des étudiants de niveau licence / DUT / BTS, pour leur donner les notions de base, le vocabulaire nécessaire à la compréhension des enjeux de sécurité et pour présenter succinctement l'organisation de la cybersécurité en France ;
- des fiches sur différents sujets pour les enseignants au niveau master : les systèmes d'exploitation, le réseau, le développement logiciel, l'authentification et les composants électroniques. L'objectif est de proposer à l'enseignant de tisser au cœur de son cours l'étude d'un sujet lié à la sécurité.

2. <https://www.ssi.gouv.fr/administration/formations/cyberedu/contenu-pedagogique-cyberedu/>

Ces documents ont été publiés sous licence Creative Commons Attribution (CC-BY). Cela permet à tous les enseignants le souhaitant de les réutiliser et de les adapter librement, y compris pour une utilisation commerciale, à la condition de rappeler la paternité des documents originaux (le projet CyberEdu).

Les colloques

Entre 2014 et 2017, cinq colloques ont été organisés par l'ANSSI pour les enseignants du supérieur en informatique afin de leur présenter la démarche et d'échanger sur l'intégration de la sécurité dans leurs cours. Plus d'une centaine de personnes venues de toute la France ont ainsi assisté à ces colloques.

Ces rencontres ont été l'occasion de partager le point de vue de l'ANSSI avec les enseignants. En plus de son objectif principal, CyberEdu a d'autres vertus qui peuvent intéresser les enseignants adhérant à la démarche. Tout d'abord, les sujets touchant à la sécurité sont plébiscités par les étudiants (encore plus depuis leur médiatisation auprès du grand public) ce qui aide à capter leur attention. Ensuite, pour comprendre un problème de sécurité, il est généralement nécessaire de bien comprendre le fonctionnement du sujet auquel le problème se rattache. Ainsi, la compréhension des vulnérabilités que peuvent introduire un débordement de tampon requiert des connaissances précises sur la représentation mémoire d'un processus.

Chaque colloque, d'une durée de trois jours, a été l'occasion d'aborder aussi bien des sujets techniques (systèmes d'exploitation, réseau, développement, notions de cryptographie) que juridiques et ont également permis de présenter quelques retours d'expériences.

L'association

Après cette première phase (2013-2016) menée par l'ANSSI, quelques limites de cette approche ont été constatées. D'une part, il semblait important d'associer durablement des enseignants à cette démarche, puisque ce sont eux les relais au cœur du dispositif. D'autre part, les colloques ont jusqu'à présent été uniquement organisés à Paris, pour des raisons logistiques.

Afin de répondre à ces limitations, l'ANSSI a proposé d'impliquer les enseignants dans la démarche, en commençant par ceux qui avaient suivi les colloques. C'est pourquoi, le 17 mai 2016, l'association CyberEdu a été fondée. Son objectif est de reprendre le flambeau en lien avec l'ANSSI. L'agence continuera en effet de s'impliquer dans la démarche CyberEdu via une participation à divers groupes de travail de l'association.

L'association a pour vocation de porter les missions suivantes :

- maintenir à jour les documents existants et en proposer de nouveaux ;
- proposer des colloques sur l'ensemble du territoire ;

- offrir un forum d'échanges entre spécialistes et non spécialistes de la sécurité ;
- labelliser des formations « CyberEdu » (voir section suivante).

L'association est jeune, mais ces différentes activités se structurent depuis quelques mois. L'actualité de l'association peut être suivie sur le site de l'association³.

L'esprit sécurité ?

Pour les spécialistes du domaine de la sécurité, certaines habitudes deviennent une seconde nature. Il leur est ainsi naturel d'étudier le comportement d'un programme face à des stimuli non prévus dans la spécification. En effet, l'attaquant n'est pas tenu de respecter les règles du jeu.

Même si CyberEdu n'a pas pour objectif de former des spécialistes en sécurité du numérique (mais la démarche peut toujours susciter des vocations !), il est néanmoins nécessaire de détruire certains mythes et d'en finir avec le côté parfois magique de l'informatique.

En réseau, un exemple classique est la capture de paquets réseau. Il est en effet toujours intéressant de découvrir la réaction des étudiants face à la quantité de données qui circulent en clair sur le réseau ! Sans cette prise de conscience, comment espérer qu'un administrateur comprenne l'importance des protocoles de sécurité ?

Lors du développement d'une application, l'objectif premier est évidemment de livrer un produit fonctionnel dans les délais impartis. Cependant, il arrive trop souvent que les développeurs oublient de tester le comportement de leur application face à des arguments invalides. Il s'agit pourtant d'une source classique de failles de sécurité. Si un utilisateur normal n'a aucune raison d'ajouter un caractère guillemet (") dans le champ login, un attaquant essaiera certainement d'introduire de tels caractères, considérés comme spéciaux par de nombreux moteurs de base de données (on parle d'injections SQL, une vulnérabilité très connue, et décrite jusque dans la BD en ligne xkcd, <https://www.xkcd.com/327/>).

L'esprit sécurité, c'est se poser des questions qui vont au-delà du fonctionnel !

3. <https://www.cyberedu.fr>

Le label CyberEdu

Afin de mettre en avant les formations incluant des contenus d'initiation à la sécurité du numérique, l'association a lancé au printemps 2017 un dispositif de labellisation.

Le processus de labellisation est le suivant⁴ :

- les organismes de formation intéressés remplissent un formulaire et le transmettent par voie électronique. Le formulaire décrit en particulier les contenus de sensibilisation et d'initiation à la sécurité du numérique inclus dans la formation candidate ;
- le dossier est instruit par l'association, qui vérifie que la demande est complète, et peut demander des précisions sur les contenus dispensés ;
- si le label est accordé, l'association publie les informations correspondant à la formation sur son site web ;
- pendant la durée de validité du label, l'organisme de formation s'engage à dispenser les contenus décrits dans le formulaire, selon des dispositions décrites dans une charte.

Attention, ce label, qui sera géré et décerné par l'association, concerne les formations de *non-spécialistes* en sécurité, et ne doit pas être confondu avec le label *SecNumedu*, délivré par l'ANSSI aux formations de *spécialistes* en sécurité du numérique.

Le futur de CyberEdu

Quelles extensions sont envisagées pour la démarche CyberEdu ? Au-delà des informaticiens, ou plus généralement des acteurs du numérique, il serait en effet logique de s'adresser à des enseignants d'autres disciplines, de communiquer notre message à d'autres populations (professions libérales, acteurs du monde de la santé par exemple), ou encore d'étendre la démarche à l'enseignement secondaire.

Mais avant cela, il faut pérenniser le travail réalisé auprès des enseignants du supérieur dans le domaine du numérique !

Pour la suite, restez à l'écoute sur cyberedu.fr !

4. <https://www.cyberedu.fr/pages/labellisation/>



Cybersécurité

Gildas Avoine¹ et Marc-Olivier Killijian²

Cet article a pour objectif de présenter quelques éléments de compréhension sur la recherche académique française en cybersécurité. L'importance de cette discipline pour notre société et les thématiques scientifiques qui la composent sont discutées dans les deux premières sections de cet article. Nous présentons ensuite les travaux de l'alliance Allistene sur la cartographie de la recherche académique française en cybersécurité, puis nous illustrons les activités de recherche de la communauté avec quelques défis scientifiques majeurs.

Cybersécurité dans la société

Évolution de la cybersécurité

L'actualité témoigne de l'importance toujours plus grande que l'on accorde à la cybersécurité, en particulier aux cyberattaques. Il ne s'agit malheureusement pas d'un biais journalistique, mais bien d'une réalité qui se traduit par un accroissement du nombre d'attaques et de leur impact.

Parmi ces attaques, il faut distinguer celles qui sont ciblées de celles qui ne le sont pas. Les premières visent des personnes ou des entités clairement identifiées et sont généralement conçues pour être efficaces sur la cible. Un exemple emblématique est le virus Stuxnet conçu pour s'attaquer aux centrifugeuses iraniennes d'enrichissement d'uranium. De manière générale, la victime peut être un individu, une entreprise, ou la société toute entière comme le laisse supposer les allégations d'ingérence dans les élections présidentielles américaines et françaises. Les attaques non

1. INSA Rennes, IRISA UMR 6074, IUF.

2. CNRS, LAAS UPR 8001.

ciblées sont quant à elles des attaques de masse qui reposent sur le fait que certaines cibles seront vulnérables. L'ampleur de l'attaque peut être considérable, comme ce fut le cas avec le réseau Mariposa qui comprenait 13 millions d'ordinateurs infectés. Plus récemment, c'est le rançongiciel WannaCry qui a défrayé la chronique : en touchant plus de 300 000 ordinateurs en seulement quelques heures, WannaCry a mis en lumière le risque bien réel lié aux rançongiciels.

Le temps des hackers isolés comme John Draper ou Kevin Mitnick n'est peut-être pas totalement révolu, mais c'est essentiellement le crime organisé, les activités cyberterroristes et le cyberespionnage qui inquiètent aujourd'hui. En formant les étudiants à la cybersécurité, nous formons ceux qui pourront lutter contre la cybercriminalité ou le cyberterrorisme, mais nous formons aussi les cybercriminels eux-mêmes. La cybercriminalité est en effet devenue une affaire d'experts très pointus, qu'ils soient des « *black hat* » ou des « *white hat* »³. Comme le soulignait David Naccache [8] lors du colloque de l'INS2I sur la cybersécurité en décembre 2016 : « les fraudeurs auxquels on fait face aujourd'hui utilisent des outils extrêmement avancés, ce sont des ingénieurs doués, ils connaissent très bien les normes et les publications académiques, ils utilisent des contre-mesures anti-expertales, et si jamais vous n'avez pas ça en tête lorsque vous mettez au point une application, elle finira par être cassée. »

Paradoxalement, la cybersécurité est aussi parfois perçue comme un obstacle à la sécurité nationale car elle est utilisée par les personnes malveillantes pour communiquer de manière confidentielle. Le compromis entre surveillance de masse et protection des libertés individuelles doit alors trouver un subtil équilibre acceptable par tous.

Informatique omniprésente

Si la sécurité a connu un tel essor, en particulier durant les 15 dernières années, c'est parce que l'informatique est devenue omniprésente. Il est aujourd'hui difficile d'avoir des activités qui ne font pas appel à l'informatique, même pour les loisirs. Le talon d'Achille est qu'il n'existe généralement pas de procédure pour poursuivre une activité en cas de problème informatique majeur. C'est ainsi qu'une gare ou un aéroport peuvent se retrouver immobilisés tant que l'attaque est en cours. La résilience des systèmes est un sujet important sur lequel les attentions se concentrent et les opérateurs d'importance vitale (OIV) sont particulièrement sensibilisés au sujet.

L'omniprésence de l'informatique implique aussi la multiplication des produits, des services et des acteurs économiques du domaine. Développer une application

3. *Black hat* et *white hat* sont des hackers informatiques. Alors que les premiers accomplissent des actions illégales ou malintentionnées, les seconds conservent un comportement éthique. Les *white hat* sont typiquement des experts en cybersécurité qui réalisent des tests de pénétration ou des recherches de vulnérabilités.

informatique sécurisée est une tâche difficile qui requiert des compétences qu'une petite structure ne peut pas toujours s'offrir. La pression financière pour réduire le temps de mise sur le marché rend la tâche encore plus difficile pour les équipes de conception et de développement. C'est ainsi que de nombreuses solutions logicielles sont développées sans prendre en compte l'état de l'art scientifique. La différence entre les connaissances académiques et ce qui est réellement implémenté dépasse souvent l'imaginable. Quelques cas réels sont par exemple : un concepteur qui « mélange » les octets des données d'une carte à puce pour en assurer la confidentialité car la cryptographie est pour lui un concept obscur et inintelligible ; un chef de projet qui ne garde que deux ou trois rondes de l'algorithme de chiffrement AES pour en accélérer l'exécution ; un fabricant de cartes à puce qui continue de vendre son produit alors qu'il est totalement cassé depuis presque 10 ans ; un constructeur de voitures qui utilise la même clef cryptographique dans des millions de véhicules, et dont la clef est accessible via le bus technique de la voiture ; etc. Même les applications les plus sensibles peuvent présenter des faiblesses liées à l'ignorance, ou la volonté d'ignorance de certains industriels. Ainsi le passeport électronique qui a vu le jour en Europe en 2004 (2006 pour la France) a souffert de nombreuses erreurs de jeunesse. Des tests d'interopérabilité internationaux orchestrés en 2008 ont montré que certains logiciels commerciaux de lecture des passeports ne vérifiaient aucune des mesures cryptographiques disponibles sur le passeport ! La cybersécurité n'est donc pas qu'un problème de recherche, c'est aussi et avant tout un problème d'ingénierie, c'est-à-dire qu'il faut pouvoir exploiter les connaissances connues en dépit des contraintes économiques et sociales qui peuvent faire pression. Il serait toutefois facile pour les chercheurs de lancer la pierre aux industriels en se dédouanant de toute responsabilité. Force est de constater qu'il y a aussi un combat qui doit être mené auprès des chercheurs pour rendre plus aisé le transfert de connaissances entre le monde académique et le monde industriel.

Il résulte de cette situation le sentiment que les problèmes vont plus vite que les solutions. C'est en partie vrai. Les primitives qui permettent d'assurer la sécurité sont de plus en plus sûres et les individus sont beaucoup mieux sensibilisés au problème. En revanche, l'omniprésence de l'informatique et de ses conséquences, ainsi que la découverte de nouvelles techniques d'attaque, par exemple celles reposant sur les canaux cachés, qu'ils soient matériels ou logiciels, augmente la surface d'attaque.

Thématiques scientifiques

Contours de la cybersécurité

Selon l'Union internationale des télécommunications (ITU), la cybersécurité est « l'ensemble des outils, politiques, concepts de sécurité, mécanismes de sécurité,

lignes directrices, méthodes de gestion des risques, actions, formations, bonnes pratiques, garanties et technologies qui peuvent être utilisés pour protéger le cyberenvironnement et les actifs des organisations et des utilisateurs. Les actifs des organisations et des utilisateurs comprennent les dispositifs informatiques connectés, le personnel, l'infrastructure, les applications, les services, les systèmes de télécommunication, et la totalité des informations transmises et/ou stockées dans le cyberenvironnement » [7].

Le terme *cybersécurité* est en fait apparu assez récemment, aussi bien en France qu'à l'étranger. Il était à l'origine peu usité par les scientifiques en raison de son manque de précision, mais il devient le terme moderne pour désigner ce que l'on appelait autrefois la sécurité des systèmes d'information ou la sécurité informatique. Il faut souligner que les frontières de la *sécurité informatique* ont beaucoup évolué ces vingt dernières années et que cette terminologie ne recouvre plus l'ensemble des disciplines concernées, notamment sur les aspects humains, sociétaux, économiques et juridiques. La désignation *sécurité du numérique* n'est pas non plus totalement appropriée ; par exemple la protection contre les signaux compromettants⁴ n'est pas limitée au monde numérique. La *sécurité de l'information* ne convient pas non plus car la cryptographie, par exemple, peut être utilisée pour du contrôle d'accès physique, sans qu'il y ait nécessairement une information à protéger à proprement parler. Face à cette difficulté de nommer cette discipline à part entière, le terme *cybersécurité* fait son chemin et s'impose progressivement dans la communauté scientifique.

La cybersécurité se situe ainsi à la croisée de nombreuses disciplines telles que l'informatique, les mathématiques, l'électronique et le traitement du signal. Récemment, les dimensions humaine, sociale, économique et juridique se sont ajoutées à l'équation car la sécurité n'a de sens au final que si elle est traitée dans sa globalité. À défaut, certaines dimensions du problème risquent d'être omises, conduisant à une sécurité réelle totalement caduque.

Classification

Les thématiques de la cybersécurité sont extrêmement variées. Plusieurs classifications existent car elles peuvent reposer sur les disciplines sous-jacentes (informatique, mathématiques, électronique...), sur les outils utilisés (méthodes formelles, cryptographie...), sur les objectifs à atteindre (protection de la vie privée, authentification, tatouage numérique...), ou encore sur les champs d'application (contrôle de processus industriels, contrôle d'accès...). Par exemple, la sécurité et la protection de la vie privée (*security and privacy*) est l'une des douze grandes thématiques de l'informatique qui apparaît dans le système de classification de l'ACM (CCS 2012). La sécurité et la protection de la vie privée est ensuite divisée en (1) cryptographie,

4. Les signaux compromettants sont des signaux électromagnétiques parasites – émis par un système logiciel ou matériel – qui peuvent laisser fuir de l'information sensible par rayonnement ou conduction. Il peut s'agir par exemple de signaux émis par un écran ou un clavier connecté par Bluetooth.

(2) méthodes formelles et aspects théoriques de la cybersécurité, (3) cybersécurité au niveau des services, (4) détection d'intrusion et protections contre les programmes malveillants, (5) sécurité des systèmes matériels, (6) sécurité des systèmes logiciels, (7) sécurité des réseaux, (8) sécurité du stockage de l'information et des bases de données, (9) sécurité des logiciels et des applications, (10) aspects humains, sociétaux et éthiques. Il est important de souligner qu'aucune classification ne peut faire l'unanimité et qu'il est difficile de concevoir une classification qui soit une partition de l'ensemble des thématiques.

Le GDR Sécurité informatique [2] (actuellement, pré-GDR), créé par l'Institut des sciences de l'information et de leurs interactions (INS2I) du CNRS en janvier 2016, considère quant à lui sept thématiques, assez proches de celles de la classification ACM, mais qui sont adaptées au paysage de la recherche française. Ces thématiques, déclinées en groupes de travail au sein du (pré-)GDR, sont les suivantes :

- *Codage et cryptographie*
- *Méthodes formelles pour la sécurité*
- *Protection de la vie privée*
- *Sécurité et données multimédias*
- *Sécurité des réseaux et des infrastructures*
- *Sécurité des systèmes logiciels*
- *Sécurité des systèmes matériels*

L'interdisciplinarité de la cybersécurité s'illustre pleinement ici, puisque trois groupes de travail (GT) sont communs avec d'autres GDR : le GT *Codage et cryptographie* (C2) est commun avec le GDR *Informatique mathématique* (IM), le GT *Sécurité et données multimédias* est commun avec le GDR *Information, Signal, Image et Vision* (ISIS), et le GT *Sécurité des systèmes matériels* est commun avec le GDR *System-On-Chip, System-In-Package* (SoC-SiP).

Thématiques

Afin de comprendre la classification utilisée par le (pré-)GDR Sécurité informatique, chacune des thématiques mentionnées est succinctement présentée dans cette section.

Codage et cryptographie.

La cryptographie vise à garantir la confidentialité, l'authenticité et l'intégrité des informations et des communications. Ces besoins remontent à la nuit des temps mais ce n'est que récemment que l'on peut parler de science. La cryptographie s'est d'abord mécanisée, puis informatisée, et ce sont enfin les mathématiques, dans les années soixante-dix, qui ont révolutionné ce champ disciplinaire. La recherche française en cryptographie est très bien développée et structurée, et elle possède une reconnaissance internationale incontestable. Dire qu'il existe une école française de la cryptographie n'est pas exagéré, et les figures de proue de cette école ont remporté de nombreuses reconnaissances nationales et internationales. Historiquement,

la recherche en cryptographie et en codage s'est structurée en France autour du GT *Codage et cryptographie* qui regroupe environ 450 personnes. Ce GT est affilié aujourd'hui au GDR IM et au (pré-)GDR Sécurité informatique, mais sa création est en fait bien antérieure à celle de ces deux structures.

Méthodes formelles pour la sécurité.

Lors du colloque sur la cybersécurité organisé par l'INS2I en décembre 2016, Hubert Comon [3] débute son intervention en faisant référence aux précédents exposés qui étaient principalement axés sur les attaques : « (...) On n'est pas seulement dans ce cycle infernal où il y a des hackers qui cherchent des attaques, je trouve des contre-mesures et je recommence, etc. On peut imaginer autre chose. Par exemple, on prouve la sécurité. Si je prouve la sécurité, point final. » Le domaine des méthodes formelles pour la sécurité vise ainsi à définir de manière formelle les propriétés qu'un système doit garantir, dans le but de mieux les comprendre et de les prouver. Le domaine des méthodes formelles pour la sécurité est très prometteur et il a déjà montré son efficacité sur des protocoles cryptographiques, notamment pour l'authentification dans le passeport biométrique et pour les protocoles de vote électronique. Le chemin est toutefois encore long avant de pouvoir prouver la sécurité d'un système dans sa globalité car une difficulté majeure est de clairement identifier les hypothèses de la preuve. Ainsi, Hubert Comon dans son exposé précise humblement : « Vous allez voir qu'il y a des bémols, ce n'est pas si simple que ça. »

Protection de la vie privée.

Malgré un arsenal législatif relativement fort (au moins en Europe) autour de la protection des données personnelles et une prise de conscience des individus de plus en plus prégnante, nous avons tous le sentiment que nos données personnelles et notre vie privée ne sont pas concrètement protégées. Il existe en effet différentes sources de risques pour la vie privée dans le monde d'aujourd'hui. En particulier, la publicité ciblée sur Internet est devenue une source de financement importante et passe par la collecte massive de données sur les utilisateurs. Les profils des internautes obtenus à partir des nombreuses traces numériques (laissées par les individus eux-mêmes de façon passive, ou collectées activement) intéressent également les acteurs traditionnels de l'économie, telles les banques et les assurances. Enfin ces profils et données personnelles peuvent être utilisés à des fins crapuleuses comme du chantage ou de l'escroquerie par ingénierie sociale. La protection de la vie privée est une question transdisciplinaire où les aspects humains, légaux, mathématiques et informatiques doivent être considérés dans leur globalité et ne peuvent pas être traités séparément.

Sécurité et données multimédias.

La protection de l'information passe aussi par le traitement du signal, notamment mais pas seulement, le traitement des images. Il s'agit d'une thématique importante de la cybersécurité qui traite par exemple de la protection des droits sur les images

et de la stéganographie⁵. La biométrie, qui consiste à utiliser les caractères biologiques intrinsèques à une personne pour garantir certaines propriétés de la sécurité, comme l'authentification, se retrouve également dans la communauté des données multimédias. Il existe évidemment une interaction forte entre cette thématique et la cryptographie.

Sécurité des réseaux et des infrastructures.

La communauté de la sécurité des réseaux et des infrastructures se retrouve en partie dans les communautés de la cryptographie de la sécurité des systèmes logiciels. Il existe toutefois des sujets bien spécifiques à cette thématique, comme, de façon non-exhaustive : la surveillance des réseaux (dénis de service distribués, menaces persistantes...), la gestion des identités et des clefs (cycles de vie), la sécurisation des protocoles distribués (par exemple, la sécurisation du protocole BGP), ou encore la sécurisation des réseaux de nouvelle génération (IoT/M2M, SDN, CDN...).

Sécurité des systèmes logiciels.

La sécurité des systèmes logiciels porte sur la conception de systèmes sécurisés et sur l'analyse des vulnérabilités. Il est essentiel dans cette thématique d'avoir un point de vue vertical qui va du logiciel jusqu'à son interface avec le matériel, mais aussi une approche horizontale qui cherche à couvrir toutes les plateformes déployées. Il faut pour cela prendre en compte la variété des attaques par logiciels malveillants (vers, virus, botnets, rançongiciels, chevaux de Troie...) mais aussi les vulnérabilités liées notamment à des canaux cachés et des corruptions de mémoire. La recherche de vulnérabilités se fait en analysant les programmes de manière statique ou dynamique et doit faire face aux protections logicielles mises en place par les fraudeurs, à savoir l'obfuscation de code⁶, le chiffrement et toutes les méthodes d'auto-modification.

Sécurité des systèmes matériels.

La sécurité n'a de sens que si elle est assurée dans sa globalité, on parle d'ailleurs de sécurité de bout en bout. Cela signifie qu'il n'est pas suffisant de concevoir des primitives qui sont sûres sur le papier, il faut qu'elles le restent après leur implémentation sur un système matériel, par exemple une carte à puce. L'informatique est aujourd'hui embarquée dans de nombreux objets de la vie de tous les jours (abonnement aux transports publics, clef de démarrage de voiture, télécommande de garage, téléphone portable...), ce qui illustre l'importance de ce champ disciplinaire. La recherche française est aux avant-postes de la sécurité des systèmes matériels et elle s'est structurée autour de l'axe « confiance matérielle » du GDR SoC-SiP. Ce GDR

5. La stéganographie consiste à dissimuler un message dans un autre message, par exemple insérer un identifiant invisible dans une image.

6. L'obfuscation est une technique de protection de code source contre les attaques par rétro-ingénierie. L'obfuscation consiste typiquement à rendre le code source incompréhensible à un attaquant, sans pour autant impacter les fonctionnalités du programme.

est rattaché à l'Institut des sciences de l'ingénierie et des systèmes (INSIS) et à l'Institut des sciences de l'information et de leurs interactions (INS2I).

Cartographie de la cybersécurité

Depuis quelques années, il y a un besoin significatif en recrutement en cybersécurité en raison d'une pénurie de techniciens et d'ingénieurs formés à ce domaine. À titre d'illustration, le nombre d'offres d'emploi en cybersécurité diffusé par l'APEC est passé de 315 en 2014 à 1133 en 2016 [1]. Il est important de souligner qu'il y a un manque d'experts en cybersécurité, mais aussi un manque d'ingénieurs formés à la sécurité sans pour autant en être experts. Cette pénurie impacte les entreprises privées, les services étatiques, mais aussi les laboratoires publics de recherche. Il est aujourd'hui difficile de recruter des enseignants-chercheurs et des doctorants car les candidats sont happés par le monde professionnel de la cybersécurité.

Bien qu'il n'y ait pas de chiffre officiel, il semble également que les étudiants à fort potentiel s'orientent majoritairement vers des thématiques en lien avec des questions théoriques. Plusieurs champs disciplinaires de la cybersécurité qui ne possèdent pas la même maturité scientifique bénéficieraient pourtant aussi de pouvoir recruter de tels profils. La situation évolue toutefois et on peut se réjouir de voir que le nombre de formations en cybersécurité a explosé ces dernières années et, surtout, qu'elles sont particulièrement attrayantes pour les étudiants.

À partir de ces constatations, il est intéressant de s'interroger sur la répartition des forces académiques en France dans le domaine de la cybersécurité. Depuis 2015, l'alliance Allistene [6] (Alliance des sciences et des technologies du numérique) s'est dotée d'un groupe de travail « Cybersécurité » dont l'une des missions a été le recensement des forces académiques en cybersécurité.

La méthodologie suivie pour réaliser cette cartographie a reposé sur l'envoi d'un questionnaire aux responsables des équipes identifiées *a priori* ou qui se sont déclarées *a posteriori*. Le questionnaire comportait essentiellement deux parties : la première, quantitative, visait à évaluer les forces en termes de chercheurs, enseignants-chercheurs, doctorants, post-doctorants et ingénieurs ; la seconde partie, qualitative, visait à identifier plus spécifiquement les domaines scientifiques de chaque équipe.

Les résultats de cette étude [5] sont riches en instructions. L'étude montre notamment que la recherche française en cybersécurité est composée, à la date du recensement (année académique 2016-2017), d'environ 852 *équivalents temps-plein pour la recherche* (ETP). Notons que cette étude a porté sur l'ensemble des employeurs académiques français : CEA, CNRS, écoles d'ingénieurs, Inria et universités. Les équivalents temps-plein pour la recherche ont été calculés de la manière suivante : un chercheur déclarant travailler à 80 % sur la cybersécurité est comptabilisé à hauteur de 0,8 ETP, alors qu'un enseignant-chercheur également impliqué à hauteur de 80 % sur la cybersécurité est comptabilisé à hauteur de 0,4 ETP.

Répartition géographique des ETP en cybersécurité

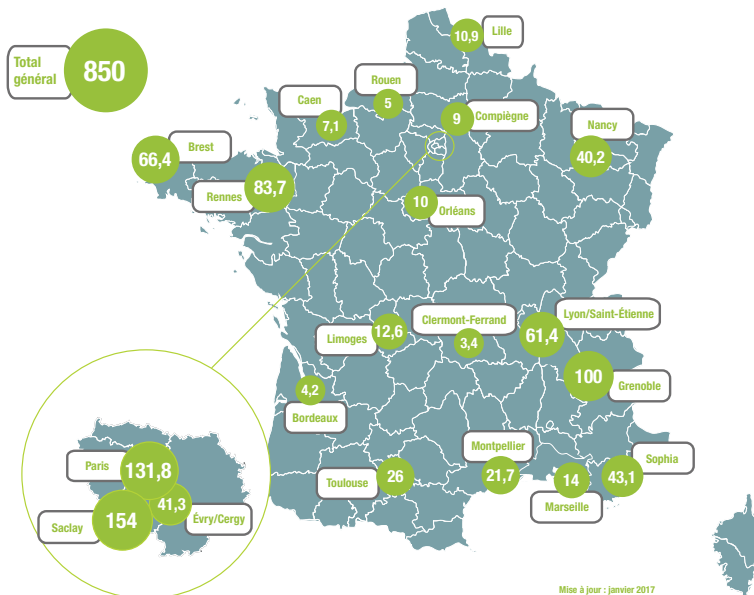


FIGURE 1. Répartition géographique des ETP de la recherche académique française en cybersécurité (janvier 2017).

Les 852 ETP correspondent à 143 ETP chercheurs, 130 ETP enseignants-chercheurs, 126 ETP ingénieurs, 82 ETP post-doctorants et 371 ETP doctorants, répartis sur 1 101 personnes physiques. Les répartitions géographiques de ces ETP et des personnes physiques sont illustrées sur les figures 1 et 2.

La répartition des forces selon les thématiques de la cybersécurité est également une information pertinente, fournie dans la table 1.

Défis scientifiques

Pour illustrer les défis scientifiques auxquels la recherche en cybersécurité devra faire face dans les années à venir, nous avons consulté les responsables des groupes de travail du (pré-)GDR Sécurité informatique. La liste présentée ne prétend pas être exhaustive. Elle n'a pour ambition que d'illustrer le domaine avec quelques défis d'envergure qui occupent les esprits des chercheurs en cybersécurité.

Répartition géographique des personnels en cybersécurité

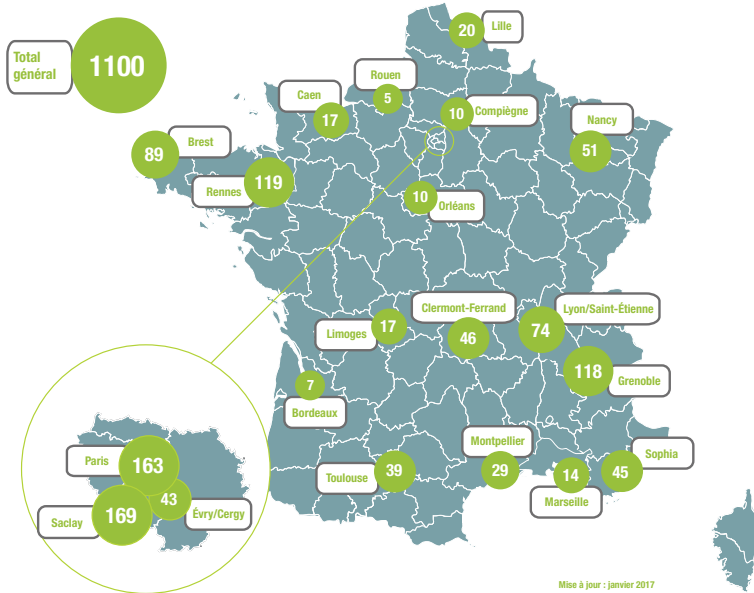


FIGURE 2. Répartition géographique des personnes physiques de la recherche académique française en cybersécurité (janvier 2017).

Logiciels malveillants. Un facteur déterminant pour mener une attaque est la présence de vulnérabilités dans le système visé. La conception d'un système sûr constitue donc un élément primordial de défense. En particulier les attaques ciblées vers une personne ou un groupe de personnes peuvent être sophistiquées, avec des codes d'attaques qui sont obfusqués, chiffrés et auto-modifiants (par exemple, le code ne se déploie qu'après une suite de décompressions et de déchiffrements). Un autre aspect à prendre en compte est la masse de codes malveillants produite, le plus souvent en réemployant une souche connue. Face à cela, un défi majeur de la cybersécurité est d'identifier une attaque en déployant et combinant des approches formelles, des heuristiques d'analyse et des méthodes d'apprentissage.

Attaques physiques. Dans le domaine de la sécurité des systèmes matériels, de nouvelles technologies d'implantation microélectronique (FDSOI, MRAM, RRAM, nanotubes...) doivent être évaluées vis-à-vis des attaques physiques existantes et vis-à-vis de leur capacité à intégrer des éléments essentiels de la sécurité, comme la

Thématique de la cybersécurité	
Cryptographie	20 %
Sécurité des systèmes matériels	17 %
Méthodes formelles et aspects théoriques de la cybersécurité	16 %
Sécurité des systèmes logiciels	10 %
Cybersécurité au niveau des services	8 %
Détection d'intrusion et protections contre les logiciels malveillants	8 %
Sécurité des réseaux	8 %
Sécurité des logiciels et des applications	6 %
Sécurité du stockage de l'information et des bases de données	3 %
Science forensique (analyse du système après un incident)	3 %
Aspects humains, sociétaux et éthiques	2 %
	100 %

TABLE 1. Répartition des activités à travers les thématiques de la cybersécurité, exprimée en pourcentage du nombre d'ETP total.

génération de nombres aléatoires (TRNG et PUF), la mémorisation sécurisée de données sensibles et les services cryptographiques. La recherche de failles de sécurité inhérentes à ces technologies doit être l'un des objectifs prioritaires des travaux à conduire dans les années à venir.

Sécurité matérielle par conception. Un autre objectif fort de ce domaine est le développement de méthodes de sécurité par conception des systèmes sur puces complexes et hétérogènes, avec de nouveaux enjeux comme la gestion des droits de propriété des composants virtuels et des circuits, la protection vis-à-vis de matériels malicieux, la surveillance comportementale des systèmes, la résilience des systèmes matériels sensibles, la garantie des fonctions de test, de diagnostic et de débogage sans préjudice sur la sécurité et la proposition de mécanismes sophistiqués d'authentification et d'identification intrinsèque du matériel (cas des fonctions physiques non clonables, PUF).

Cryptographie post-quantique. Concevoir des algorithmes cryptographiques résistants à l'ordinateur quantique, tel est l'objectif de la cryptographie post-quantique. Étant donné que les problèmes de factorisation et de logarithme discret peuvent être résolus facilement avec un ordinateur quantique, il est important de concevoir des algorithmes alternatifs. Des solutions reposant notamment sur des réseaux euclidiens, des codes correcteurs et des polynômes multivariés existent. Le défi consiste à en améliorer l'efficacité, en réduisant notamment la taille des clefs.

Chiffrement complètement homomorphe. Un défi important dans le domaine de la cryptographie est aussi la conception de chiffrement complètement homomorphe utilisable dans des applications réelles. L'objectif d'une telle primitive est de permettre d'appliquer tout type de fonction calculable sur des messages, sans nécessiter pour cela leur déchiffrement. Une telle primitive permet par exemple de déléguer à une tierce partie (par exemple un serveur dans le *cloud*) l'exécution d'opérations sur des messages, sans pour autant lui révéler le contenu de ces messages.

Vote électronique. Un autre défi important pour notre société est le vote électronique [4]. Le vote électronique présente de nombreux avantages mais il soulève aussi de nombreux défis scientifiques, même si des avancées récentes ont été obtenues dans ce domaine. Par exemple, lors d'un vote par Internet, une difficulté est de s'assurer de l'identité de l'électeur afin de garantir que ce dernier est bien la personne à l'origine du vote et que c'est bien le choix de celui-ci qui a été pris en compte. Cette difficulté doit également faire face aux logiciels malveillants potentiellement installés sur l'ordinateur de l'électeur.

Manipulation d'images. Dans le domaine de la sécurité des données multimédias, garantir une fiabilité très élevée de détection de manipulation d'images (copier-coller, effacement de zones, stéganographie...) constitue l'un des sujets majeurs pour les années à venir. Garantir qu'une image n'a pas été manipulée est, par exemple, fondamental pour les experts judiciaires afin qu'elle puisse constituer une preuve recevable. Plus généralement, l'image prend une place de plus en plus importante dans le monde numérique et CISCO prévoit que les données visuelles (images, vidéos...) représenteront 80 % du trafic sur Internet en 2019.

Conclusion

La recherche scientifique en cybersécurité est étonnante car elle mêle des communautés scientifiques bien établies, structurées et reconnues depuis longtemps, avec des communautés qui posent les premières briques de leur structuration. Afin d'animer ces communautés et leur offrir la possibilité de mieux se connaître, le CNRS a décidé de créer un (pré-)GDR Sécurité informatique. Cette structure qui recouvre des chercheurs de tous horizons permet de mettre en place un réseau scientifique national à travers ses groupes de travail, qui ont d'ores et déjà organisé plusieurs événements scientifiques en 2017. Le (pré-)GDR a aussi mis en place des activités transverses aux groupes de travail, notamment les journées nationales, organisées au siège du CNRS à Paris en 2017, une école pour jeunes chercheurs organisée à Rennes en 2016 et à Bourges en 2017, une semaine de rencontre entre entreprises et doctorants (REDOCS) organisée dans les locaux du CNRS à Gif-sur-Yvette en 2016 et 2017, et enfin un colloque ouvert à un plus large public, « Sécurité informatique : mythes et réalité », en décembre 2016. Prochainement, c'est aussi un club de

partenaires qui verra le jour pour renforcer les liens entre recherche académique et recherche industrielle.

Références

- [1] APEC. Cybersécurité en Bretagne : l'enjeu des compétences. Juin 2017.
- [2] CNRS/INS2I. Groupement de recherche (pré-GDR) Sécurité informatique. (<http://gdr-securite.irisa.fr>).
- [3] Hubert Comon. Peut-on prouver la sécurité des communications ? Colloque du CNRS sur la sécurité informatique : mythes ou réalité (<http://gdr-securite.irisa.fr/colloque.html>), décembre 2016.
- [4] Véronique Cortier. Vote électronique. 1024 – *Bulletin de la société informatique de France*, (9):95–109, novembre 2016 (<http://www.societe-informatique-de-france.fr/wp-content/uploads/2016/11/1024-no9-vote-electronique.pdf>).
- [5] Groupe de travail Cybersécurité de l'alliance Allistène. Cartographie de la recherche académique française en cybersécurité. (https://www.allistene.fr/files/2015/04/2017-06-10-cartographie_final.pdf), 2017.
- [6] Allistene (Alliance des sciences et technologies du numérique). (<https://www.allistene.fr/>).
- [7] ITU. X.1205 série X : Réseaux de données, communication entre systèmes ouverts et sécurité – sécurité du cyberspace – cybersécurité – présentation générale de la cybersécurité, avril 2008.
- [8] David Naccache. De la puce au cloud, comment fonctionne l'expertise judiciaire ? Colloque du CNRS sur la sécurité informatique : mythes ou réalité (<http://gdr-securite.irisa.fr/colloque.html>), décembre 2016.



Le logiciel libre, de l'usage à la recherche

Entretien avec Roberto Di Cosmo (réalisé par Valérie Schafer)

Actuellement détaché auprès d'Inria de l'université Paris-Diderot où il est professeur d'informatique, Roberto Di Cosmo se consacre pleinement à Software Heritage dont il est le directeur. Ce projet patrimonial dédié aux codes et logiciels libres concrétise des années de pratique, d'investissement et de recherches dans le domaine des logiciels libres.

Retour avec Roberto Di Cosmo sur le parcours qui l'a mené d'Italie en France, de la découverte des ordinateurs à la recherche en informatique, enfin de l'agacement face aux systèmes fermés et propriétaires à l'action dans le domaine des logiciels libres et à leur préservation.

Valérie Schafer¹

Valérie Schafer : *Avant que nous abordions le projet Software Heritage pour lequel vous êtes en détachement depuis un an chez Inria, j'aimerais que l'on revienne sur votre parcours dans l'informatique, débuté en Italie.*

Roberto Di Cosmo : Le contexte italien est alors différent du contexte français. Cela va sans doute vous surprendre, mais, à mon époque, en Italie on mettait beaucoup plus l'accent sur les filières littéraires que scientifiques, contrairement à l'usage en France.

Au lycée je poursuis des études de philosophie, d'histoire, de lettres, de latin et de grec. Et c'est par ce biais que je vais me pencher sur l'informatique. En effet je remporte un prix lors d'un concours de traduction de textes du grec au latin. Et

1. Valérie Schafer, ISCC, CNRS/Paris-Sorbonne/UPMC. Entretien co-publié avec la revue *Technique et science informatiques* (TSI).

c'est avec l'argent de ce prix que je m'achète mon premier ordinateur ! On est en 1979 et c'est un ZX 80, avec un processeur 8080 cadencé à quelques mégahertz qui génère aussi le signal vidéo, et un seul kilo-octet de RAM, partagé entre mémoire de programme et mémoire vidéo. C'est un ordinateur extrêmement rudimentaire : si le programme que l'on écrit est trop long, la partie d'écran affichée se réduit ; quand on appuie sur une touche, on désynchronise l'écran, etc. Et pourtant, c'était absolument magique !

V. S. : *Vous apprenez sur le tas ?*

R. D. C. : Il n'y avait pas de cours d'informatique au lycée, et il y avait une seule université d'Informatique dans toute l'Italie, mais pour un lycéen motivé, quelques magazines informatiques, nés des milieux hobbyistes notamment, suffisaient pour commencer à écrire quelques petits programmes, et faire surgir plein de questions qui attendaient réponse.

Très naturellement, quand j'intègre la *Scuola Normale Superiore di Pisa* (École normale supérieure de Pise) je choisis l'option informatique. Nous devions être trois ou quatre en informatique et les promotions étaient très petites, autour de 25 personnes. Contrairement au système français où l'entrée à l'ENS se fait après une classe préparatoire, en Italie le concours était (et est encore) à la sortie du lycée : il n'y a pas de classe préparatoire, parce qu'il y a très peu de grandes écoles.

La sélection doit donc se faire parmi tous les lycéens, peu importe qu'ils soient issus de filières littéraires ou scientifiques. Et les épreuves ne peuvent pas s'appuyer sur un programme touffu et balisé comme en France pour déceler les meilleurs candidats. Si en France les sujets en mathématiques font plusieurs pages et demandent maîtrise et rapidité, à Pise tout tient sur un feuillet et il faut faire preuve de créativité et d'imagination. Si en France l'admission à une grande école est la preuve qu'on est parmi les meilleurs, à Pise après le passage du concours on doit constamment prouver qu'on est bien le bon candidat, et pas quelqu'un qui a eu la bonne idée au bon moment par hasard, avec des conditions draconiennes sur les résultats attendus à la fois dans les cours à la Faculté et les cours supplémentaires dispensés en interne. Dans ma promotion à Pise, nous étions une vingtaine à entrer, et on était un peu plus d'une dizaine à la fin.

Je fais mes études d'informatique entre 1982 et 1986. On pouvait avoir encore l'illusion de tout comprendre de l'informatique. Je dis bien l'*illusion* de tout savoir. Les matières enseignées l'étaient chacune alors par un luminaire, un maître dans sa discipline. Bien sûr aucun d'entre eux ne venait directement de l'informatique, ce champ était en construction. Ils n'avaient pas étudié l'informatique, mais ils l'avaient créée ! La plupart d'entre eux étaient partis à l'étranger puis revenus ensuite pour créer cette formation à l'université de Pise. Cette période est véritablement révolutionnaire ! Deux ans avant que je commence ces études d'informatique à la Faculté, la promotion était seulement d'une cinquantaine d'étudiants. En 1982, nous étions

environ 1200. Le lancement du PC IBM en 1981 avait notamment contribué à l'explosion du nombre d'étudiants.

V. S. : *Pas évident à gérer pour les enseignants...*

R. D. C. : Les moyens informatiques à disposition des étudiants d'informatique de l'université étant effectivement totalement sous-dimensionnés pour cette nouvelle foule, les cours étaient essentiellement théoriques dans ces années, ce qui n'est pas forcément une chose si mauvaise, vu le grand nombre d'étudiants qui sont devenus des informaticiens théoriques de niveau mondial.

À l'université, nous avions accès à un ordinateur, mais nous étions encore dans un système de *batch processing*². On pouvait entrer des calculs, mais nous n'avions les résultats que le lendemain.

Mais heureusement, à l'École Normale de Pise notre tout petit groupe de normaliens informaticiens avait sympathisé avec les responsables du centre de calcul, qui se révéla une véritable caverne d'Ali Baba. On a passé des heures à analyser le fonctionnement d'un vieil ordinateur HP décommissionné : il y avait une magnifique plaque en bronze avec gravées les instructions à insérer péniblement à la main pour mettre en marche le lecteur de rubans perforés, et heureusement une mémoire à noyaux magnétiques qui maintenait l'information des mois après avoir éteint la machine. De la même façon lorsque le PDP-11 de la comptabilité a été remplacé, nous l'avons récupéré, en découvrant les subtilités d'un système multitâche et multi-utilisateur.

Et très rapidement on rentre dans la période de développement ultrarapide de l'informatique personnelle de l'Apple 2 au Macintosh, des premiers PC IBM aux Atari et Amiga. C'est aussi le moment de la découverte des premiers jeux d'aventures, des premiers programmes d'intelligence artificielle et si nous sommes bien sûr tous intéressés à y jouer, on finit par s'intéresser beaucoup plus à comprendre comment tout cela fonctionne, en faisant du *reverse engineering* et en s'amusant comme des fous.

Il y avait un autre aspect intéressant qui compliquait la tâche des enseignants : l'apparition des premiers ordinateurs « de poche », comme le Sharp PC-1500, qui pouvaient servir à calculer automatiquement les réponses aux exercices de mathématiques.

Je me rappelle que je trouvais insupportable l'idée de devoir faire à la main des calculs qui sont automatisables, comme les dérivées. J'avais donc programmé un dérivateur formel pour faire le calcul des examens plutôt que de les effectuer à la main : comme il n'y avait pas assez de mémoire, je ne pouvais pas programmer toutes les fonctions prévues au programme, et j'ai dû finir par me résigner à me

2. Dans le *batch processing* ou traitement par lots, les programmes ou travaux devant être exécutés sont placés les uns derrière les autres. Ils constituent ainsi des lots qui sont traités par l'unité centrale de manière séquentielle.

convertir en calculette humaine pour passer l'examen d'analyse. Mais cette envie d'automatiser n'en était que plus vivante en moi.

V. S. : *Vous finissez ces études en 1986 ?*

R. D. C. : Absolument, et je pars faire mon service militaire ensuite, pendant lequel, après des longs mois d'entraînement, je me retrouve aussi à... enseigner l'informatique aux sous-officiers. En 1988 je commence une thèse sur la théorie des types sous la direction de Giuseppe Longo, professeur de logique à Pise et aujourd'hui directeur de recherche CNRS à Ulm. En 1989 j'ai passé une période à Cornell, une excellente université américaine, mais à mon retour mon directeur de thèse était en visite à l'ENS Ulm, donc je l'ai suivi, comme bien d'autres de ses doctorants, et je me suis retrouvé à Paris. J'ai découvert l'Inria Rocquencourt, où l'on allait religieusement suivre le séminaire du vendredi organisé par Gérard Huet, et j'ai pu croiser des chercheurs de grande renommée comme Jean-Jacques Lévy, Maurice Nivat, Gérard Berry, Pierre-Louis Curien, Jean Berstel, Jean Vuillemin, et bien d'autres qui j'espère ne se vexeront pas du fait que j'arrête ici la liste à la Prévert. C'était inespéré, et j'ai naturellement prolongé mon séjour, devenant d'abord ATER et ensuite Maître de Conférences à l'ENS Ulm, vivant quelques années magiques, dans une atmosphère de réelle pluridisciplinarité.

Je me souviens par exemple d'avoir un jour entendu des étudiantes littéraires qui parlaient des œuvres de Shakespeare et cherchaient à étudier la fréquence des mots utilisés : certaines avaient passé des jours entiers à pointer le nombre d'occurrences d'un mot donné dans toutes les œuvres. On était alors en 1990 ou 1991, et j'utilisais un ordinateur NeXT : nouvelle aventure de Steve Jobs après avoir été éjecté d'Apple, cette machine magnifique intégrait non seulement des logiciels en avance sur leur temps (certains écrits par Jean-Marie Hulot, que Steve Jobs était venu chercher exprès chez Inria), mais aussi une copie numérique de toutes les œuvres de Shakespeare. Bref en une heure j'avais construit la table des fréquences non pas d'un, mais de tous les mots qui apparaissaient dans les textes de Shakespeare, et convaincu un certain nombre de littéraires que l'Informatique n'était pas qu'un petit jeu de *geeks*.

Je suis resté à Ulm, jusqu'en 1999, et pendant ces années j'ai eu le privilège de rencontrer des étudiants exceptionnels, auxquels j'ai donné tout ce que j'avais : je garde encore le contact avec beaucoup d'entre eux. Ma recherche est très théorique : sémantique des langages de programmation, réécriture, théorie des nombres et théorie des catégories. Mais en parallèle je conserve un indomptable esprit *hacker* : j'adore les machines, j'adore jouer avec et je ne supporte pas qu'elles soient verrouillées, m'empêchant de corriger les erreurs qu'on rencontre, et me forçant à suivre un chemin pré-tracé.

Je commence donc à m'intéresser aux logiciels libres : ce n'était pas du pain béni, la variété et la qualité des logiciels libres disponibles à l'époque étant plutôt limitées,

mais c'était bien un moyen pour récupérer un peu de sa liberté, et il y avait une vraie dynamique qui émergeait. C'est à ce moment-là que je découvre que, même si je ne veux pas du système d'exploitation propriétaire qui est vendu avec les machines, je dois le payer quand même. C'était un comble.

V. S. : *Ce sont ces constats qui vous conduiront à écrire en 1998 Piège dans le cyberspace puis, avec la journaliste Dominique Nora, Le Hold-up planétaire : la face cachée de Microsoft*³ ?

R. D. C. : Je me rendais compte que l'informatique était en train de changer le monde. Mais on ne nous proposait pas les meilleurs outils. Au contraire on essayait de rentabiliser ce qui était déjà sur le marché et de freiner les concurrents. *Piège dans le cyberspace* était d'abord une thérapie psychologique personnelle qui se voulait un manifeste de révolte, plutôt humoristique, destiné à circuler entre amis.

Mais les amis ont commencé à le faire circuler : il se retrouve en ligne sur un site web au Canada et le nombre de lecteurs – dont témoigne alors le compteur de visites – explose, avec des milliers de téléchargements par jour. Il est ensuite traduit en plusieurs langues, et je me retrouve propulsé contre toute attente porte-parole d'une majorité silencieuse qui s'ignorait. Ce texte comportait deux volets : une analyse critique fine des problèmes de l'industrie informatique de l'époque, et des propositions alternatives, comme l'accès au code source et l'adoption des logiciels libres.

Piège dans le cyberspace

Ce texte est depuis le 20 mars 1998 sur le Web
<http://www.mmedium.com/dossiers/piege>

« Pendant les vacances de Noël, j'ai été frappé une fois de plus par l'engouement croissant des médias pour cet obscur objet du désir qui se cache derrière les mots « ordinateur », « multimédia », « web », « Internet » et leurs dérivés. À croire ces médias et bon nombre d'experts improvisés, on ne peut se prétendre citoyen à part entière sans posséder le matériel informatique flambant neuf (et très cher) donnant accès au paradis féérique du « cyberspace ».

Difficile aussi de ne pas remarquer l'étrange et omniprésent amalgame qui nous incite à penser que le seul type d'ordinateur existant est le PC, à condition bien entendu qu'il soit équipé d'une puce Intel, et que sur ce PC, il y a un seul logiciel indispensable, Windows produit par Microsoft [...].

Par contre, notre public à nous est bien loin de se réveiller : bercé par la voix douce du conformisme ambiant, il s'endort de plus en plus dans les bras de Microsoft. Il rêve

3. R. Di Cosmo et D. Nora, *Le Hold-up planétaire : la face cachée de Microsoft*, Paris, Calmann-Levy, 1998. Le livre est disponible sous licence Creative Commons Attribution-NoDerivs-NonCommercial. Voir <http://www.dicosmo.org/HoldUp/HoldUpPlanetaire.pdf>

d'un monde joyeux où un grand philanthrope distribue à tous les écoliers de France des copies gratuites de Windows 95, dans le seul but de les aider à rattraper leur retard technologique ; il sourit en pensant aux écrans bleus pleins de messages rassurants qui expliquent comment le programme machin a provoqué l'exception bidule dans le module truc non pas par la faute de Windows, bien entendu, mais par celle du programme ; il dort heureux sans se demander pourquoi un ordinateur bien plus puissant que celui qui a servi à envoyer des hommes sur la lune, et à les ramener vivants, n'est pas en mesure de manipuler correctement un document d'une centaine de pages quand il est équipé par ce Microsoft Office qui fait la joie de tous nos commentateurs [...] ».

Retrouvez l'intégralité du texte sur

<http://www.dicosmo.org/Piege/cybersnare/piege.html>

La critique à l'époque portait surtout sur Microsoft, qui était assise dans une position dominante exploitée sans réserve, et malheureusement beaucoup de lecteurs ont juste retenu un message négatif en général vis-à-vis de cette entreprise. En réalité, l'analyse était plus fine : ce que je critiquais n'était pas tant l'entreprise Microsoft en tant que telle, mais la stratégie de verrouillage et de contrôle absolu de toute la chaîne de l'information qu'elle poursuivait à l'époque, et les moyens de la mettre en œuvre. Aujourd'hui, bien d'autres acteurs sont plus inquiétants, et l'analyse de 1999 s'appliquerait à eux bien plus qu'à ce que Microsoft est devenue.

Le succès de ce livre m'a valu d'être invité un peu partout pendant plus d'un an, et même si je n'étais pas formé à toute cette exposition médiatique, je considérais que j'avais une responsabilité, une mission d'explication, et aussi celle de proposer une autre vision de la société de l'information : pour une fois qu'on écoutait le message, je n'avais pas le droit de me dérober. C'était pourtant épuisant pour moi et éprouvant pour ma famille : j'avais une fille de deux ans, une épouse brillante aussi impliquée que moi dans la recherche, que je remercie infiniment de m'avoir soutenu et supporté pendant cette période. J'ai donc décidé de m'arrêter en juin 1999 et de revenir à une vie et un rythme plus normaux. Mais c'était sans compter une dernière interview avec Dominique Nora : ce rendez-vous devait durer 30 minutes. Il a duré la journée ! Elle est sortie en me disant qu'il fallait faire quelque chose ! Et c'est comme cela que nous avons écrit *Le Hold-up planétaire : la face cachée de Microsoft*.

En relisant ce livre, je peux dire que même si nous avons une compréhension imparfaite du futur, on y retrouve une analyse lucide et fine des dangers, et notamment du risque d'une maîtrise de la chaîne complète de l'information qui est aujourd'hui enfin devenu un sujet central du débat public.

Le plan de retour à une vie normale avait définitivement échoué : ces ouvrages m'avaient transformé en une sorte d'apôtre du logiciel libre et cela a été dur de ne pas tomber dans le travers des « experts permanents ».

Piège dans le cyberspace

« Une alternative possible : les logiciels libres »

Quand il s'agit de choisir le logiciel à fournir à nos collègues et à nos lycées pour initier nos enfants à l'informatique, on n'est pas obligés de s'en tenir aux cadeaux louches des cybermonopolistes : plutôt qu'un système propriétaire qui se plante très souvent, change tout le temps de version sans raison et dont le code source n'est pas disponible, on peut choisir un système libre ouvert et stable (il faut savoir que contrairement au préjugé populaire, le logiciel libre a eu amplement l'occasion de faire ses preuves [...]) qui permettrait à tout jeune de travailler et apprendre en toute sécurité et à tout esprit curieux d'acquérir une formation informatique avancée et intelligente, car la disponibilité du code source lui permet d'ouvrir le capot et même, s'il le veut, de démonter le moteur.

Et quand il s'agit d'équiper des grands comptes, comme on les appelle, de systèmes informatiques, mieux vaut se fier à du logiciel dont on a le code source et la documentation, qui est constamment vérifié et mis à jour par une communauté techniquement compétente et que l'on peut adapter à ses besoins au moindre coût [...] ».

Retrouvez l'intégralité du texte sur

<http://www.dicosmo.org/Piege/cybersnare/piege.html>

V. S. : *Et en parallèle vous poursuivez votre carrière. Vous passez notamment à l'université Paris 7 sur un poste de professeur...*

R. D. C. : En effet il y a une loi non écrite à Ulm chez les mathématiciens, que l'on essayait d'appliquer aussi en informatique : l'idée qu'il ne faut pas que tous les meilleurs étudiants, qui convergent vers Ulm, soient exposés à la même école et donc que les enseignants tournent. Ainsi les enseignants ne restent pas plus de dix ans à Ulm, au moins en mathématiques, et souvent en informatique. Pierre-Louis Curien a donc décidé de partir à l'université Paris 7 créer en 1999 le laboratoire PPS (Preuves, Programmes, Systèmes), qui est devenu IRIF (Institut de recherche en informatique fondamentale⁴) par une fusion récente avec le LIAFA (Laboratoire d'informatique algorithmique : fondements et applications).

Nous avons été plusieurs à le suivre, et ce changement n'a fait qu'augmenter la charge. J'ai dû construire beaucoup de cours nouveaux et en parallèle faire des centaines d'exposés sur le logiciel libre, les standards ouverts, et l'importance de reprendre la maîtrise d'une technologie informatique qui change le monde. J'ai dépensé une énergie considérable avec mes deux premiers doctorants, Vincent Balat et Jean-Vincent Loddo, pour construire DémoLinux : la première distribution de Linux qui fournissait un environnement graphique complet en démarrant à partir d'un

4. <https://www.pps.jussieu.fr>

lecteur de CD, sans installation. Cela avait le but de donner la possibilité à tous de tester les logiciels libres sans changer leurs machines, mais cela s'est révélé utile dans d'autres occasions, y compris pour les cours à l'université.

Enseignant, chercheur, responsable de formation, vulgarisateur, hacktiviste, tout à la fois, cela ne pouvait pas durer. Il fallait rapidement réconcilier ces différentes casquettes et c'est comme ça que j'ai commencé à orienter ma recherche sur les problèmes scientifiques posés par les logiciels libres.

V. S. : *Quand vous dites que vous vous concentrez sur des recherches dédiées au logiciel libre, cela veut-il dire que les logiciels libres posent des questions particulières différentes des autres logiciels ?*

R. D. C. : Au tout début, beaucoup de collègues étaient surpris quand je tenais ce discours, mais maintenant on sait bien qu'il y a des problèmes scientifiques liés aux modalités de développement logiciel et à la complexité croissante de systèmes construits de façon décentralisée à partir de grandes masses de codes sources : beaucoup de ces problèmes se manifestent seulement, ou plus clairement, quand le code source est disponible et modifiable. Ce que nous entreprenons dans le cadre du projet *Software Heritage* est dans la continuité⁵. Il y a une croissance très rapide de la taille des codes dans le logiciel libre et d'énormes réseaux d'interdépendance. Bien sûr il faut se garder de cette vision illusoire d'une communauté magique qui déciderait démocratiquement. Je ne résume pas le monde du logiciel libre naïvement à l'idée que c'est un système parfait, mais c'est celui qui permet le plus de liberté. De même que la démocratie, c'est une base de construction, même si cela ne garantit pas forcément qu'il en sortira le meilleur.

V. S. : *Vous parlez de la communauté du logiciel libre. Participez-vous aux différentes associations qui se créent pour soutenir le logiciel libre ?*

R. D. C. : J'ai eu l'occasion de rencontrer beaucoup d'acteurs associatifs avec lesquels je garde des contacts étroits et parfois même une véritable amitié. Mais je n'ai pas été directement très impliqué faute de temps et d'énergie, et aussi parce que j'essaie d'être le plus objectif possible dans mes prises de parole : je n'ai aucune peur d'assumer une position atypique ou anticonformiste, et je l'ai montré pendant toute ma carrière, mais j'ai horreur qu'on me range dans une case. Cela a été le cas aussi lors du débat sur l'*open access* dans la loi sur la république numérique, où j'ai

5. « *Our ambition is to collect, preserve, and share all software that is publicly available in source code form. On this foundation, a wealth of applications can be built, ranging from cultural heritage to industry and research* ». Pour un descriptif complet des missions du projet *Software Heritage* voir le site : <https://www.softwareheritage.org/mission/>.

Voir également R. Di Cosmo, *Software Heritage : pourquoi et comment construire l'archive universelle du code source*, 1024 n° 10, avril 2017, pp. 67–72, <http://www.societe-informatique-de-france.fr/bulletin/1024-numero-10/>.

maintenu ma position⁶ jusqu'à la fin : pourtant plébiscitée par les collègues, elle n'a finalement pas été retenue.

En 2004-2005 j'ai introduit les premiers cours sur les logiciels libres dans la maquette de l'université et c'étaient des cours non techniques pour des étudiants en informatique dont je m'étais rendu compte qu'ils ne connaissaient rien au droit des logiciels, à son histoire, etc. Un des premiers projets a été de réaliser une matrice de compatibilité des licences dans le monde du logiciel libre, qui allait de la BSD⁷, qui offre un maximum de liberté, à la GPL⁸, qui dit que la liberté que l'on acquiert doit être compatible avec le fait de la donner aux autres. Cela crée bien sûr des débats par exemple sur quelle licence est au fond la plus libre.

V. S. : *Vous avez aussi une expérience dans l'industrie ?*

R. D. C. : Dans mon rôle de « sage » du logiciel libre, je me suis retrouvé mêlé au montage d'un pôle de compétitivité en 2005-2006, qui est finalement devenu le Groupe Thématique Logiciel Libre dans le pôle Systematic⁹. Pendant deux ans nous avons été plusieurs à porter ce projet, dont je me suis retrouvé finalement responsable afin d'assurer la confiance entre les différents acteurs de la galaxie un peu agitée des entreprises du logiciel libre de l'époque. Là aussi, cela ne correspondait pas à ma mission (le porteur et président d'un pôle doit être normalement un industriel, pas un académique), mais j'ai investi quatre années au service de la communauté du Logiciel libre française, avant de pouvoir enfin passer la main à un président industriel, Stefane Fermigier.

Malgré ce que cet investissement m'a coûté, presque dix ans après, je ne le regrette pas : cela a donné un groupe thématique qui compte sur un comité de pilotage de bénévoles varié et engagé, avec de nombreuses femmes, et on a plus de cent membres (entreprises et laboratoires) qui ont pu faire financer à travers le pôle une cinquantaine de projets de recherche et développement totalement orientés vers le logiciel libre. Nous avons joué un rôle bénévole, mais précieux de *technological gatekeeper* en mettant en relation les différents acteurs de la recherche et de l'innovation dans la région Île-de-France.

V. S. : *J'ai vu aussi que vous aviez reçu un Microsoft Research Distinguished Artefact Award, cela peut sembler ironique !*

6. Voir <http://www.dicosmo.org/FSP>

7. *Berkeley Software Distribution License*.

8. *General Public License*.

9. Voir encart page 48.

R. D. C. : En effet en 2011 avec Jérôme Vouillon nous avons envoyé à la conférence ESEC/FSE¹⁰ un article qui décrit un algorithme sophistiqué lié à certains problèmes dans les logiciels libres, même si nous n'étions pas dans notre domaine traditionnel. Les organisateurs de cette conférence ont été les premiers à noter un problème : trop souvent, les articles décrivaient des résultats obtenus avec des outils logiciels mais ne fournissaient pas le moyen de les vérifier. En 2011, ils ont décidé de créer un prix pour les artefacts logiciels associés aux articles scientifiques. L'idée était de soumettre, si on le souhaitait, l'artefact logiciel, qui serait évalué par d'autres que ceux ayant évalué les papiers. Le jury devait dire si l'artefact *meets expectations, exceeds expectations*, etc. Jérôme a fait un travail spectaculaire, comme d'habitude, et quelques mois après nous avons appris que nous avions remporté le prix du concours et découvert que ce prix était financé par... Microsoft. C'était bien un signe que les temps étaient en train de changer !

Le Groupe thématique logiciel libre (GTLL)

« [...] Le logiciel libre ou *open source* est utilisé de plus en plus largement par les entreprises. Il représente en France un marché de plus de 3 milliards d'euros et en croissance annuelle de l'ordre de 30 %. En fournissant des briques technologiques sur lesquelles il est possible de construire des offres à valeur ajoutée, en accélérant l'émergence et l'adoption de nouveaux standards, en catalysant la collaboration entre des acteurs complémentaires, voire dans certains cas concurrents, le logiciel libre s'est imposé aussi bien dans des secteurs industriels comme le militaire et l'aérospatial, que dans l'informatique d'entreprise, où il joue un rôle primordial dans l'émergence de l'offre autour du *cloud*, du *Big Data*.

La France, et plus spécifiquement l'Île-de-France, possède un leadership reconnu internationalement dans le domaine. Des centaines de PME spécialisées, éditrices ou prestataires de services en logiciel libre, plusieurs grands intégrateurs, et un tissu d'universités et d'écoles d'ingénieurs constituent un écosystème actif qui s'est rassemblé pour constituer le Groupe Thématique Logiciel Libre (GTLL) au sein du Pôle Systematic.

Dans ce contexte, la mission du GTLL est de développer durablement l'écosystème du logiciel libre en Île-de-France, à travers 4 axes d'actions :

- Animer l'écosystème du logiciel libre ;
- Susciter et faire croître des projets innovants sur le logiciel libre (et doter les entreprises de technologies innovantes les rendant compétitives) ;
- Aider à pérenniser et développer les industriels du logiciel libre ;

10. R. Di Cosmo & J. Vouillon, « *On software component co-installability* », in Tibor Gyimóthy and Andreas Zeller, editors, SIGSOFT/FSE'11 19th ACM SIGSOFT Symposium on the Foundations of Software Engineering (FSE-19) and ESEC'11 : 13rd European Software Engineering Conference (ESEC-13), Szeged, Hungary, September 5–9, 2011, pp. 256–266. ACM, 2011. <http://www.dicosmo.org/Articles/coinstallability.pdf>

— Faire progresser la réflexion et communiquer sur les fondamentaux du logiciel libre.

Pour atteindre ces objectifs, le Groupe thématique s'appuie sur la communauté du Logiciel libre francilien, caractérisée par la diversité de ses acteurs : PME, grands groupes industriels et universités et établissements de recherche, associations et individus ».

<http://www.systematic-paris-region.org/fr/logiciel-libre/propos/presentation>

V. S. : *Je vous propose d'en venir au projet qui vous occupe actuellement et que j'ai découvert avec beaucoup d'intérêt, à savoir le projet Software Heritage.*

R. D. C. : On commence au printemps 2014 une série de discussions, d'abord avec Guillaume Rousseau, et un peu plus tard avec Stefano Zacchiroli, autour de la machine à café et de la thèse de l'Irill¹¹. C'est juste une idée vague au départ : on se rend compte qu'on aurait bien besoin d'un endroit unique qui serve de référence à tout le code source de la planète, comme base pour une série d'applications.

Mais c'était juste une idée : dans ce monde il y a plein d'idées qui n'ont rien donné, la difficulté est de les transformer en réalité. D'autant que nous nous sommes dit que cette archive devait forcément exister, que c'était déjà fait.

V. S. : *Par exemple chez Internet Archive ?*

R. D. C. : Je me souviens très bien de la façon dont j'ai découvert Internet Archive. Ted Nelson¹² faisait une tournée en France et il nous a dit : « J'ai un ami qui a lancé un truc... ». C'était le tout début d'Internet Archive, un projet magnifique qui nous permet de revisiter vingt ans d'histoire du Web à travers leur *WayBack machine*. On a donc tout naturellement regardé ce qui était fait pour les logiciels, tant par l'Internet Archive que par d'autres initiatives de préservation numérique.

Mais on a vite compris qu'archiver du code source pose des problèmes spécifiques qu'on ne rencontre pas dans d'autres domaines. Le code source est une connaissance à la fois exécutable par une machine et faite pour être lue et comprise par un être humain (le programmeur). Et il a vocation à changer très rapidement : le code source est vivant, et il est important de savoir ce qu'on y a changé et quand, si on veut comprendre son évolution.

On s'est finalement rendu compte que la préservation du code source avec ses spécificités n'était vraiment au cœur de la mission de personne : on préservait des logiciels exécutables, jouables, des jeux vidéos, c'était notamment fait par Internet

11. Initiative de recherche et innovation sur le logiciel libre.

12. Ted Nelson a forgé le terme *hypertext*. Ce sociologue américain a lancé le projet Xanadu dès les années 1960, rêve de bibliothèque universelle, dans laquelle tous les documents seraient reliés les uns aux autres, préfigurant le Web. Voir notamment : <http://xanadu.com/xUniverse-D6>

Archive qui a une grosse sélection de jeux vidéo. On préservait des pages web qui parlaient de logiciels et de codes sources. Mais les codes sources, comme objet noble, non.

On s'est aussi tourné vers le *Computer History Museum* : Len Shustek¹³, qui préside son conseil d'administration, avait déjà tout compris de ce problème, et écrit sa vision, extrêmement proche de la nôtre, dans un article de 2006 dans *IEEE Annals of the History of Computing*. Mais 2006 était trop tôt pour pouvoir imaginer archiver tout le code source de la planète : le *Computer History Museum* s'est donc spécialisé dans la conservation d'une petite sélection de codes sources de logiciels, analysés en profondeur, avec des interviews des architectes et développeurs.

Nous espérons tisser bientôt des liens avec tous ces acteurs, qui partagent en partie notre mission et notre vision, tout en adoptant des approches différentes, tant sur le plan technique qu'organisationnel.

« When we started Software Heritage we did not expect that publicly available source code was going to be endangered so quickly. Among other major source code hosters, Gitorious¹⁴ was already being taken down in mid 2015, before our initial infrastructure was ready!

Luckily fellow hackers of the Archive Team¹⁵ were already working full speed, and they managed to salvage Gitorious content in time ».

*Gitorious content is safe thanks to the Archive Team,
site Software Heritage, 21/07/2016*

<https://www.softwareheritage.org/2016/07/21/gitorious-retrieved/>

V. S. : *Ce projet est soutenu par Inria. Cela a-t-il été difficile de les convaincre ?*

R. D. C. : Inria est un institut de recherche unique. D'abord parce que c'est un institut entièrement dédié à l'Informatique, mais qui reste à taille humaine et où les décisions peuvent être prises rapidement. Il y a plus de vingt ans, Inria a accepté de devenir un des trois piliers du *World Wide Web Consortium* (W3C), en donnant notamment mission à Jean-François Abramatic.

Il y a un peu plus de deux ans, l'actuel PDG d'Inria, Antoine Petit, a immédiatement compris la portée du projet, et l'adéquation parfaite de l'institut pour l'amorcer, avec la même volonté de donner son envol à une structure qui a vocation à devenir internationale et multipartenaire.

13. Il préside le *Board of trustees* du *Computer History Museum*. Voir <http://www.computerhistory.org/atc/m/author/lshustek/>, Len Shustek, « What Should We Collect to Preserve the History of Software? », *IEEE Annals of the History of Computing*, 2006, vol. 28, n° 4.

14. <https://gitorious.org>

15. http://www.archive-team.org/index.php?title=Main_Page

Et les planètes se sont rapidement alignées : j'ai eu la chance d'attirer dans l'équipe Stefano Zacchirolli, enseignant-chercheur passionné de logiciels libres et ancien leader du projet Debian¹⁶, et puis d'autres passionnés. Jean-François Abramatic, de retour de l'industrie, est devenu président du comité de pilotage, où siègent Serge Abiteboul, Gérard Berry et Julia Lawall. Nous avons à la fois réalisé un prototype, qui montre qu'on prend la chose au sérieux, mais aussi expliqué comment on comptait faire et suscité la confiance. L'équipe actuelle est petite, mais le support précieux d'Inria permet de faire des miracles.

V. S. : *Quels sont vos objectifs en patrimonialisant le code source ?*

R. D. C. : Il y a essentiellement trois objectifs : l'archivage, le référencement et l'étude du code source.

La *mission d'archivage* est claire : il s'agit de ne plus perdre le code source des logiciels qui changent notre société.

La *mission de référencement* est concernée par le problème d'associer à chaque code source un identifiant unique pérenne, et de construire un catalogue de tous les logiciels existants, avec les informations sur leur développement et l'origine depuis laquelle on les a obtenus.

Enfin la *mission d'étude* est probablement celle qui a le plus de potentiel d'impact sur notre futur : nous voulons construire un grand instrument de recherche qui permette d'observer l'évolution du développement du logiciel à l'échelle planétaire ; cela permettra de développer un grand nombre d'applications pour améliorer la qualité et la sécurité de nos logiciels et comprendre comment mieux développer notre code.

V. S. : *Être exhaustif ?*

R. D. C. : Nous pensons que c'est nécessaire. Contrairement au regard que portent par exemple les bibliothèques sur le patrimoine, qui se méfient de l'auto-édition et où l'on préfère sélectionner plutôt que tout archiver, nous avons une optique différente, et cela pour deux raisons. D'abord, techniquement, nous ne sommes pas obligés de faire une sélection : la taille du code source est grande, mais petite par rapport aux images ou aux vidéos. Ensuite, d'un point de vue plus philosophique, c'est vraiment ce que nous voulons faire : pouvoir observer la naissance du code source et le voir grandir. Avec les logiciels libres, c'est possible, on peut voir le code à sa naissance, comme pour un embryon, et il est trop tôt pour savoir si ce sera un paysan ou un roi.

Donc nous archivons tout, et nous pensons que le filtrage se fera *a posteriori* : les logiciels intéressants vont être mentionnés et référencés, et donc émerger d'eux-mêmes, alors que les scories ne seront pas mentionnées et vont lentement décanter.

16. <http://epsilon.cc/~zack/>

Enfin, n'oublions pas l'aspect *Big Data* : on ne sait pas *a priori* ce qu'on voudra chercher dans cette masse de logiciels, peut-être même les raisons des échecs de certains d'entre eux, et pour cela il faut tout garder.

V. S. : *Quels sont les publics envisagés ?*

R. D. C. : Ce projet intéresse plusieurs publics. Ceux qui sont intéressés au patrimoine de l'Humanité bien évidemment découvrent petit à petit qu'une partie grandissante de notre connaissance scientifique, technique, organisationnelle et sociétale est inextricablement embarquée dans les logiciels, et commencent à s'y intéresser. Il faut non seulement le préserver, mais aussi le rendre facilement accessible afin d'améliorer l'éducation.

Ensuite, il y a le monde de la recherche scientifique. Rendre disponible son code source est en train de devenir indispensable pour permettre de comprendre les procédés décrits dans les articles et contribuer à la reproductibilité scientifique. Il faut une archive de référence où retrouver ces codes source : *Software Heritage* est tout désigné pour cela et répond en même temps aux besoins des scientifiques pour lesquels le code est un objet d'étude.

Il y a enfin un très fort intérêt industriel : *Software Heritage* peut remplir le rôle de base de connaissance universelle ouverte sur tous les logiciels libres qui sont massivement utilisés aujourd'hui, avec un grand nombre d'applications à la clé.

V. S. : *Avez-vous déjà une garantie de pérennité du projet ?*

R. D. C. : Pour pérenniser *Software Heritage*, nous souhaitons parler à tous ces publics et les réunir derrière une fondation indépendante, qui va leur permettre de contribuer à l'évolution du projet et de s'y appuyer pour leurs différentes applications. Nous espérons que le nombre de soutiens grandira très rapidement, avec la prise de conscience du fait qu'il s'agit d'une infrastructure essentielle.

Grâce à Inria, nous disposons de quelques années pour y arriver, et on compte déjà des soutiens variés, comme la Société Générale, Huawei, Microsoft, Nokia et DANS.

Mais il n'est pas nécessaire d'être une grande entreprise ou une institution pour aider : notre propre code source est ouvert, et a besoin de pas mal de contributions ; des défis scientifiques nouveaux se posent avec l'arrivée de cette archive unique et ont besoin de chercheurs qui s'y intéressent ; retrouver tout le code source, même ancien, qui est disponible est une tâche complexe et a besoin de personnes qui signalent les sources pertinentes ; expliquer l'intérêt du projet à tous les publics est important, et là il y a de la place pour tous. Enfin, nous savons que nous ne sommes pas à l'abri d'erreurs, et l'on en commettra sûrement, mais en adoptant une approche ouverte, en collaborant avec une multiplicité de publics, nous espérons pouvoir les réduire et les corriger rapidement.



Défendre une éducation qui permet de décoder le numérique

Collectif¹

L'introduction de l'enseignement de l'informatique, à l'école, au collège et au lycée nécessite à la fois la formation assez rapide de tous les enseignants et une formation assez longue de spécialistes. La concomitance de ces deux besoins et des propositions publiques contradictoires en apparence peut parfois prêter à confusion ou donner l'impression que la SIF, qui soutient ces deux approches, se contredit. Il faut encourager les uns à apprendre – donc en leur présentant cela comme quelque chose à la portée de chacun-e et réclamer de la même voix la création de postes d'enseignants ayant suivi une formation longue (CAPES, Agrégation).

Les deux discours sont tout à fait compatibles et compréhensibles simultanément puisqu'il s'agit de régler deux questions différentes. Il est par contre de la responsabilité de chacun de rappeler régulièrement, sans animosité, que si la formation de base des enseignants du primaire et des enseignants de l'ensemble des disciplines doit être proposée de façon à correspondre au temps que ceux-ci peuvent lui consacrer, se satisfaire de ce choix conduirait à la catastrophe. Le niveau à atteindre, pour les enseignants qui ont à enseigner l'informatique au collège et au lycée, doit correspondre à celui qui est demandé pour enseigner les mathématiques, la physique ou la biologie.

Et si le lecteur de 1024 a pu lire dans ces colonnes à plusieurs reprises les positions concernant l'enseignement disciplinaire, il a peut-être eu moins accès à ce

1. Rédaction collective des acteurs des projets Capprio, Class'Code, D-clics numériques et Voyageurs du code.

qu'on peut penser vouloir enseigner à l'école primaire et donc aux éléments de formation des enseignants correspondants. Class'Code et les partenaires de projets liés à ce qu'on appelle la « culture code » nous en disent plus à ce sujet dans le texte ci-dessous.

Colin de la Higuera

La France à décidé que la formation aux enjeux numériques et aux compétences qui y sont liées (informatiques et au-delà) serait proposée à toutes et à tous dès la rentrée 2016.

« Pour relever ce défi. Il y a deux enjeux majeurs : la formation des professionnels de l'éducation (enseignants, animateurs) et celle de la mobilisation de la société civile (faire comprendre la nécessité de cette éducation). En effet, nous vivons une époque qui est traversée par des révolutions technologiques et sociétales. Il faut donc permettre à chacune et chacun de jouir pleinement de sa citoyenneté en maîtrisant fondements et outils des mondes numériques, et en se construisant un esprit critique sur notre société en mutation et la place des mondes numériques. »

Que l'EPI² ou la SIF défendent cela ne surprend pas. Mais qui exprime cette vision ? C'est plus d'une trentaine de partenaires académiques, industriels et de l'éducation populaire (voir Figure 1) qui agissent concrètement au niveau de plusieurs projets du programme d'investissement d'avenir qui ont en commun :



- le partage d'une culture scientifique et technique,
- une démarche de créativité numérique, critique et responsable,
- le développement d'une pédagogie active, bienveillante et participative.

Ces projets, financés par la force publique, répondent de manière concertée et complémentaire aux vrais besoins des éducateurs sur ces sujets et plusieurs milliers de professionnel-le-s de l'éducation commencent à profiter de cette quadruple offre. Class'Code, pour ne citer qu'un exemple, a touché plus de 20 000 personnes³, y

2. Association Enseignement public & informatique, <http://epi.asso.fr>

3. Quelques chiffres sur Class'Code : fin avril, 15 000 personnes sont inscrites sur les formations en ligne liées au primaire et collège (dont plus de 50 % ont profité des formations) et 4 500 inscrites sur les temps de rencontre (dont quelques centaines déjà formées). De plus, 7 500 sont inscrites pour la formation en ligne au MOOC ICN (dont plus de 750 ayant achevé la formation au bout de deux mois). Le site de ressources associé, <https://pixees.fr>, reçoit environ 500 visites par jour (depuis janvier, 50 000



FIGURE 1. Partenaires principaux des quatre projets.

compris au niveau du MOOC⁴ de formation à l'ICN⁵, et cela est vrai d'autres projets⁶, qui ensemble se mettent au service des personnes qui éduquent nos enfants dans ces domaines (voir Figure 2).

sessions de 37 000 utilisateurs, avec un taux de rebond inférieur à 56 % ; par exemple, 4 500 visites de 10 minutes en moyenne sur le module 1 et plus de 12 000 visites sur la page d'accueil, et le bureau d'accueil en ligne gère plusieurs demandes par semaine (plus de 500 demandes depuis le lancement du projet).

4. Extension MOOC-ICN de Class'Code : il y plus 7 500 inscrits au 1^{er} mai ; selon les 500 premières réponses au « questionnaire sur votre profil », il y a 47.53 % d'enseignants, 11 % de salariés d'une entreprise et 10 % de salariés de la fonction publique, 10 % en recherche d'emploi, 8 % d'étudiants et 5 % de retraités... Pour les enseignants du secondaire qui sont directement concernés par ce MOOC, c'est majoritairement des enseignants de mathématiques (46 %) et de physique (15 %), 11 % en STI, 8 % en technologie. À propos d'informatique et de sciences du numérique, ils-elles se déclarent plutôt de niveau débutant (38.40 %) ou intermédiaire (40.49 %), peu en avancé (14.07 %). Le taux de satisfaction mesuré à trois semaines est de 84 % de satisfaits ou très satisfaits, et les retours qualitatifs sont unanimement positifs, plusieurs centaines de messages sur le forum.

5. <https://pixees.fr/mooc-icn-de-linformatique-de-la-creation-du-numerique-des-le-20-fevrier/>

6. Le projet D-Clics numériques c'est aujourd'hui :

- 7 parcours éducatifs clés en main correspondant à 70 heures d'activités accessibles gratuitement, y compris sans compétences préalables ;
- sur le volet formation : 250 formateurs répartis sur tout le territoire métropolitain et en outre-mer, déjà 2 500 professionnels de l'éducation formés dont environ 30 % d'enseignants ;
- sur le volet mobilisation : 2 000 volontaires, bénévoles et réservistes de l'Éducation nationale mobilisés, qui mènent des actions d'inclusion numérique dans 41 départements.



FIGURE 2. Les quatre projets proposent ensemble une offre complémentaire au service de professionnelles et professionnels de l'éducation.

Quelle vision commune et quelle volonté partagée par ces projets ?

Ce qui rassemble ces projets est une vision commune et une volonté partagée de faire avancer l'éducation au numérique. *« C'est-à-dire la nécessité d'un apprentissage du code qui s'intègre dans une éducation plus large par et avec le numérique ; qui aide à une vraie diversité et équité sociale face au numérique et contribue à redonner les libertés de créer et de faire :*

- *L'éducation, en soi, n'est pas et ne peut être un bien marchand.*
- *L'égalité d'accès, de pratique et d'éducation est un défi prioritaire pour une république numérique qui soit réellement inclusive.*
- *Cette égalité doit s'implémenter dans l'accompagnement et la formation de tous les acteurs qui interviennent dans l'École et dans la Cité.*
- *Par sa nature, le numérique redéfinit les rôles et postures de chacun au sein des (techno-)structures, ouvrant la porte à la co-gestion, aux modèles collaboratifs ou ouverts qui les rassemblent. »*

Ces quatre projets disent encore : *« Contribuer à ce que tout le monde soit acteur du numérique, signifie que chaque personne peut atteindre son meilleur niveau de développement, participer pleinement à sa citoyenneté à l'ère numérique, vivre le numérique de manière inclusive au quotidien. C'est aussi développer un esprit critique vis à vis du paysage et des outils numériques, par exemple sur les aspects environnementaux, ou la concentration de nos données dans quelques mains.*

Il s'agit de servir l'intérêt public et de produire un bien commun éducatif : les outils qui y sont produits en commun sont libres et ouverts, et la genèse de ces outils est accessible aux usagers de manière transparente et participative. La collaboration est au centre de cette dynamique. Ces projets s'appuient tant sur des modèles économiques divers et innovants, que sur une démarche au service de l'intérêt général. »

Les valeurs communes⁷ (voir par exemple Figure 3) ?

#équité-sociale, #égalité-des-genres, #solidarité, #citoyenneté.

Nous voilà bien loin de simplement proposer d'apprendre quelques modes d'emploi d'outils numériques ou de « jouer » avec Scratch ou des objets connectés. On parle bien de développer de nouvelles compétences pour décoder et maîtriser le numérique (voir Figure 4).

Et quel positionnement face aux mouvements politiques ?

En ce printemps électoral 2017, tous les mouvements politiques⁸ parlaient d'éducation et de numérique dans leur programme. Mais la vision et les valeurs présentées ici ne se retrouvaient que de manière morcelée dans ces programmes politiques. Pourtant les quatre projets sont à la fois (i) des actions concrètes déjà au service des citoyennes et citoyens, et (ii) des actions transitoires vers une montée en charge de l'éducation au numérique (informatique et littératie numérique) telle que proposée ici.

Ainsi les formations des professionnel-le-s de l'éducation que réalisent ces projets se donnent un objectif précis et limité : permettre aux enseignant-e-s et animateur-e-s d'initier les jeunes à l'informatique, sans oublier la nécessité de cette initiation effectuée à tous âges, et tous les jeunes (donc à travers l'éducation nationale). On y vise ici à former aussi celles et ceux qui seront en interaction avec les futur-e-s professeur-e-s d'informatique.

7. On parle ici d'une vraie égalité des chances (https://fr.wikipedia.org/wiki/%C3%89galit%C3%A9_des_chances#L.27.C3.A9galit.C3.A9_des_chances_en_France:_effet_de_la_loi_de_2006), donc de donner plus pour aider ceux qui en ont le plus besoin.

8. En regardant les programmes des formations politiques (qui ne font pas injure à notre démocratie) qui étaient présentes aux législatives, on relève les éléments suivants :

- Le programme des républicains proposait de mettre en place un CAPES informatique pour développer le recrutement d'enseignants en informatique de niveau master, qui l'enseigneraient au collège, et initier à la programmation, au codage et à l'algorithmique sur le temps consacré à la technologie.
- Du côté du mouvement « En Marche ! », il n'y avait rien d'explicite dans le programme quant à une éducation aux fondements du numérique (on parle uniquement d'éducation avec le numérique), mais l'informatique semblait devoir être liée spécifiquement aux mathématiques.
- Le parti socialiste et ses alliés avaient eux-même mis en place une véritable éducation au code dans le domaine péri-scolaire et scolaire, avec cette volonté d'ouverture d'une initiation à l'informatique à la fois bien identifiée et ouverte sur la multi-disciplinarité et des approches pédagogiques en rupture.
- Le mouvement de la France insoumise promettait une agrégation d'informatique et s'inscrivait dans le mouvement initié par les autres mouvements politiques de l'éducation dite du code, en mettant en priorité la réduction des inégalités.



FIGURE 3. L'informatique c'est aussi pour les garçons !). Dans les faits, ces activités techniques se réalisent sans aucune asymétrie de genre. © vsp.fr.



FIGURE 4. Apprendre à travers un travail de recherche collectif : « si tu m'enseignes je me souviens, si tu m'impliques j'apprends ». © vsp.fr.

Cependant, même une formation minimale⁹ de plusieurs heures n'est possible que si de vrais moyens en temps sont donnés aux enseignants. La mise en place

9. Croire que l'on résout le problème de l'éducation au numérique par une ou deux heures de sensibilisation revient simplement à penser qu'il suffit d'un brevet de secourisme pour exercer la médecine. Ce sont souvent ceux qui n'ont *même pas* leur brevet de secourisme (qui ne mesurent donc pas les choses) qui pourraient le croire.



FIGURE 5. La notion de classe libérée, avec des enfants partie prenante de la façon dont va s'organiser l'activité. © vsp . fr.

d'une formation hybride (en ligne et présentielle) grâce au numérique permet de mieux travailler ensemble, de lisser ce temps de formation mais pas de le contracter.

Il faut réaliser aussi qu'à terme les jeunes ne se satisferont pas d'être (par exemple) initiés à Scratch du primaire au lycée, mais voudront aller plus loin à la fois (i) en informatique, et (ii) au niveau des compétences nouvelles qui émergent avec le numérique.

Pourquoi mélanger informatique et littératie numérique ?

La réponse tient une phrase : parce que l'un est le fondement de l'autre.

Cette vision ouverte de l'apprentissage de l'informatique a de multiples vertus : apprendre autrement (voir Figure 5), mieux comprendre le numérique et se donner les moyens de le co-créeer et pas uniquement le consommer, réduire la fracture culturelle face au numérique, mobiliser les personnes sur ces sujets, et c'est vrai que cela se produit dès l'initiation.

Mais ce n'est pas l'amalgame de l'apprentissage au numérique et avec le numérique. Pour illustrer ce point, prenons l'exemple des jeux sérieux. On apprend l'informatique en jouant, mais il y a une disruption pédagogique entre l'apprentissage avec des jeux et l'apprentissage dit du code :

- D'une part c'est l'enfant lui-même qui crée son propre jeu : en effet, elle ou il sait programmer, faire des objets numériques ; ce jeune se trouve donc en situation complète d'acteur de sa propre formation, dans une démarche de recherche, en travaillant par projet, en position de « *maker* ». On ne consomme pas le jeu, on le fait. On va ainsi apprendre à aller chercher et valider les savoirs et savoir-faire dont on a besoin pour construire les compétences qu'on propose d'acquérir, et à organiser ses connaissances comme cela nous convient au mieux.

— D'autre part, paradoxalement peut-être, on apprend les concepts de l'informatique certes en jouant mais souvent sans ordinateur : ce sont des activités débranchées, il s'agit de se décoller des écrans et de prendre au sens propre comme au sens figuré du recul par rapport au numérique ; le jeu se fait entre humains, souvent avec un scénario ludique, et de manière coopérative, c'est un jeu où on gagne si tout le monde gagne. La machine reprend alors une place d'instrument.

— Enfin, il y a une réelle fracture entre deux modèles économiques : les jeux sérieux sont des objets vraiment coûteux (de l'ordre de 100 k€, parfois plus) qui nécessitent, quelle que soit la démarche de co-construction entre enseignants et développeurs, de laisser à un moment la main, de manière qualitativement plus importante que pour les manuels scolaires, à des personnes qui ne sont pas des professionnels de l'éducation.

*Capprio, Class'Code, D-clics numériques et Voyageurs du code,
en lien avec les autres partenaires des projets cités ici.*



Un enseignement de l'informatique de culture générale pour tous les élèves – Pourquoi et comment

Jean-Pierre Archambault ¹

Entamé il a une quarantaine d'années, le lent et chaotique cheminement de l'enseignement de l'informatique se poursuit dans le système éducatif, au lycée en particulier.

Rappelons qu'il existait dans les années 80 et 90 une option d'informatique dans les lycées d'enseignement général. Présente dans un lycée sur deux et en voie de généralisation, elle fut supprimée une première fois en 1992, rétablie en 1995 puis à nouveau supprimée en 1998. Ce fut alors le temps du « désert explicatif » avec le B2i dont l'échec était prévisible.

Et puis ce furent les actions conjuguées en faveur de cet enseignement, actions qui continuent, de l'EPI (association Enseignement public et informatique), de l'ASTI, de SPECIF puis de la SIF ((Société informatique de France), d'Inria... et de personnalités du monde informatique, comme Gérard Berry, professeur au Collège de France.

En 2007, l'entrevue de l'EPI à l'Élysée enclencha une dynamique nouvelle². Les actions alors menées permettront de modifier le paysage, avec des résultats, la légitimité d'un enseignement de l'informatique étant ainsi reconnue par l'institution éducative.

1. Président de l'EPI (Association Enseignement Public et Informatique). Membre du Conseil d'administration de la SIF.

2. Dépêche de l'AEF du 27 septembre 2007 « L'association EPI défend à l'Élysée l'enseignement de l'informatique à l'École », https://www.epi.asso.fr/revue/editic/aef_jb-jpa.htm

Et il y a eu ce rapport de l'Académie des sciences publié en mai 2013 « L'enseignement de l'informatique en France. Il est urgent de ne plus attendre »³. L'Académie préconise la création dans le secondaire d'une discipline de science informatique pour tous et pour l'ensemble des niveaux. Son enseignement doit débiter dès le collège au même titre que celui de la physique ou de la biologie, après une sensibilisation à l'école primaire. Mais il ne doit pas être à caractère optionnel, puisque ce sont bien tous les citoyens qui sont déjà confrontés à des questions qui ne pourront se résoudre que grâce à une véritable compréhension du monde numérique, rendue possible par une initiation à la science informatique.

1. Un état des lieux

Au lycée, se sont installés progressivement⁴ :

- un enseignement de spécialité optionnel « Informatique et sciences du numérique » (ISN) en Terminale S (rentrée 2012),
- un enseignement d'informatique pour tous les élèves des CPGE scientifiques (rentrée 2013),
- un enseignement d'exploration « Informatique et création numérique » (ICN) en classe de seconde (rentrée 2015),
- un enseignement facultatif « Informatique et création numérique » (ICN) en Premières ES, L et S (rentrée 2016).

Un enseignement optionnel en Terminales ES et L verra le jour à la rentrée 2017.

Et, depuis la rentrée 2016, au collège, un enseignement d'informatique est confié aux professeurs de mathématiques et de technologie : algorithmique, programmation, machines, réseaux. La question de la cohérence des deux enseignements est posée. Et, toujours depuis la rentrée 2016, les élèves peuvent dès le CE1 coder des déplacements à l'aide d'un logiciel de programmation adapté, ce qui les amènera au CE2 à la compréhension et la production d'algorithmes simples⁵.

3. http://www.academie-sciences.fr/pdf/rapport/rads_0513.pdf

4. Voir notamment :

- ISN-TS : BOEN n° 8 du 13-10-2011, http://www.education.gouv.fr/pid25535/bulletin_officiel.html?cid_bo=57572
- ISN au bac : http://www.education.gouv.fr/pid25535/bulletin_officiel.html?cid_bo=57489
- Programme de seconde : <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000030964520>
- Programmes des premières et terminales, BO n° 29 du 21-07-2016, http://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=104657
- BO n° 29 du 21-07-2016, grilles horaires : http://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=104635

5. BOEN n° 11 du 26-11-2015, http://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=94753

Ces avancées doivent se consolider et s'étendre car elles sont encore (très) insuffisantes au regard des enjeux majeurs dans notre société où l'informatique et le numérique sont omniprésents. En effet, il y a là d'abord pour tous les élèves, et donc pour la société et le système éducatif, un enjeu fort de culture générale dont on sait que l'enseignement scolaire relève fondamentalement. Or, paradoxe, les élèves sont encore peu nombreux à rencontrer, lors de leur parcours à l'École, l'informatique en tant que matière scolaire scientifique et technique. Il faut donc un enseignement d'informatique pour tous les élèves.

On parlera ici de l'enseignement général. Dans les enseignements techniques et professionnels, l'informatique et les ordinateurs se sont banalisés depuis une trentaine d'années : les mutations profondes de la production, des entreprises, des administrations, des métiers et des qualifications obligent⁶.

2. Pourquoi faut-il enseigner aujourd'hui la science et technique informatique ?

Pour répondre à cette question, il faut examiner la place prise par l'informatique dans la société au prisme des trois missions traditionnelles de l'École, à savoir former l'homme et la femme, le travailleur et le citoyen.

Le citoyen

Des parallèles sociétaux

Lors des débats sur l'énergie, un citoyen sait « de quoi il retourne » dans la mesure où il peut se référer à ses cours de sciences physiques. Idem pour les débats sur les OGM avec ce qu'il a appris en SVT. Notre quotidien est fait de machines (toujours plus sophistiquées) et de moyens de communication (de plus en plus complexes) qui paradoxalement, parce qu'ils banalisent l'utilisation des ondes ou celle de l'électricité, font écran aux connaissances complexes nécessaires à leur production. Contrairement à des idées reçues en la matière, pourtant, il ne s'agit pas seulement de maîtriser les savoirs technologiques de base pour pouvoir suivre les dernières évolutions des voitures, machines à laver et robots divers mais bien de déterminer les composantes scientifiques qui les sous-tendent et peuvent éclairer les débats actuels

6. Cela concerne notamment les formations suivantes :

- Terminale STMG (sciences et techniques de management et de gestion – bac technologique) : la spécialité SIG (systèmes informatiques de gestion).
- Première et Terminale STI2D (sciences et techniques de l'ingénieur et développement durable – bac technologique) : la spécialité SIN (systèmes d'information et numérique).
- BTS SIO (services informatiques aux organisations), un peu conçus comme suite de la STMG spécialité SIG.
- BTS SN (systèmes numériques) avec l'option IR. Ça donne BTS SN-IR : informatique et réseaux (leur référentiel recouvre *a priori* l'informatique de L1, L2 et L3).

et surtout à venir. Comment espérer maîtriser des évolutions cruciales pour l'avenir sans s'appropriier les éléments déterminants qui font notre environnement et le modifient ? Comment alors accepter que des savoirs indispensables pour structurer les connaissances nécessaires soient abandonnés en cours de scolarité ou relégués à un supplément d'âme ?

Effectivement, comme si l'on pouvait parler avec pertinence des effets des sciences, de ce qu'elles sous-tendent et entraînent, du monde qu'elles contribuent à créer, sans les connaître ! Il faut « ouvrir le capot ». Les sciences ont une contribution spécifique et irremplaçable à l'exercice de la citoyenneté. Il y a en France un réel problème des enseignements scientifiques. Cela étant, le citoyen qui a fait des études secondaires a (encore) des connaissances pour se faire son opinion. Ces apprentissages se sont faits dans un cadre disciplinaire : la physique, la chimie sont des matières scolaires en tant que telles.

Et l'informatique ?

Mais, concernant l'informatique, il n'en va pas de même pour tous les élèves, loin de là. Le numérique est partout. Présent dans une multitude de domaines, il est la numérisation de l'information. Or l'informatique est la science de la représentation et du traitement de l'information numérisée. Elle est la science du numérique comme la biologie est celle de l'étude du vivant. Quid de tous les élèves qui n'ont pas étudié cette science ?

Que signifie « code source » pour quelqu'un qui n'a jamais écrit la moindre ligne de programme ? « Neutralité du Net » pour celui qui n'a aucune représentation mentale d'un réseau informatique ? Le trafic sur Internet ne cesse de croître, une évidence que cette rançon du succès ! La vidéo, gourmande en bande passante, sature les réseaux. Le marché mobile des terminaux explose. Les infrastructures doivent évoluer et se développer, le haut débit en premier lieu. Qui doit payer ? Qui pourrait payer ? Des mesures de discrimination, blocage et filtrage (pour les flux illicites), antinomiques avec la philosophie du Net, sont mises à l'ordre du jour, issues de problématiques comme la lutte contre la cybercriminalité, les modèles économiques de l'immatériel, des industries culturelles. La neutralité du Net rencontre ici les débats qui ont accompagné en leur temps la transposition de la DADVSI, la loi Hadopi... La question se pose également de savoir si le Net est vraiment neutre. Et si la vision d'un cyberspace « idéal » et insensible aux réalités géopolitiques de la planète est réaliste et pertinente.

Ce débat sur la neutralité du Net est un débat de société qui concerne d'évidence tout le monde. Il mêle d'une manière inextricable des questions politiques, économiques, juridiques et des concepts scientifiques et techniques : couches et protocoles de l'Internet, adresses IP, réseaux d'accès et de transit, réseaux privés virtuels, routage, gestion du trafic, interconnexion des réseaux, chiffrement, standards ouverts... S'y impliquer suppose des connaissances scientifiques informatiques, des représentations mentales efficaces.

L'on se souvient que lors des votes sur la transposition de la directive européenne DADVSI et de la loi Hadopi, s'il fut abondamment question de copie privée, de propriété intellectuelle, de modèles économiques...., ce fut sur fond d'interopérabilité, de DRM, de code source, de logiciels en tant que tels. Dans un cas comme dans l'autre on n'a pu que constater un sérieux déficit global de culture informatique largement partagé... La question se pose bien de savoir quelles sont les représentations mentales opérationnelles, les connaissances scientifiques et techniques informatiques qui permettent à tout un chacun d'être en phase et en prise sur la société dans laquelle il vit.

Il est bien d'autres débats sociétaux qui requièrent, eux aussi, une culture informatique : les libertés numériques, le respect de la vie privée, l'espionnage généralisé de la NSA, la protection des données personnelles, le vote avec les ordinateurs et les réseaux, etc. Dans les colonnes du *Monde diplomatique*, en décembre 2002, John Sulston, prix Nobel de médecine, évoquant les risques de privatisation du génome humain, indiquait que « *les données de base doivent être accessibles à tous, pour que chacun puisse les interpréter, les modifier et les transmettre, à l'instar du modèle de l'open source pour les logiciels* ». *Open source*, logiciels libres, code source... Encore une fois, c'est quoi le code source pour quelqu'un qui n'a jamais écrit une ligne de programme ? Le libre est aussi un outil conceptuel qui aide à appréhender les problématiques de l'immatériel. Il suppose une culture générale informatique.

Dans un monde où le quotidien, l'entreprise et les administrations, la société sont massivement faits d'objets informatiques, d'ordinateurs, de réseaux, de logiciels, de données, l'informatique est, pour chaque citoyen, la clé de la compréhension de ce monde numérique en construction. Tout élève doit avoir une culture générale informatique même s'il ne sera pas informaticien. Comme il ne deviendra pas nécessairement médecin, infirmier ou ingénieur électricien.

Le travailleur

L'informatisation est la forme contemporaine de l'industrialisation

La France doit former des techniciens, des ingénieurs et des scientifiques qui sachent développer des outils numériques : il y va de l'innovation et de l'emploi, de la création des richesses. L'informatique est une science et technique majeure du XXI^e siècle (elle représente de l'ordre de 30 % de la R&D de par le monde, 18 % seulement en Europe), omniprésente dans l'entreprise et les administrations. Elle intervient dans l'économie de plusieurs façons essentielles, aux niveaux suivants : la production de biens manufacturés ou agricoles, de par l'automatisation de plus en plus poussée des processus de production ; la création de nouveaux produits ou l'amélioration de produits anciens par l'introduction de puces et de logiciels ; la communication entre les personnes⁷. L'industrie automobile recrute aujourd'hui

7. « L'informatique à l'école : il ne suffit pas de savoir cliquer sur une souris », paru dans *Rue89*, Tribune du 28 juin 2012. Jean-Pierre Archambault, Gérard

plus d'informaticiens que de mécaniciens. On ne fait plus d'essais en vol pour fabriquer des avions. Tous les secteurs d'activités sont concernés : chirurgie, médecine, arts, architecture, droit... On a besoin de spécialistes informaticiens bien sûr – on en manque – mais tous les personnels, peu ou prou utilisateurs au quotidien de l'informatique, doivent avoir la culture qui donne le recul et l'efficacité. Gare aux difficultés de dialogue quand sévit l'illettrisme numérique...

Pour illustrer à nouveau l'omniprésence de l'informatique dans le monde du travail, quelques exemples parmi beaucoup d'autres. Anthropologues, ethnologues, sociologues, géomaticiens... n'ont pas pu échapper ces dernières années à la montée en puissance de l'algorithmique. Alors qu'il a fallu plus de 10 ans et 3 milliards de dollars pour réaliser le premier séquençage humain complet, il est aujourd'hui possible d'en réaliser un en quelques jours pour moins de 1 000 dollars. Ainsi la taille des bases de données génétiques a-t-elle explosé. L'ère du « super-numérique » est bien là⁸. Sans superordinateur, plus de prévisions météorologiques, et pas de modèles de changement climatique. L'innovation sera digitale ou ne sera pas. Des pans entiers de notre économie sont déjà dépendants du calcul intensif. Il est par exemple impensable de se passer de cet outil pour concevoir un avion. Ce n'est donc pas par hasard que la Chine à son tour soit entrée dans la course – et de suite par la grande porte – en fabriquant le supercalculateur le plus puissant au monde. Celui-ci peut réaliser 30 millions de milliards d'opérations par seconde en consommant une puissance électrique équivalant à six TGV à pleine vitesse. L'Inde, la Russie et d'autres ne sont d'ailleurs pas en reste. L'Europe n'a plus que 5 % des parts de marché dans ce domaine.

Des emplois vont disparaître, d'autres vont apparaître (somme nulle, positive ou négative ?), quasiment tous les métiers existants vont être peu ou prou transformés par la numérisation de l'information. Comment espérer exercer les métiers du XXI^e siècle sans avoir une culture générale en informatique, des savoirs appris à l'École avec une discipline en tant que telle, comme pour les autres sciences ?

L'homme et la femme

On entend souvent dire que, l'informatique irriguant la vie quotidienne de tout un chacun, les nouvelles générations, qui baignent dans Internet depuis leur plus jeune âge, n'auraient pas besoin d'une formation spécifique de nature scientifique et technique. Leurs utilisations d'Internet, dans et hors l'école, suffiraient. Qu'en

Berry, Maurice Nivat. <https://www.epi.asso.fr/revue/articles/a1209e.htm>, <http://tempsreel.nouvelobs.com/rue89/rue89-nos-vies-connectees/20120628.RUE0924/l-informatique-a-l-ecole-il-ne-suffit-pas-de-savoir-cliquer-sur-une-souris.html>

8. Didier Schmitt, Commission européenne, « Bienvenue dans l'ère du super-numérique », <http://www.latribune.fr/opinions/tribunes/20140214trib000815451/bienvenue-dans-l-ere-du-super-numerique.html>

est-il exactement ? Dans le cadre de sa thèse de doctorat, Cédric Fluckiger a réalisé une étude dans un collège de la région parisienne⁹. Lucas, élève de troisièmes, pense qu'il est nécessaire d'avoir plusieurs abonnements à Internet pour accéder à toutes les pages, car les moteurs de recherche proposés sur les différents portails n'indiquent pas la même liste de sites : « *Wanadoo ils ont pas les mêmes pages. Si je cherche quelque chose, j'aurai pas les mêmes choses dans Wanadoo et dans quelque part d'autre. (...) Ça change tout, c'est pour ça qu'on en a pris trois différents.* » Cet exemple d'utilisation approximative, qui n'est pas unique loin s'en faut, traduit manifestement une représentation mentale erronée de l'environnement numérique dans lequel le collégien évolue. Des pratiques spontanées et sans recul ne suffisent pas à devenir un utilisateur averti. Une bonne appropriation de notions scientifiques fondamentales est indispensable car elle conditionne une utilisation rationnelle de l'outil conceptuel qu'est l'ordinateur et la résolution des problèmes rencontrés au fil du temps présent et à venir dans la société et l'économie numériques. Il faut relativiser fortement les compétences acquises hors de l'École, qui restent limitées aux usages quotidiens. Elles sont difficilement transférables dans un contexte scolaire plus exigeant. Les pratiques ne donnent lieu qu'à une très faible verbalisation. Les usages reposent sur des savoir-faire limités, peu explicites et laissant peu de place à une conceptualisation. Il ne faut pas confondre « consommation » et « création » d'informatique, utilisation « intelligente » des outils.

Pour leur vie de tous les jours aussi, les élèves doivent donc acquérir des connaissances en informatique à l'École, le seul lieu où ils rencontrent la connaissance sous une forme structurée. L'École doit leur donner une culture générale informatique.

À méditer

En décembre 2013, le président Obama appelant tous ses compatriotes à étudier la programmation¹⁰, déclarait « *Ne vous satisfaites pas de l'achat d'un nouveau jeu vidéo : fabriquez-en un !* » leur disait-il. C'est clair : utilisateur mais aussi créateur. Avec un enjeu fondamental, pour le moins à méditer : selon le président des États-Unis, « *l'apprentissage des compétences en jeu n'est pas seulement utile pour le futur des jeunes, il est important pour le futur de notre pays. Si nous voulons que les États-Unis restent en tête, nous avons besoin que de jeunes Américains comme vous aient la meilleure maîtrise possible des outils et de la technologie* ». Rester la première puissance mondiale... Quoi qu'on en pense par ailleurs, vraiment à méditer pour ceux qui ne seraient pas convaincus de l'importance de la chose...

9. Voir « Internet et ses pratiques juvéniles », Cédric Fluckiger, Médialog n° 69, <http://www.epi.asso.fr/revue/articles/a0905d.htm>

10. « Le président Obama appelle tous les Américains à étudier la programmation », <https://www.epi.asso.fr/revue/lu/11312p.htm>

3. L' « essence » des disciplines enseignées

Autre manifestation de l'omniprésence de l'informatique et qui concerne aussi l'enseignement au premier chef : elle contribue à faire évoluer l'« essence » des autres disciplines (objets, méthodes, outils). Prenons l'exemple de la biologie. En une phrase, Ernest Rutherford (1871-1937) énonce la suprématie de la physique en son siècle, en l'opposant à la biologie : « *En sciences, il y a deux choses : la physique et la collection de timbres.* » À l'époque, la technique n'avait pas sa place et les mathématiques se justifiaient par leur utilité en physique. Aujourd'hui, la biologie modélise ses objets en insistant sur leur caractère algorithmique (il faut former les élèves à la pensée informatique qui ne se confond pas avec la pensée mathématique). Les nouveaux déploiements de la démarche scientifique accordent une place importante au choix du langage utilisé pour décrire les objets. Le séquençage du génome, par exemple, repose sur une abstraction nouvelle de ce qu'est un brin d'ADN ou d'ARN : une suite finie à valeurs dans un ensemble à quatre éléments – les informaticiens disent « un mot dans un alphabet de quatre lettres ». Cette question de la représentation informationnelle des objets est une question récurrente en informatique. La biologie s'est dotée d'un appareillage théorique qui se fonde beaucoup sur l'informatique.

De fait, cette évolution est peu ou prou le cas pour toutes les disciplines. Peut-on par exemple imaginer aujourd'hui la géographie sans les systèmes d'information géographique ou la linguistique sans le traitement automatique de la langue ? Ou l'astronomie sans ses pipelines de calculs informatiques, ou la génomique sans ses algorithmes d'analyse de séquences ADN ? Numérisation, bases de données, hypertexte et Internet investissent les sciences humaines. Le théorème des « quatre couleurs » (qui suffisent pour colorer n'importe quelle carte) a été démontré par ordinateur.

Quelque part, à des degrés très diversifiés, l'enseignement des disciplines scolaires doit en tenir compte. Pas de sciences expérimentales sans simulation et EXAO. C'est particulièrement vrai pour les enseignements techniques et professionnels où le traitement de texte s'est substitué à la machine à écrire, la base de données au fichier-carton, le logiciel de DAO à la planche à dessin, la machine à commandes numériques à l'étau-limeur, etc.

4. Le lent et chaotique cheminement de l'enseignement de l'informatique

Les applications de la science informatique sont partout dans la société mais c'est loin d'être le cas pour cette science dans l'enseignement scolaire. Un paradoxe.

Contrairement à ce que l'on pourrait penser de prime abord, ce paradoxe ne saurait surprendre. Il est même la loi du genre dans tous les domaines. C'est un grand classique : le nouveau émerge toujours dans la douleur. Et cela ne date pas d'hier.

Déjà, Confucius mettait en garde : « *Lorsque tu fais quelque chose, sache que tu auras contre toi ceux qui voulaient faire la même chose, ceux qui voulaient faire le contraire et l'immense majorité de ceux qui ne voulaient rien faire.* » Au début du XX^e siècle, un lobby du courant continu s'évertuait à « prouver », force arguments « scientifiques » à l'appui, que le courant alternatif constituait une impasse...

La création d'une nouvelle discipline scolaire s'avère toujours difficile. On se souvient de discours des années 80 qui, péremptaires, affirmaient que l'informatique était une mode qui passerait comme passent les modes, que la France n'avait plus besoin d'informaticiens. On pourrait multiplier sans fin les exemples... une véritable litanie. Cela étant, on n'est pas obligé de persévérer dans l'erreur. L'enjeu pour la société française est primordial.

Discipline ou pas

Depuis de nombreuses années, concernant le contenu de la culture générale informatique et les modalités pour la donner à tous les élèves, deux approches coexistent. Discipline ou pas ? Pour l'une, les apprentissages doivent se faire d'une manière exclusive à travers les usages de l'outil informatique dans les différentes disciplines existantes. Pour l'autre, l'informatique étant partout, car elle sous-tend le numérique, elle doit être quelque part en particulier, à un moment donné, sous la forme d'une discipline scolaire en tant que telle. Pour les uns, l'utilisation des technologies de l'information et de la communication (TIC) suffit. Pour les autres, l'usage d'un outil, matériel ou conceptuel, ne suffit pas pour le maîtriser, ni pour donner l'indispensable culture générale de notre époque. Le B2i a traduit l'approche par les « usages ». L'examen de plus de 10 ans de mise en œuvre du B2i, ou d'essai de fait limité de mise en œuvre, montre un échec patent, un échec prévisible.

Apprendre à utiliser un outil ne suffit pas. Pour André Leroi-Gourhan, l'outil n'est rien sans le geste qui l'accompagne et l'idée que se fait l'utilisateur de l'outil de l'objet à façonner. Et d'ajouter : « *Ce qui était vrai de nos lointains ancêtres du Néandertal, quand ils fabriquaient des lames de rasoir en taillant des silex, est toujours vrai : l'apprentissage de l'outil ne peut se faire sans apprentissage du geste qui va avec ni sans compréhension du mode de fonctionnement de l'outil, de son action sur la matière travaillée, ni sans formation d'une idée précise de la puissance de l'outil et de ses limites.* » Cela vaut pour tous les outils, notamment les outils numériques omniprésents dans la vie d'aujourd'hui, et encore beaucoup plus de demain.

5. Les programmes scolaires et la culture générale scolaire

Les programmes scolaires traduisent pour le système éducatif la culture générale d'une époque. Ils fabriquent ainsi une culture générale scolaire. S'il y a des fondamentaux qui demeurent, la culture générale scolaire n'en est pas pour autant immuable. Elle évolue car la société évolue. C'est pour cela qu'il faut aller vers un enseignement de l'informatique pour tous les élèves. Ce qui n'exclut évidemment

pas l'utilisation de l'« outil » informatique dans les différentes disciplines et activités. Faut-il répéter une fois de plus que les deux démarches sont complémentaires ?

Le latin et le grec n'occupent plus la place qu'ils avaient antan. En mathématiques, la géométrie descriptive et les coniques ont disparu, remplacées par les probabilités et les statistiques. Dans les années 1960, la discipline sciences économiques et sociales a été créée, etc.

L'emploi du temps des élèves

Çà et là on nous dit que l'emploi du temps des élèves est déjà très chargé. Mais l'allongement de la scolarité est une tendance lourde. Il est loin le temps où la majorité d'une génération quittait l'école à 10, 12 ans. La marge de manœuvre est réelle pour sortir du cercle (de plus en plus) vicieux qui veut que comme c'est difficile d'introduire une nouvelle discipline, eh bien on ne le fait pas. Si l'on appliquait ce principe dans le futur, on est certain que dans 1000 ans les enseignements seraient les mêmes qu'aujourd'hui ! Et l'on peut essayer d'imaginer ce que serait l'enseignement en 2014 si les pédagogues du passé avaient développé et mis en œuvre des thèses similaires. Les élèves d'aujourd'hui apprendraient, comme dans les écoles médiévales, d'abord la Grammaire, la Dialectique et la Rhétorique, puis l'Arithmétique, la Musique, la Géométrie et l'Astronomie. Pas de Physique, pas d'Histoire, pas de Biologie !

Culture générale et formation professionnalisante

Les disciplines scolaires de l'enseignement général concernent tous les élèves. Elles donnent la culture générale, des fondamentaux que les élèves retrouveront d'une manière diversifiée dans leurs formations ultérieures et leur vie professionnelle. Toutes les disciplines sont des outils au service des autres, et aussi des fins en soi. Cela vaut par exemple pour les mathématiques qui sont au service des sciences physiques ou des sciences économiques ou... Et pour l'informatique bien sûr. Plus les disciplines sont au service des autres, plus elles deviennent une fin en elles-mêmes. Plus elles sont des composantes majeures de la culture des hommes et des femmes. Informatique et littérature même combat ! Écrire un programme ou écrire un texte sont deux activités d'égale dignité, tout aussi passionnantes l'une que l'autre : une fin en soi !

Une formation structurée sur une longue durée doit être organisée comme une fusée à deux étages : les premières années doivent être consacrées à l'apprentissage de savoirs fondamentaux, puis doivent venir les savoirs spécialisés, qui ont vocation à être directement utilisés dans les premières années d'une activité professionnelle. Par exemple, la formation d'un médecin consiste à apprendre d'abord (dès l'école primaire, le collège et le lycée) des généralités sur l'anatomie et la physiologie humaine, avant d'apprendre tel ou tel geste chirurgical ou la posologie de tel ou tel médicament. Cette seconde phase de la formation est très variable en fonction du métier que l'on souhaite exercer : les mêmes savoirs spécialisés ne sont pas nécessaires à

un ophtalmologiste et un anesthésiste, alors que l'un et l'autre doivent savoir que le cœur est à gauche et le foie à droite ou qu'une cellule humaine contient vingt-trois paires de chromosomes.

Si l'enseignement scolaire, à l'école, au collège et au lycée, relève fondamentalement de la culture générale – certains donnant l'impression de ne plus savoir ce que c'est alors qu'ils en ont bénéficié pendant leur scolarité –, il a bien aussi un rôle de propédeutique aux formations professionnalisantes.

6. Quoi enseigner ? Des considérations d'ordre pédagogique

Explicitons les relations entre la culture générale et la culture générale scolaire (n'oublions pas par ailleurs l'importance de la précocité des apprentissages).

Un détour par les mathématiques

À notre époque, par exemple, tout un chacun doit savoir lire une courbe ou un graphique (ainsi la représentation de l'évolution du chômage, de la décélération de son accélération si c'est le cas) et, pour cela, savoir qu'une grandeur peut dépendre d'une autre grandeur. Il étudiera donc les fonctions en cours de mathématiques. Il apprendra des notions comme la continuité, la dérivation. Il résoudra des équations, activité à laquelle, en général, il ne se livrera plus dans sa vie d'adulte, sauf dans le cas où son métier l'exige.

Restons sur les fonctions. L'actualité de la campagne présidentielle fait que le citoyen doit pouvoir se faire son opinion sur les causes de la crise financière latente. Vaste problème. Dans une interview au journal *Le Monde*, le 19 octobre 2009, Benoît Mandelbrot, père de la théorie des fractales, indiquait qu'« *il était inévitable que des choses très graves se produisent* ». Dès 1964, il avait perçu que les modèles mathématiques utilisés par les financiers étaient erronés et avait tenté d'alerter sur leurs dangers. « *Les gens ont pris une théorie inapplicable... Elle ne prend pas en compte les changements de prix instantanés qui sont pourtant la règle en économie. Elle met des informations essentielles sous le tapis. Ce qui fausse gravement les moyennes. Cette théorie affirme donc qu'elle ne fait prendre que des risques infimes, ce qui est faux...* » S'il est difficile au citoyen lambda de discuter sur le fond ce point de vue, il doit néanmoins comprendre la problématique posée, se faire son opinion dans un débat contradictoire. B. Mandelbrot ajoute que « *les catastrophes financières sont souvent dues à des phénomènes très visibles mais que les experts n'ont pas voulu voir* ». Savoir pourquoi il en va ainsi est une question fort intéressante... Un certain niveau de culture mathématique constitue d'évidence une condition d'exercice de la citoyenneté. Il vaut mieux avoir appris à étudier des fonctions à l'École que de ne point l'avoir fait. Il ne s'agit bien évidemment pas pour le citoyen lambda d'être en mesure de mener un débat sur un pied d'égalité avec des spécialistes de haut niveau, mais de percevoir de quoi on parle, d'avoir des intuitions et de pouvoir se faire une opinion dans le cadre d'un débat pluraliste. Plus on en sait mieux ça vaut.

Et rappelons que la démocratie est le gouvernement des affaires de la cité par des « ignorants » !

La formule est bien connue selon laquelle la culture est ce qui reste quand on a tout oublié. Effectivement, à l'École, on apprend des choses dont, ultérieurement, on ne se servira pas tous les jours, loin de là, ou même dont on ne servira plus mais qui, dans des processus d'apprentissage, auront été nécessaires pour avoir une bonne culture générale.

La programmation, excellent outil pédagogique

Et il y a même une cerise sur le gâteau pédagogique. La programmation est un excellent outil pédagogique au service des disciplines, des enseignants et des élèves. Elle favorise l'activité intellectuelle. En effet, on constate une transposition des comportements classiques que l'on observe dans le domaine de la fabrication des objets matériels. À la manière d'un artisan qui prolonge ses efforts tant que son ouvrage n'est pas complètement terminé, un lycéen, qui par ailleurs se contentera d'avoir résolu neuf questions sur dix de son problème de mathématiques (ce qui n'est déjà pas si mal !), s'acharnera jusqu'à ce que fonctionne le programme de résolution de l'équation du second degré que son professeur lui a demandé d'écrire, pour qu'il cerné mieux les notions d'inconnue, de coefficient et de paramètre. Ce surcroît d'activité se révèle être très précieux pour des apprentissages solides. Il enrichit la panoplie des outils pédagogiques de l'enseignant.

7. Des contenus scientifiques

Nous avons vu comment la culture générale scolaire en mathématiques (les fonctions) contribue à l'exercice de la citoyenneté (lecture d'une courbe ou compréhension des crises financières). Le détour est certes grand, mais incontournable, et l'effet lointain. L'élève doit savoir faire confiance à son professeur. Poursuivons. Regardons ce que fait l'École, comment, pour former l'homme et la femme, le travailleur et le citoyen, elle leur donne une culture générale dans les autres domaines de la connaissance. Depuis longtemps, nous savons qu'il est indispensable que tous les jeunes soient initiés aux notions fondamentales de nombre et d'opération, de vitesse et de force, d'atome et de molécule, de bactérie et de virus, de chronologie et d'événement, de genre et de nombre, etc. Pour différentes raisons. La thermodynamique, la mécanique, l'électricité, la chimie sous-tendent les réalisations de la société industrielle, c'est pour cela que les sciences physiques sont devenues discipline scolaire il y a une centaine d'années. Cela concerne effectivement les futurs spécialistes. Mais tout le monde ne sera pas technicien ou ingénieur. En revanche, tout le monde a besoin d'une culture de base en la matière. Au travail mais aussi dans le quotidien car il faut connaître l'environnement moderne. Se connaître aussi, savoir de quoi est fait l'être humain et comment son corps fonctionne, même si, répétons-le, tout le monde n'est pas médecin ou infirmier ou infirmière. Et il y a les débats de société

que nous avons déjà mentionnés. Ces initiations se font dans un cadre disciplinaire. Il est indispensable aujourd'hui d'initier tous les élèves de la même façon aux notions centrales de l'informatique, devenues tout aussi indispensables : celles d'algorithme, de langage et de programme, de machine et d'architecture, de réseau et de protocole, d'information et de communication, de données et de formats, etc. Cela ne peut se faire qu'au sein d'une vraie discipline informatique. Pas d'incompréhensible sort fait à la science informatique au nom d'on ne sait trop quoi¹¹.

Les enseignements à tirer de l'échec du B2i

En effet, l'expérience a montré l'échec du B2i et de son approche exclusive par les usages. Échec prévisible disions-nous. En effet, un tel dispositif suppose implicitement un apport de connaissances mais ne dit pas où les trouver, dans quelles disciplines. Il faut « nommer » les connaissances que l'humanité a pu mettre des siècles ou des millénaires à élaborer. Il n'est déjà pas évident d'organiser des apprentissages progressifs sur la durée lorsque les compétences recherchées sont formulées de manière très générale (du type « maîtriser les fonctions de base » ou « effectuer une recherche simple »), éventuellement répétitives à l'identique d'un cycle à l'autre, et que les contenus scientifiques, savoirs et savoir-faire précis permettant de les acquérir, ne sont pas explicités. Mais, quand, en plus, cela doit se faire par des contributions multiples et partielles des disciplines, à partir de leurs points de vue, sans le fil conducteur de la cohérence didactique des outils et notions informatiques, par des enseignants insuffisamment formés, on imagine aisément le caractère ardu de la tâche au plan de l'organisation concrète. Pour se faire une idée de ces difficultés, il suffit d'imaginer l'apprentissage du passé composé et du subjonctif qui serait confié à d'autres disciplines que le français, au gré de leurs besoins propres (de leur « bon vouloir »), pour la raison que l'enseignement s'y fait en français. Idem pour les mathématiques, outil pour les autres disciplines, avec les entiers relatifs enseignés lors de l'étude de la période avant-après J.-C. en histoire et les coordonnées de celle de la latitude et la longitude en géographie ! Répétons-le, les pratiques seules ne suffisent pas à maîtriser un outil.

Les contenus informatiques à enseigner

11. « La science informatique doit être enseignée dès le secondaire au même titre que la physique ou la biologie », paru le 12 novembre 2013 dans *Le Monde Informatique*, Jacques Baudé, <http://www.lemondeinformatique.fr/actualites/lire-la-science-informatique-doit-etre-enseignee-des-le-secondaire-au-meme-titre-que-la-physique-ou-la-biologie-55631.html>

On pourra se référer au programme de l'enseignement de spécialité optionnel « Informatique et sciences du numérique » de Terminale S¹² ainsi qu'au manuel scolaire correspondant édité par Eyrolles¹³.

Dans le cadre des actions menées en faveur de l'enseignement de l'informatique, notamment lors de rencontre avec les responsables du ministère de l'Éducation nationale, l'EPI et la SIF ont été « forces de proposition », répondant à des « commandes » du MEN. On pourra consulter des propositions de programme pour les trois années du lycée, le collège et l'école primaire¹⁴.

Une proposition de formation à l'intention des professeurs habilités à enseigner dans l'option de spécialité de Terminale S a également été faite¹⁵. Un ouvrage correspondant « Introduction à la science informatique » a été édité par le CRDP de Paris¹⁶.

12. « Enseignement de spécialité d'informatique et sciences du numérique de la série scientifique – classe terminale », *Bulletin officiel* spécial n° 8 du 13 octobre 2011, http://www.education.gouv.fr/pid25535/bulletin_officiel.html?cid_bo=57572

13. *Un manuel scolaire pour ISN en Terminale S*, https://wiki.inria.fr/wikis/sciencinfolycee/images/7/73/Informatique_et_Sciences_du_Num%C3%A9rique_-_Sp%C3%A9cialit%C3%A9_ISN_en_Terminale_S.pdf, et https://wiki.inria.fr/sciencinfolycee/Informatique_et_Sciences_du_Num%C3%A9rique_-_Sp%C3%A9cialit%C3%A9_ISN_en_Terminale_S

14. Communiqué de l'EPI : « Audience de l'association Enseignement Public et Informatique au cabinet du Ministre de l'Éducation nationale », <http://www.epi.asso.fr/revue/articles/a0712f.htm>
 « Pour un enseignement de l'Informatique et des Technologies de l'Information et de la Communication au lycée », Groupe ITIC de l'ASTI, *Epinet* n° 100, décembre 2007, http://www.epi.asso.fr/revue/editic/asti-itic-txt_0711.htm

« Proposition d'orientations générales pour un programme d'informatique à l'école primaire », Serge Abiteboul, Jean-Pierre Archambault, Gérard Berry, Colin de la Higuera, Gilles Dowek, Maurice Nivat, http://www.epi.asso.fr/revue/editic/itic-ecole-prog_2013-12.htm

« Esquisse d'un programme d'informatique pour le Collège », Serge Abiteboul, Jean-Pierre Archambault, Gérard Berry, Colin de La Higuera, Gilles Dowek, Maurice Nivat, <http://www.epi.asso.fr/revue/docu/d1402a.htm>

15. Communiqué de l'Association Enseignement Public et Informatique (EPI), *Epinet* n° 124 d'avril 2010, <http://www.epi.asso.fr/revue/docu/d1004a.htm>

« Proposition de programme de formation pour les enseignants chargés de la spécialité *Informatique et sciences du numérique* en terminale S », http://www.epi.asso.fr/revue/editic/asti-itic-prog-prof_1004.htm

16. « Introduction à la science informatique », <https://framablog.org/2011/09/06/manuel-informatique-sciences-numeriques/>, <https://framablog.org/2011/09/18/manuel-science-informatique-commentaires/>, <http://www.cndp.fr/crdp-paris/Introduction-a-la-science>

8. La formation des enseignants : une question décisive

Si des avancées ont été réalisées, il reste une question décisive en suspens : la formation des enseignants. Comment espérer donner une culture informatique scientifique et technique sans professeurs spécialisés, sans professeurs d'informatique. Formés comme leurs collègues des autres disciplines (pendant cinq années et non en quelques dizaines d'heures, quand ils sont formés, ce qui est loin d'être toujours le cas), c'est-à-dire titulaires d'un Capes ou d'une agrégation d'informatique qu'il faut créer sans plus attendre.

Cette création rapide (concours externes et internes, listes d'aptitude) constituerait un signal fort au plan politique traduisant la prise de conscience, l'engagement, la volonté d'aller résolument vers un enseignement pérenne de culture générale informatique pour tous les élèves. De ce point de vue, la création d'une option informatique dans le Capes de mathématiques (50 % de mathématiques et 50 % d'informatique), en 2017, si elle constitue une avancée, n'est pas à la hauteur des enjeux.

Pendant une période transitoire, la formation continue dans les ÉSPÉ, à renforcer, doit aussi, de manière complémentaire, faire office de formation initiale avec des habilitations du type de celles prévues pour les enseignants d'ISN, des certifications (certifications aussi pour les professeurs des écoles). Il faut que ces formations aient des retombées positives pour la carrière des enseignants. Une reconnaissance claire d'un bénéfice pour les enseignants suivant ces formations sera indispensable pour construire leur motivation. Une organisation institutionnelle serait, par ailleurs, de nature à favoriser la reconversion dans l'enseignement des professionnels de l'informatique qui le désirent.

Tous les enseignants

Il y a les professeurs qui enseignent l'informatique mais tous les enseignants sont concernés et doivent être formés, avec des objectifs distincts. Ils doivent avoir une bonne culture générale informatique. Cela suppose de préciser les choses. En effet, les débats sur la nécessité d'une discipline scolaire informatique sont souvent allés de pair avec une certaine confusion sur les statuts éducatifs, divers et différents, de l'informatique et des TIC. Il est beaucoup question d'usages et d'utilisations pertinents, nécessaires, souhaitables... mais qui ne peuvent pas suffire à donner la culture générale scientifique et technique dont tous les élèves ont besoin, comme l'expérience l'a montré. L'informatique est à la fois¹⁷ :

- objet d'enseignement ;
- outil pédagogique transversal ou spécifique à une discipline, dont on ne dira jamais assez la complémentarité avec le statut précédent, les deux se renforçant mutuellement ;

17. « La diversité de l'informatique à l'École », Jean-Pierre Archambault, <http://www.epi.asso.fr/revue/articles/a1402b.htm>

- facteur d'évolution des disciplines enseignées, de leur « essence » (objets, méthodes et outils); c'est plus ou moins le cas pour toutes les disciplines et particulièrement vrai pour les enseignements techniques et professionnels ;
- outil de travail personnel et collectif des élèves, des enseignants et de la communauté éducative dans son ensemble.

La diversité de l'informatique à l'école, ses statuts éducatifs amènent à distinguer les profils de formation suivants :

- l'ensemble des enseignants pour qui c'est une formation à l'exercice de leur métier ; avec deux niveaux, les enseignants et les formateurs ;
- les enseignants d'une discipline donnée (peu ou prou, toutes les disciplines, d'une manière spécifique, à donner dans un cadre disciplinaire) ;
- les professeurs de la discipline scientifique et technique informatique.

Les enjeux sont essentiels pour l'avenir de tous et l'avenir du pays. Il est vraiment urgent de ne plus attendre.



L'informatique en CPGE

Judicaël Courant ¹

Nous décrivons l'organisation de l'enseignement de l'informatique en classes préparatoires aux grandes écoles (CPGE). Si l'enseignement d'informatique commun à toutes les filières scientifiques (hors filières à dominante biologique) mis en place en 2013 a constitué un net progrès vers la mise en place d'un véritable enseignement informatique de qualité, on verra que l'informatique garde encore une place très marginale en classes préparatoires. Nous concluons par quelques suggestions pour qu'elle prenne véritablement la place qu'elle mérite.

1. Introduction

Ce texte décrit l'organisation de l'enseignement de l'informatique en CPGE. On peut classer les classes préparatoires dans les quatre catégories suivantes :

- (1) littéraires (classes A/L et B/L),
- (2) préparatoires économiques et commerciales (ECE, ECS, ECT),
- (3) à dominante biologique (BCPST, TB),
- (4) scientifiques, hors BCPST et TB (on trouvera Figure 1 un schéma des différentes filières scientifiques et en annexe A la signification des sigles des différentes filières). Les classes à étoile (MP*, PC*, PSI*, PT*) préparent en pratique aux concours les plus difficiles mais ont les mêmes programmes et

1. Docteur en informatique et agrégé de mathématiques sur poste informatique au lycée la Martinière-Monplaisir de Lyon.

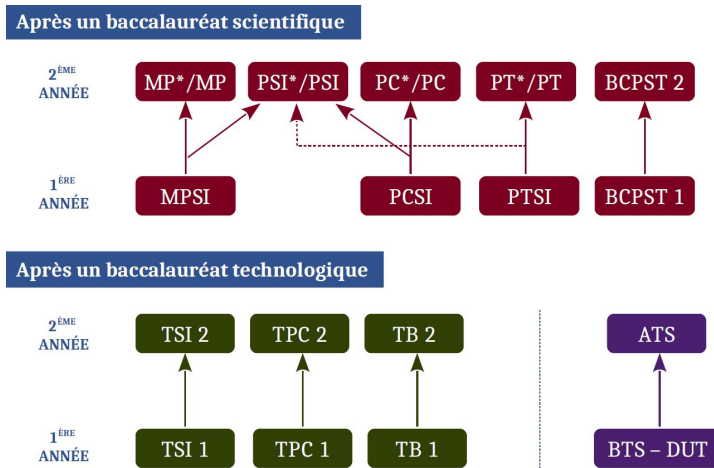


FIGURE 1. Filières scientifiques depuis 1995 [4]

horaires que les classes de même nom sans étoile (MP, PC, PSI, PT). Nous ne les distinguerons donc plus dans le reste de ce texte.

La classe d'ATS dure un an et s'effectue après un BTS ou un DUT. Toutes les autres classes (scientifiques ou non) se déroulent sur deux années, chacune découpée en deux semestres.

À l'intérieur de chacune de ces quatre grandes catégories, l'enseignement de l'informatique est très similaire voire le même.

Nous présenterons rapidement Section 2 la situation dans ces quatre grandes catégories, avant de montrer, Section 3, en quoi le tronc commun des filières scientifiques (hors BCPST et TB) est un net progrès pour l'enseignement de l'informatique en CPGE. Nous montrerons Section 4 que l'informatique en CPGE reste malheureusement encore une discipline avec un statut très marginal, avant de proposer Section 5 quelques pistes pour la faire évoluer.

2. Vue d'ensemble des programmes

2.1. Lettres : pas d'informatique

L'informatique est absente des filières littéraires : aucun des mots « informatique », « algorithme », « programmation » n'apparaît dans les programmes officiels ; dans le programme de mathématiques de la classe B/L, la seule évocation d'un algorithme se trouve dans le programme de mathématiques : il s'agit de la méthode du

pivot de Gauss, qui est appliquée à la recherche d'une forme triangulaire, de l'inverse d'une matrice carrée et des solutions d'un système linéaire.

Il est à noter cependant que les « objectifs de la formation » A/L comportent le développement de « l'aptitude [...] à se servir des instruments et des ressources numériques dans une perspective de construction, d'appropriation et de partage des connaissances », sans que l'on trouve cependant dans le reste du programme de moyen explicite de parvenir à cet objectif.

2.2. Prépas commerciales : informatique dans les maths

L'enseignement de l'informatique est intégré à l'enseignement de mathématiques en classes économiques et commerciales et est réparti sur les quatre semestres des deux années de ces prépas. À travers l'utilisation de Scilab, il s'agit essentiellement de modéliser, simuler et visualiser des problèmes liés à des questions de statistique (simulation de lois aléatoires, méthode de Monte-Carlo). Le savoir-faire algorithmique le plus difficile qui soit exigible paraît être le calcul de l'encadrement du zéro d'une fonction par dichotomie.

2.3. Prépas à dominante bio : un programme spécifique

L'enseignement de l'informatique fait l'objet d'un programme spécifique en classes à dominante biologique [2] : l'objectif affiché est « d'introduire puis de consolider les concepts de base de l'informatique, à savoir l'analyse et la conception [...] des algorithmes et la question de la représentation des données ». Il est réparti sur les quatre semestres. Concrètement :

- cet enseignement utilise le langage Python ;
- une unique structure de données est introduite : le tableau dynamique (tel que Python le propose) ;
- les algorithmes qu'il s'agit de savoir programmer sont très simples : recherche séquentielle et tri par insertion ; la recherche de la médiane est mentionnée mais il faut probablement comprendre qu'il s'agit de la rechercher par essais successifs des différentes valeurs ou par tri préalable des valeurs ;
- des algorithmes plus évolués sont présentés : tri par pivot, algorithme de Dijkstra.

Cet enseignement est pour la majorité dispensé sous la forme de la réalisation d'un projet informatique. Les projets typiques réalisés par des binômes de deuxième année vont de la simulation de la propagation d'un feu de forêt ou d'une épidémie pour les plus simples à des jeux mettant en œuvre des algorithmes évolués comme des parcours en largeur de graphes, *union-find*, minimax ou negamax pour les plus complexes (certains projets utilisant même plusieurs d'entre eux).

On n'attend cependant des étudiants aucune considération de complexité algorithmique.

2.4. Autres filières scientifiques : tronc commun et option

2.4.1. Le tronc commun informatique

Dans toutes les filières scientifiques, hors filières à dominante biologique, l'informatique fait l'objet d'un unique programme commun [3], que nous appellerons « tronc commun » par la suite, à l'exception de la classe ATS, où ce programme est allégé (la formation en ATS ne durant qu'un an).

Ce programme se déroule sur les deux semestres de première année et sur le premier semestre de deuxième année. Il comporte les parties suivantes :

- Présentation du système informatique utilisé et éléments d'architecture des ordinateurs : il s'agit essentiellement d'apprendre à utiliser les principales fonctions d'un système d'exploitation et d'un environnement de développement.
- Représentation des nombres et conséquences : il s'agit d'étudier la représentation des nombres entiers et les principes du calcul en virgule flottante.
- Programmation : il s'agit d'apprendre à écrire des programmes impératifs simples en Python (usage de boucles et de fonctions). La seule structure de données concrète proposée par Python qui soit employée dans le programme est le tableau dynamique. On présente la notion de pile, qu'on implante avec ces tableaux. La récursivité est introduite en seconde année. Le programme mentionne la notion d'invariant de boucle.
- Algorithmique : notion de complexité d'algorithmes (en mémoire et en temps, dans le meilleur et dans le pire des cas). En première année, l'algorithme le plus délicat étudié est la recherche par dichotomie dans un tableau. En seconde année, on étudie le tri par insertion, le tri rapide (par pivot) et le tri par fusion.
- Ingénierie numérique et simulation. Il s'agit d'apprendre à résoudre numériquement des problèmes comme la recherche d'un zéro d'une fonction (par dichotomie et par la méthode de Newton), la résolution approchée d'une EDO² par la méthode d'Euler, ou l'inversion d'une matrice par la méthode du pivot de Gauss.
- Initiation aux bases de données : il s'agit essentiellement de savoir consulter une base de données en utilisant SQL et/ou le langage de l'algèbre relationnelle. Les problématiques de conception d'une base, de contraintes de cohérence ou d'accès concurrents ne sont pas étudiées.

2. Équation différentielle ordinaire.

2.4.2. L'option informatique

Dans la filière MPSI/MP, et dans cette seule filière, est proposée, dans certains établissements seulement, une option informatique en plus du tronc commun informatique. Celle-ci se déroule sur le deuxième semestre de MPSI et les deux semestres de MP. Le programme de cette option comporte les parties suivantes :

- Méthodes de programmation : itération, récursivité, diviser pour régner, programmation dynamique ; le langage utilisé est Caml-Light.
- Structures de données : notions de structures de données abstraites (piles, files, dictionnaires, files de priorité), tableaux, listes, arbres (arbres binaires de recherche, tas binaires, réalisation d'une file de priorité avec un tas stocké dans un tableau).
- Calcul propositionnel (tables de vérité, tautologies, satisfiabilité).
- Graphes : représentation par matrice d'adjacence ou listes d'adjacence, parcours en profondeur et en largeur, recherche des composantes connexes d'un graphe non-orienté, algorithmes de Floyd-Warshall et Dijkstra.
- Langages rationnels et automates : langages rationnels, langages locaux, automates finis déterministes et non déterministes (sans ε -transitions), automates locaux, automate de Glushkov (procédure de Berry-Sethi), détermination, clôture par intersection et complémentaire.

3. Le tronc commun : un net progrès

La dernière réforme des programmes en CPGE s'est appliquée à la rentrée 2013. Nous présentons Section 3.1 le programme de tronc commun informatique qui existait antérieurement et Section 3.2 le programme de tronc commun introduit en 2013. Nous verrons ensuite qu'il a été pris en compte par les différents concours (Section 3.3), que ces heures ont maintenant le statut de véritables heures d'enseignement (Section 3.4) et enfin que cette réforme a permis la création de postes dédiés à l'informatique (Section 3.5). Sur tous ces plans, c'est un progrès notable.

3.1. Le programme de 1995 à 2013

Le programme du tronc commun informatique de 1995 à 2013 [1] disait ceci :

« L'enseignement d'informatique dispensé dans ces classes a pour principaux objectifs d'offrir :

- Une familiarisation avec l'utilisation d'outils informatiques évolués (logiciel de calcul formel et numérique, logiciels d'acquisition et de traitement de données, logiciels de modélisation, logiciels de simulation...) en vue de permettre l'approfondissement des disciplines scientifiques et techniques.

— Une introduction à l’informatique en tant que discipline, par une initiation élémentaire au traitement automatique de l’information, à l’algorithmique et à la programmation structurée (illustrée à l’aide du langage du logiciel de calcul formel retenu).

[...]

Aucune connaissance n’est exigible sur la complexité des programmes et sur les techniques de preuve de programmes. La mise en œuvre de la programmation n’est pas séparée de l’utilisation du logiciel de calcul formel en tant qu’outil et s’effectue à l’occasion des séances d’interrogations orales, appliquées à la résolution de problèmes de mathématiques, de physique, de chimie, de mécanique et automatique.

[...]

L’outil informatique n’est pas une fin en soi mais un moyen efficace pour faire des mathématiques, des sciences physiques ou des sciences industrielles. »

En pratique, cet enseignement reposait typiquement sur l’utilisation de Maple et, dans la plupart des filières, n’était évalué que dans deux concours :

- Le concours Centrale, où une épreuve d’oral de mathématiques s’effectuait sur machine. Il s’agissait alors essentiellement d’utiliser Maple comme une super-calculatrice pour résoudre un problème de mathématiques.
- Le concours X-ENS qui comportait une épreuve écrite de programmation et algorithmique élémentaire. Cette épreuve, corrigée uniquement pour les candidats admissibles, comptait pour l’admission mais non pour l’admissibilité.

Cet enseignement était administrativement considéré comme des heures d’interrogations orales, ce qui avait les conséquences suivantes :

- Il était moins payé que des heures d’enseignement d’une autre discipline.
- Il pouvait être effectué par des personnes n’ayant pas le statut d’enseignant ; il était donc typiquement assuré ou bien par l’enseignant de mathématiques de la classe ou bien par des étudiants en informatique ou en mathématiques.
- Il n’était pas envisageable de créer des postes d’enseignants à temps plein sur cet enseignement, les heures d’interrogations orales n’entrant pas en compte dans le service d’un enseignant.

La filière PT constituait cependant une exception notable car d’une part, l’essentiel des concours y est regroupé dans une unique banque d’épreuve, la Banque Filière PT, et d’autre part, celle-ci comportait une épreuve maths/info comportant un exercice de mathématiques au tableau ainsi qu’un exercice sous Maple.

3.2. En 2013 : un vrai programme d'informatique

Le programme de 2013 évoque des points essentiels pour une formation à l'informatique :

- structures de données,
- invariants de boucles pour montrer la correction des programmes,
- notions de complexité (en temps et en espace),
- algorithmes fondamentaux (recherche par dichotomie, tris).

Il s'agit d'un net progrès par rapport à l'antérieur.

3.3. Une prise en compte par tous les concours

Même si l'on verra qu'il convient de relativiser cette proposition, tous les concours évaluent ce nouvel enseignement. C'est un progrès par rapport à la situation antérieure où peu de concours l'évaluaient vraiment (essentiellement Polytechnique et Centrale-Supélec).

3.4. Statutairement : de vraies heures de cours

Les heures du tronc commun informatique ne sont plus des heures d'interrogations orales mais de véritables heures d'enseignement. C'est une forme importante de reconnaissance de l'informatique comme une véritable discipline par l'institution, puisqu'en particulier, ces heures ne peuvent être dispensées que par des enseignants déjà en poste.

Ce passage d'un statut d'interrogations orales à des heures d'enseignement a eu les conséquences suivantes :

- il n'est plus possible d'embaucher des étudiants,
- le vivier d'enseignants potentiels s'est alors essentiellement restreint aux enseignants de CPGE,
- en raison du nombre d'heures concernées, il a fallu créer des postes spécifiques « fléchés informatique ».

3.5. Des postes spécifiques

En raison du manque d'enseignants pour ce tronc commun informatique, les services de l'Éducation nationale ont été contraints de créer des postes spécifiques.

Les réticences de l'institution en 2013 étaient suffisamment fortes pour interdire toute création d'un corps spécifique. Ces postes sont donc officiellement des postes d'enseignants de mathématiques, de sciences physiques ou de sciences industrielles « fléchés informatique ».

Les enseignants ayant pris de tels postes ont dû assurer leurs inspections qu'ils continueraient à garder le lien avec leur discipline d'origine, par exemple en s'engageant à encadrer des TIPE (Travaux d'initiative personnelle encadrés) et/ou à donner des heures de TD ou TP dans leur discipline d'origine. Ce qui, dans certains cas, n'a

pas été sans difficultés vis-à-vis des enseignants qui se faisaient ainsi « déposséder » d'un TD ou TP.

4. Un enseignement encore trop marginal

Les enseignements en CPGE dépendent :

- des moyens horaires qui leur sont dévolus (volumes horaires, effectifs des groupes),
- des programmes officiels,
- des enseignants qui leur sont affectés,
- de l'organisation des concours.

Si les programmes officiels paraissent relativement satisfaisants *a priori*, nous verrons que sur les trois autres plans, l'informatique reste très marginale en CPGE.

Nous présentons Section 4.1 la question des horaires, Section 4.2 celle de la prise en compte de l'informatique aux concours, Section 4.3 la question de la compétence des enseignants et enfin Section 4.4 celle de la lisibilité de l'enseignement de l'informatique pour les étudiants.

4.1. Les moyens horaires

4.1.1. Un très faible volume horaire pour les étudiants

Comme nous l'avons vu Section 2.1, le volume horaire hebdomadaire d'informatique pour les étudiants des filières littéraires est nul, au sens propre du terme : le nombre d'heures d'enseignement d'informatique dispensé aux étudiants dans ces filières est 0, y compris dans la filière B/L qui comporte par ailleurs un programme de mathématiques évolué. Pour mémoire, le volume d'enseignement hebdomadaire reçu par un étudiant d'une filière littéraire est supérieur à 30 heures.

Dans les filières économiques et commerciales, seule une heure hebdomadaire est consacrée à l'informatique. Cet enseignement est intégré à l'enseignement de mathématiques et est rétribué comme une heure d'interrogation orale (« colle »). Le volume d'enseignement hebdomadaire reçu par un étudiant d'une de ces filières est au moins de 33 heures hebdomadaires.

Le volume horaire hebdomadaire d'informatique pour les étudiants des filières scientifiques a été fixé par un arrêté du 4 avril 2013 [6] modifiant l'arrêté du 10 février 1995 fixant les horaires des classes préparatoires scientifiques [5].

La répartition des horaires est résumée dans l'histogramme de la Figure 2. Les horaires donnés correspondent à des moyennes hebdomadaires en fonction des parcours possibles (indiqués sur la Figure 1), sur les deux années de CPGE (une seule pour la classe ATS). Ces horaires ont été calculés à partir des informations données par l'UPS [19].

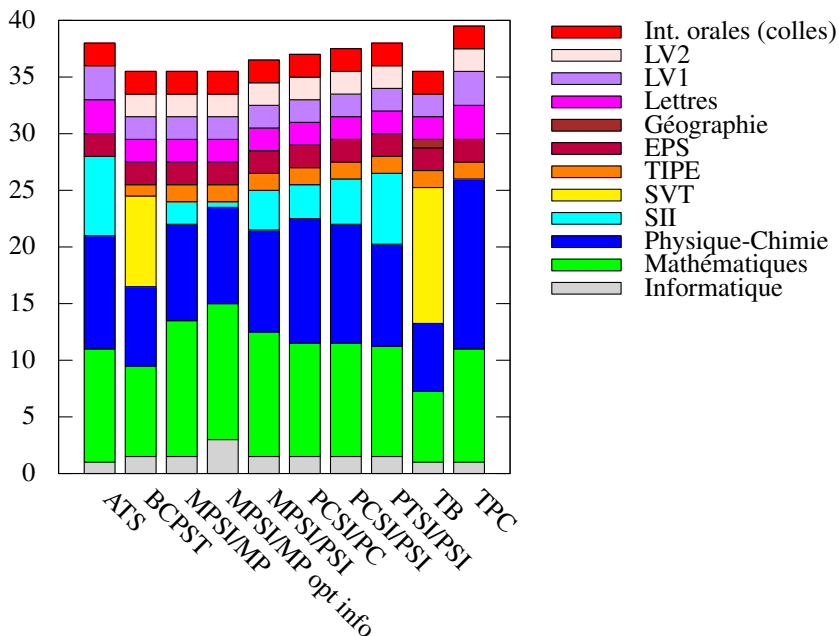


FIGURE 2. Horaires des prépas scientifiques

À une exception près, dans tous ces parcours le volume hebdomadaire moyen de l'informatique sur les deux années de CPGE est inférieur ou égal à une heure et demie. Ce qui représente au plus de l'ordre de 4,5 % du volume d'enseignement dispensé, soit 1/22^e. Dans tous ces parcours, ce volume est inférieur à celui de la 2^e langue vivante (facultative et dotée de deux heures hebdomadaires dans toutes les filières scientifiques).

L'exception est le parcours MPSI/MP avec option informatique. La filière MPSI/MP comporte en effet la possibilité, offerte seulement dans certains établissements, de suivre une option informatique (en plus du tronc commun informatique obligatoire), avec un horaire hebdomadaire de deux heures sur le deuxième semestre de la classe de MPSI et de deux heures sur l'année de MP.

Ainsi, dans la filière MPSI/MP, un étudiant choisissant l'option informatique aura un horaire total d'informatique (tronc commun plus option) de trois heures hebdomadaires, représentant un peu moins de 9 % des 33,5 heures hebdomadaires qu'il

reçoit s'il ne suit pas l'enseignement de LV2 optionnel. Cela reste très faible pour un parcours à coloration informatique.

4.1.2. *Effectifs : un cadre insuffisant*

Les effectifs en CPGE sont encadrés réglementairement :

- 48 étudiants maximum pour un cours,
- 24 étudiants maximum pour un TD.

Étonnamment, aucun texte n'encadre spécifiquement les effectifs de TP (en dehors des réglementations de sécurité pour les TP de physique-chimie). Ceux-ci ont donc réglementairement un effectif maximum de 48 élèves.

Les heures de tronc commun en première année consistent en une heure de cours et une heure de TP hebdomadaires (ce TP peut se réaliser sous forme de deux heures toutes les quinzaines).

En 2013, la direction générale pour l'enseignement supérieur et l'insertion professionnelle (DGESIP) a recommandé des effectifs de 12 étudiants maximum en première année pour les TP [8], recommandation rappelée dans le rapport de l'IGEN sur la mise en place de la réforme [10, paragraphe 2.4]. Malheureusement il ne s'agit que d'une recommandation et non d'un texte réglementaire. Cette recommandation est donc soumise au bon vouloir de la direction de l'établissement, qui dépend en partie de la dotation que celui-ci reçoit du rectorat. Il est donc possible, dans la mesure où les salles informatiques sont suffisamment grandes, de faire TP à 48 étudiants (effectif maximum pour une classe). Il ne semble pas qu'une situation aussi extrême se produise actuellement mais il est inquiétant de constater que les effectifs peuvent être une variable d'ajustement des lycées en cas de pénurie. Le rapport de l'IGEN sus-cité constate d'ailleurs que les effectifs vont de 12 à 24 étudiants par TP.

En seconde année, les horaires sont d'une heure de cours et d'une heure de TD hebdomadaires, pendant le premier semestre uniquement. Les effectifs de TD sont encadrés par les textes : 24 étudiants maximum.

4.1.3. *Une organisation semestrielle aberrante*

Les volumes horaires que nous avons donnés plus haut sont des moyennes annuelles. Comme nous l'avons dit, chacune des deux années d'enseignement est en fait organisée en deux semestres.

Or l'enseignement d'option informatique comme l'enseignement du tronc commun ne sont présents que sur trois semestres [5, 6] :

- L'option n'est proposée qu'à partir du début du second semestre. La justification est que le premier semestre doit permettre aux étudiants de choisir entre SI et informatique.

— Le tronc commun n'est prévu que sur les trois premiers semestres. On peut conjecturer que la raison est essentiellement budgétaire : le passage d'heures d'interrogations orales à des heures d'enseignement aurait entraîné un surcoût si cet enseignement avait eu lieu sur quatre semestres. La réduction à trois semestres a probablement été une façon de compenser ce surcoût.

L'enseignement du tronc commun s'arrête donc officiellement fin janvier de la deuxième année, soit environ trois mois avant les écrits (s'étalant de mi-avril à mi-mai) et cinq mois avant les oraux (mi-juin à mi-juillet). C'est tout simplement une aberration.

4.2. Une faible prise en compte aux concours

4.2.1. *Le cas du concours CCP : de très faibles coefficients*

Les concours « standard » du taupin (filières MP, PC, PSI) sont les Concours Communs Polytechniques (anciennement concours des ENSI). Dans toutes les filières, ce concours est noté sur 98 coefficients, dont 58 pour l'admissibilité aux oraux.

L'informatique n'est évaluée à l'oral dans aucune de ces filières.

À l'écrit, le tronc commun informatique n'est évalué par une épreuve spécifique que dans les filières PSI et TSI. Dans les autres filières, il est évalué lors d'épreuves de modélisation, à l'exception de la filière MP, où il est officiellement évalué dans l'épreuve d'option informatique pour ceux qui suivent cet enseignement optionnel et dans l'épreuve de SI pour les autres.

Il est donc difficile de savoir précisément quel est le poids de l'informatique à ces concours. Néanmoins, on peut estimer le coefficient de l'informatique en multipliant le coefficient d'une épreuve susceptible de contenir des questions d'informatique par la proportion des questions relatives à l'informatique dans cette épreuve, puis l'importance relative dans le concours en divisant par le nombre de coefficients du concours.

Par exemple, un candidat en filière MP n'ayant pas pris l'option informatique passe une épreuve de sciences industrielles qui a pour coefficient 7 et dont un quart est constitué de questions d'informatique. On peut donc considérer que l'informatique a pour coefficient 1,75, soit environ 1,8 % des 98 coefficients du concours.

À partir de la notice du concours de 2017 [11], on obtient ainsi, en considérant les sujets d'écrits pour les épreuves susceptibles de contenir des questions d'informatique [18, 14, 15, 12, 17, 13, 16], les valeurs suivantes :

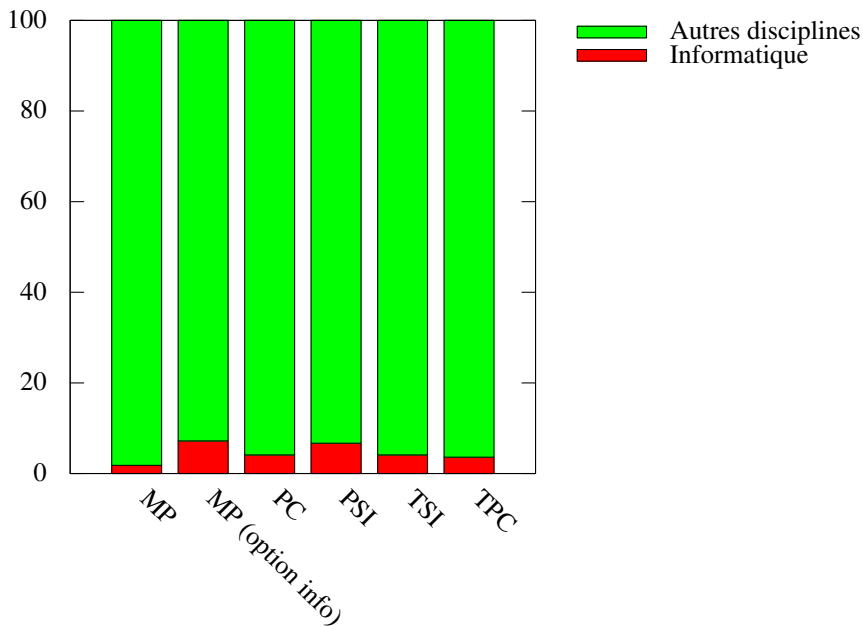


FIGURE 3. Informatique aux concours CCP (en pourcentage des coefficients)

Filière	type épreuve	coeff	info/épreuve	info/concours
MP	sciences industrielles	7	25 %	1,8 %
PC	mod. syst. phys. ou chim.	8	50 %	4,1 %
PSI	mod. et ingé. numérique	9	6 %	0,6 %
PSI	informatique	6	100 %	6,1 %
TSI	modélisation	8	0 %	0 %
TSI	informatique	4	100 %	4,1 %
TPC	modélisation	7	50 %	3,6 %

Les candidats en MP ayant suivi l'option informatique passent une épreuve d'informatique, de coefficient 7, à la place de l'épreuve de sciences industrielles.

L'importance de l'informatique dans ces différentes filières est résumée par la Figure 3.

Notons que si le coefficient de l'informatique en PSI semble relativement important comparé aux autres, cette épreuve comporte pour l'essentiel des questions liées à une simulation numérique. La seule question d'algorithmique consiste à reconnaître

un tri par insertion, à donner sa complexité dans le meilleur et le pire cas, et à choisir quelles instructions il convient d'effectuer dans la boucle la plus interne parmi quatre propositions (dont on peut voir que trois sont mauvaises sans réfléchir finement).

4.2.2. Une évaluation biaisée vers les applications

L'épreuve de tronc commun du concours X-ENS propose en général une épreuve d'informatique intéressante : on y présente souvent un problème qu'on résout d'abord avec un programme naïf avant de le faire résoudre avec un programme d'une meilleure complexité asymptotique.

Malheureusement, ce type d'épreuve est plutôt l'exception : dans les autres concours, l'épreuve porte en général sur la simulation numérique d'un problème, avec usage de la méthode d'Euler (voire d'une autre méthode de résolution approchée d'une équation différentielle). Les questions de complexité y sont la plupart du temps réduites à la portion congrue et les épreuves n'abordent que rarement le programme de deuxième année. Ainsi, sur les écrits du concours CCP 2017, il semble que seule une épreuve [13] portait (très marginalement) sur des notions de seconde année : il s'agissait d'une part de commenter un tri par pivot dont le code était donné au candidat³, puis de le modifier pour que la fonction de tri renvoie le nombre total d'appels récursifs effectués.

4.2.3. L'opposition option informatique - SI

Avoir une filière où cohabitent deux options bien distinctes — SI et option informatique — pose problème. Les étudiants de MPSI doivent choisir au bout d'un semestre s'ils prennent l'option informatique (deux heures hebdomadaires), l'option SI renforcée (quatre heures hebdomadaires), ou aucune option (auquel cas, ils continuent la SI deux heures par semaine).

L'option informatique est en général choisie par les étudiants ayant les meilleurs résultats. Or les concours ont tendance à mettre des moyennes similaires en option informatique et en SI, ce qui ne semble pas très juste au regard des autres résultats de ces étudiants. En conséquence, certains étudiants qui auraient pu réussir en option informatique choisissent SI alors qu'ils préféreraient l'informatique. C'est probablement dommageable pour eux sur le long terme et dommageable pour la nation, alors qu'elle manque d'informaticiens compétents.

On entend aussi des opérateurs de concours expliquer qu'ils ne veulent pas mettre, dans les épreuves du tronc commun informatique de MP, des questions qui seraient trop proches de l'option informatique pour ne pas favoriser les candidats ayant pris cette option. Pourtant, jamais aucun concours ne s'est ému de ce que, dans les épreuves de sciences physiques, les étudiants ayant choisi SI puissent être favorisés.

3. Ce code comportait manifestement une erreur dans la condition d'arrêt d'une des boucles de la fonction de partition, sans conséquence pour l'exécution du programme mais fort troublante pour le lecteur.

4.3. L'absence d'exigence de compétence des enseignants

Le tronc commun informatique est en général enseigné par des enseignants de l'établissement. Tout enseignant peut se déclarer compétent pour cet enseignement, sans aucun prérequis de diplôme.

Pour former des étudiants, il faut être soi-même compétent dans la discipline qu'on enseigne. Alors qu'on exige des enseignants titulaires en collège un master dans leur discipline, il est aberrant qu'aucun diplôme ne soit exigé pour l'enseignement de l'informatique.

Une procédure de certification des enseignants permettant de vérifier qui était capable d'enseigner l'informatique avait été évoquée par le ministère, mais rien n'a été mis en place à ce jour.

En revanche, les inspections générales ont décrété (sans que cela fasse l'objet d'un écrit à notre connaissance) qu'il conviendrait de respecter des quotas disciplinaires : ces enseignements devraient être partagés à parts égales entre enseignants de mathématiques, de physique et de SI. Sans prérequis de compétences.

Pire, l'UPS a reçu des témoignages d'enseignants de mathématiques, de physiques et de SI affirmant que leur inspection générale les a fortement incités à enseigner l'informatique, discipline dans laquelle eux-mêmes se considèrent incompetents.

Bien sûr, ces enseignants, pour la plupart, ont fourni un énorme travail d'auto-formation, bien aidés en cela par les écoles d'ingénieurs qui proposent régulièrement différents stages de formation. Le ministère en revanche n'a, à notre connaissance, pas organisé la moindre formation. N'étant pas l'organisateur de ces stages, il ne décharge pas de leur enseignement les enseignants qui y participent et ne les rembourse pas de leurs frais.

La formation des enseignants repose donc sur la bonne volonté et l'investissement — en temps et en argent — de ceux-ci. Est-ce bien raisonnable ?

4.4. Une faible lisibilité pour les étudiants

Tous les points précédents rendent ce dispositif d'enseignement peu lisible pour les étudiants : l'informatique en CPGE est tellement marginale qu'elle les conduit à penser que les écoles d'ingénieurs ne leur permettront pas de faire des études d'informatique solides. On observe ainsi des étudiants désireux d'étudier l'informatique quitter la prépa au bout d'une année pour entrer sur dossier dans des écoles spécialisées en informatique.

5. Conclusion : quelques pistes pour avancer

Comme on le voit, beaucoup reste à faire pour que l'informatique prenne une place digne de l'importance qu'occupe aujourd'hui cette discipline dans nos vies. Nous proposons ici trois pistes d'évolution :

- Corriger les aberrations les plus flagrantes de l'organisation de l'enseignement de l'informatique en CPGE (Section 5.1). Il s'agit de points certes mineurs au regard des enjeux globaux mais qu'il est facile de mettre en œuvre rapidement pour un coût faible voire nul.
- Créer une nouvelle filière scientifique (Section 5.2) : une filière maths-informatique (MI) ou maths-physique informatique (MPI). Le coût de la création d'une telle filière MI ou MPI se résume à celui de la création de nouveaux programmes. Le coût de la mise en place elle-même peut être faible si l'on étale dans le temps la création de ces nouvelles classes, voire nul si l'on ouvre ces classes en reconvertissant des classes d'une autre filière scientifique.
- Créer une agrégation d'informatique (Section 5.3) : cela semble indispensable pour disposer d'enseignants réellement compétents.

5.1. Corriger les aberrations flagrantes de l'organisation

Nous avons noté deux aberrations flagrantes :

- l'absence de cours de tronc commun au second semestre de la deuxième année ;
- l'absence de TP en option informatique.

Pour ce qui est de l'enseignement du tronc commun au second semestre de deuxième année, cela peut se faire en ajoutant des heures à cet enseignement au second semestre et/ou en étalant les heures sur l'année. À l'échelle d'un établissement, le surcoût est en général très faible. Si nécessaire, monter les effectifs des TP de première année de 12 à 16 étudiants pour compenser ce surcoût est une solution qui semble acceptable pédagogiquement.

Pour ce qui est des TP d'option informatique, on peut imaginer créer un TP de deux heures tous les quinze jours pour les étudiants, au moins en première année. Pour un effectif d'une trentaine d'étudiants en option informatique (effectif typique dans les établissements où cette option est proposée), cela représente un volume horaire de deux heures hebdomadaires, ce qui est en général négligeable sur la dotation horaire globale de l'établissement.

Du reste, ces solutions sont déjà mises en œuvre, au cas par cas, dans les établissements, ce qui ne manque pas de poser un autre problème : celui de l'équité pour les candidats aux concours puisque tous les établissements ne corrigent pas ces aberrations de la même façon. Tant que le cadre officiel sera aberrant, il sera difficile d'empêcher les établissements de tenter de le corriger.

5.2. Créer une filière maths-informatique

La création d'une véritable filière à dominantes mathématiques et informatique serait sans nul doute fort bénéfique :

- dans une telle filière, un enseignement d’informatique pourrait avoir lieu sur quatre semestres avec un volume conséquent (six ou huit heures par semaine par exemple) ;
- les questions d’équité aux concours vis-à-vis des SI disparaîtraient d’elles-mêmes ;
- enfin, cela donnerait une visibilité à l’informatique pour les étudiants des CPGE, mais aussi indirectement aux élèves des lycées et aux enseignants de CPGE et de lycée, dont beaucoup ignorent encore que l’informatique est une science.

Une fois une telle filière créée, les ouvertures de classes peuvent se faire de façon progressive, en lien avec le recrutement d’enseignants qualifiés.

Notons le succès de l’enseignement de spécialité ISN (informatique et sciences du numérique) en terminale, dont les effectifs ont doublé entre 2012 et 2016 pour atteindre 11,4 % des spécialités de terminale S [9].

5.3. Créer une agrégation d’informatique

Pour disposer d’enseignants formés en informatique, on dispose de deux moyens : recruter des enseignants déjà formés à cette discipline, ou former des enseignants d’autres disciplines.

Faire atteindre un niveau Master à des enseignants d’autres disciplines est probablement difficile (il faut trouver suffisamment de candidats) et coûteux (si on peut espérer former en trois ans à temps plein des enseignants ayant déjà une formation solide dans une autre discipline, pouvoir les former en moins de deux ans à temps plein paraît illusoire).

Recruter des enseignants déjà formés semble plus aisé. Se reposer comme actuellement sur l’option informatique de l’agrégation est tentant mais pose plusieurs problèmes :

- Les admis sur cette option ne sont pas légion ; il est difficile de croire qu’on pourra ainsi fournir un grand nombre de postes, le vivier potentiel étant essentiellement celui des étudiants sur un parcours mathématiques et informatique des ENS.
- Les postes « fléchés informatique » ne sont pas très attractifs actuellement : passer d’un poste de mathématiques dans une classe scientifique (hors BPCST et TB) à un poste d’informatique, c’est passer du statut d’enseignant le plus important de la classe (en volume horaire, donc aussi en reconnaissance de la part des étudiants) au statut d’enseignant le moins important, à cheval sur plusieurs classes.
- D’autres agrégations risquent de s’ouvrir à l’informatique : c’est déjà le cas de l’agrégation de SI ; on risque donc se retrouver avec une discipline morcelée, dont les enseignants auront sans doute trop peu en commun pour pouvoir dialoguer.

Il conviendrait donc de mettre en place une véritable agrégation d'informatique, qui pourrait fédérer les différentes visions de l'informatique plutôt que de conduire à un éclatement de la discipline.

À ceux qui craindraient que la mise en place d'une telle agrégation ne pose des difficultés insurmontables, on peut citer l'exemple du Maroc. Les programmes de CPGE y sont très proches des nôtres. Alors que l'organisation du concours de l'agrégation de mathématiques y repose en partie sur le concours français, le Maroc vient d'annoncer qu'il créait une agrégation d'informatique. Seul. Combien de temps la France restera-t-elle à la traîne ?

Références

- [1] Ministère de l'Enseignement supérieur et de la Recherche. Programme informatique filières scientifiques. BO n° 1 20 juillet 1995, 1995. Disponible sur <http://prepas.org/ups.php?document=771>.
- [2] Ministère de l'Enseignement supérieur et de la Recherche. Programmes des classes préparatoires aux grandes écoles. Filière : scientifique. Voie : biologie, chimie, physique et sciences de la terre (BCPST), 2013. Annexe 4 de [7], pages 100 et suivantes des annexes.
- [3] Ministère de l'Enseignement supérieur et de la Recherche. Programmes des classes préparatoires aux grandes écoles. Filière : scientifique. Voies : mathématiques, physique et sciences de l'ingénieur (MPSI), physique, chimie et sciences de l'ingénieur (PCSI), physique, technologie et sciences de l'ingénieur (PTSI), technologie, sciences de l'ingénieur (TSI), technologie, physique, chimie (TPC), mathématiques et physique (MP), physique et chimie (PC), physique et technologie (PT), physique et sciences de l'ingénieur (PSI). Discipline : Informatique première et seconde années. http://cache.media.education.gouv.fr/file/special_3_ESR/50/5/programme-informatique_252505.pdf, 2013. Annexe de [6].
- [4] Ministère de l'Éducation nationale et Ministère de l'Enseignement supérieur et de la recherche. Arrêté du 10 février 1995 définissant la nature des classes composant les classes préparatoires scientifiques aux grandes écoles, 1995. NOR : RESK9500108A.
- [5] Ministère de l'Éducation nationale et Ministère de l'Enseignement supérieur et de la recherche. Arrêté du 10 février 1995 fixant l'organisation générale des études et les horaires des classes préparatoires scientifiques aux grandes écoles, accessibles aux titulaires d'un baccalauréat ou d'un titre admis en équivalence ou d'une dispense, 1995. NOR : RESK9500109A.
- [6] Ministère de l'Éducation nationale et Ministère de l'Enseignement supérieur et de la recherche. Arrêté du 4 avril 2013 relatif au programme d'informatique des classes préparatoires scientifiques mathématiques, physique et sciences de l'ingénieur (MPSI), physique, chimie et sciences de l'ingénieur (PCSI), physique, technologie et sciences de l'ingénieur (PTSI), technologie et sciences industrielles (TSI), technologie, physique et chimie (TPC), mathématiques et physique (MP), physique et chimie (PC), physique et sciences de l'ingénieur (PSI), physique et technologie (PT), 2013. NOR : ESR51306084A. http://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=71586.
- [7] Ministère de l'Éducation nationale et Ministère de l'Enseignement supérieur et de la recherche. Programmes de la classe préparatoire scientifique biologie, chimie, physique et sciences de la terre (BCPST), 2013. NOR : ESR51306359A. Arrêté du 4-4-2013. J.O. du 30-4-2013. http://www.enseignementsup-recherche.gouv.fr/pid20536/bulletin-officiel.html?cid_bo=71645&cbo=1.

- [8] Direction générale pour l'enseignement supérieur et l'insertion professionnelle. Courrier du 19 février de la DGESIP à Mesdames et messieurs les proviseurs des lycées à classes préparatoires aux grandes écoles, 2013. Se trouve en annexe 8, dernière page de [10].
- [9] IGEN Laurent Chéno. Enseignement de l'informatique. Présentation aux journées pédagogiques de la SIF, mars 2017. Disponible sur <https://frama.link/1Dt8YmHc>.
- [10] Dominique Obert, Robert Cabane, Frédéric Carluer, Laurent Carroué, Laurent Chéno, Brigitte Flamand, Raymond Nicodème, Bertrand Pajot, and Norbert Perrot. Mise en place des nouveaux programmes de première année en CPGE. rapport à Monsieur le ministre de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, juillet 2014. Rapport n°2014-052. <http://www.education.gouv.fr/cid82710/mise-en-place-des-nouveaux-programmes-de-premiere-annee-en-cpge.html>.
- [11] Concours Communs Polytechniques. Notice relative aux modalités d'admission session 2017. http://ccp.scei-concours.fr/cpge/notice/pdf/notice_ccp_2017.pdf. Consultée le 29 mai 2017.
- [12] Concours Communs Polytechniques. Informatique, épreuve spécifique filière PSI. <http://ccp.scei-concours.fr/cpge/sujet/2017/PSI/PSI-Informatique.pdf>, mai 2017.
- [13] Concours Communs Polytechniques. Informatique, épreuve spécifique filière TSI. <http://ccp.scei-concours.fr/cpge/sujet/2017/TSI/TSI-Info.pdf>, mai 2017.
- [14] Concours Communs Polytechniques. Modélisation de systèmes physiques ou chimiques, épreuve spécifique filière PC. <http://ccp.scei-concours.fr/cpge/sujet/2017/PC/PC-Modelisation.pdf>, mai 2017. À l'exception de la page de garde, cet énoncé semble le même que [16].
- [15] Concours Communs Polytechniques. Modélisation et ingénierie numérique, épreuve spécifique filière PSI. <http://ccp.scei-concours.fr/cpge/sujet/2017/PSI/PSI-Modelisation.pdf>, mai 2017.
- [16] Concours Communs Polytechniques. Modélisation, épreuve spécifique filière TPC. http://ccp.scei-concours.fr/cpge/sujet/2017/TPC/TPC_modelisation.pdf, mai 2017. À l'exception de la page de garde, cet énoncé semble le même que [14].
- [17] Concours Communs Polytechniques. Modélisation, épreuve spécifique filière TSI. <http://ccp.scei-concours.fr/cpge/sujet/2017/TSI/TSI-Modelisation.pdf>, mai 2017.
- [18] Concours Communs Polytechniques. Sciences industrielles, épreuve spécifique filière MP. <http://ccp.scei-concours.fr/cpge/sujet/2017/MP/MP-SI.pdf>, mai 2017.
- [19] UPS. Horaires en CPGE. Site prepas.org, <http://prepas.org/ups.php?article=53>. Consulté le 29 mai 2017.

Annexe A. Signification des sigles des filières scientifiques

MPSI : Mathématiques, physique et sciences industrielles

PCSI : Physique, chimie et sciences industrielles

PTSI : Physique, technologie et sciences industrielles

TSI : Technologie et sciences industrielles

TPC : Technologie physique chimie

ATS : Adaptation technicien supérieur

MP/MP* : Mathématiques et sciences physiques

PC/PC* : Physique et chimie

PSI/PSI* : Physique et sciences industrielles

PT/PT* : Physique et technologie

BCPST : Biologie, chimie, physique et sciences de la terre

TB : Technologie biologie



Décès d'Alain Colmerauer, pionnier de l'intelligence artificielle

Frédéric Benhamou¹

Prologue

Pionnier de l'intelligence artificielle, Alain Colmerauer restera dans l'histoire de l'informatique comme le créateur du langage de programmation Prolog, qu'il a développé avec son équipe au début des années 70 et sur lequel il a travaillé, littéralement, jusqu'à la fin de sa vie.



À la fois mathématicien et informaticien, chercheur infatigable et visionnaire, il a perçu très tôt l'intérêt d'inverser la logique de la programmation classique, où chaque action de l'ordinateur doit être programmée très précisément, pour créer la programmation logique. Cette programmation déclarative décrit un univers auquel on s'intéresse et l'utilisateur soumet au programme des requêtes qui sont analysées par un algorithme qui fournit les réponses déduites du modèle programmé. Cet algorithme a été inspiré de ceux développés par Alan Robinson, puis Robert Kowalski.

Dans ces années héroïques qui voient naître les fondements de l'intelligence artificielle, Prolog n'en est pourtant pas le seul langage. En effet, dix ans avant la création de Prolog, John McCarthy fonde le laboratoire d'intelligence artificielle de Stanford

1. Professeur des universités, vice-président de l'université de Nantes.

et crée le langage LISP pour développer la robotique et l'intelligence artificielle. Les deux hommes se rencontreront et une amitié en naîtra, faite d'admiration et de respect mutuel.

Mais ni Alain, ni John n'auraient imaginé alors, qu'en janvier 2017, presque 50 ans plus tard, une « intelligence artificielle » nommée AlphaGo jouerait contre les meilleurs joueurs mondiaux du jeu de Go... et les battrait tous !

Prolog

Prolog a été initialement conçu pour traiter le langage naturel, celui des humains, à une époque où l'une des grandes ambitions de l'intelligence artificielle était déjà de programmer un ordinateur qui puisse traduire des textes entre plusieurs langues. Lors de son séjour à Montréal de 1967 à 1970, en 1968, son implication dans le projet de traduction automatique de l'université de Montréal l'amène de fait à développer un formalisme, les « systèmes-Q », qui lui permettra d'obtenir des résultats remarquables sur le plan linguistique, mais sera surtout le début de l'aventure Prolog.

De retour à Marseille, où il est nommé sur le premier poste de professeur d'informatique de l'université, Alain acquiert la conviction que l'avenir appartient à des langages de programmation de beaucoup plus haut niveau que Fortran, Cobol, ou même Algol 60. Les langages qu'il pressent seront logiques, capables de s'attaquer aux problèmes les plus difficiles. Le premier Prolog est créé en 1971 avec Philippe Roussel, Henri Meloni et Gérard Battani.

Prolog II

Cette première réussite académique, très remarquée, amène sur Prolog et son inventeur les projecteurs de nombreuses universités et le logiciel est diffusé dans le monde entier. L'attention portée à Prolog le conduit alors en 1979 à créer Prolog II, en renversant à nouveau complètement les acquis de la programmation logique pour définir une sémantique qui s'appuie sur une algèbre des arbres rationnels, qu'il invente, et des facilités de programmation multiples qui en font un véritable langage de programmation. Il a déjà en tête d'améliorer Prolog II pour qu'il soit un jour utilisé dans le monde de l'industrie informatique.

C'est à cette période, en 1982, qu'il est invité au tout nouveau Centre mondial informatique et ressource humaine, situé au 22 de l'avenue Matignon à Paris. Il y rencontre des chercheurs célèbres du MIT : Nicholas Negroponte, fondateur du Media Lab du MIT, Richard M. Stallman, créateur de la fondation pour le logiciel libre, Seymour Papert ou Alan Kay.

Cette époque est aussi celle où Prolog est soudainement projeté sur le devant de la scène médiatique, grâce à l'annonce par le MITI (Ministry of International Trade and Industry) japonais d'un projet d'ordinateur de 5^e génération conduit sur une décennie et s'appuyant sur les ordinateurs parallèles et sur un système d'exploitation basé sur

le langage Prolog. À cette occasion Alain fera partie d'une visite officielle au Japon, où il voyagera dans l'avion du président Mitterrand.

Alain confiera par la suite que ce séjour d'un an au centre mondial lui avait surtout permis de s'isoler et de travailler d'arrache-pied sur la suite de son projet : la programmation par contraintes.

En effet, une fois émancipé de l'approche logique, à ses yeux mortifère, il voyait Prolog comme une machine à résoudre des problèmes d'algèbre. Sur le modèle des arbres rationnels, il pouvait alors étendre son langage à d'autres algèbres : les booléens, les listes, les nombres rationnels. La programmation logique par contraintes était née !

Prolog III

De retour à Marseille à la rentrée 1983, il démarre un nouveau grand projet : Prolog III, une extension de Prolog II qui permettait d'étendre de façon spectaculaire les capacités déductives du langage en introduisant mais surtout en mixant de façon extrêmement élégante des algèbres diverses, en introduisant des aspects continus (l'algèbre des nombres rationnels) et en créant la notion de solveur de contraintes. Le premier prototype de Prolog III est tout de suite un succès. La start-up PrologIA est créée en janvier 1984 par Alain et ses collaborateurs initiaux (Bob Pasero, Michel Van Caneghem, Henri Garetta), et fin 1989 la première version commerciale de Prolog III est enfin disponible. La première conférence internationale sur la programmation par contraintes est organisée par Alain en 1995 à Cassis près de Marseille.

Prolog IV

La dernière étape de cette histoire de Prolog est celle de la conception et du design de Prolog IV qui a occupé son équipe pendant cinq bonnes années. Prolog IV s'inscrit dans la continuité de Prolog III, avec une avancée fondamentale qui est celle d'intégrer et à nouveau de mixer élégamment deux univers distants que sont les mathématiques discrètes, à la base de l'informatique théorique, et les mathématiques continues, celles qui fondent la physique, par exemple. Alain a passé de très longs mois à concevoir mathématiquement la construction sémantique du langage qui était d'une très grande complexité en raison de la richesse des possibilités qui y étaient incluses et il en a tiré un long article théorique, qui accompagne le manuel d'utilisation du logiciel.

Cette dernière production, en partie inachevée, met un terme à la lignée des Prolog marseillais, l'œuvre de sa vie. Elle est pour moi très représentative du scientifique et de l'homme qu'il était. Elle met en évidence son opiniâtreté, son engagement, son humilité, la volonté constante d'apporter autant de soin à la théorie qu'à la mise en

œuvre, de tout remettre en cause, toujours, de tout essayer, toujours, de tout expliquer, toujours, de tout maîtriser, toujours, tout en gardant le doute chevillé au corps. Je pense après toutes ces années que c'est ce qu'il nous faut garder de lui et je pense aussi que c'est ce qui fait de lui un immense scientifique et un homme attachant.

Alain Colmerauer était

- professeur d'informatique de classe exceptionnelle de l'université Aix-Marseille,
- correspondant de l'Académie des sciences, chevalier de la Légion d'honneur, et Fellow de l'American Association for Artificial Intelligence.

Ses derniers articles sont consultables sur son site alain.colmerauer.free.fr.

Il est décédé le 12 mai 2017 à Marseille.

Dates clés

- 24 janvier 1941 : naissance à Carcassonne
- 1968 : Directeur de TAUM (Traduction Automatique Université de Montréal)
- 1971 : Création de Prolog
- 1984 : Création de PROLOGIA
- 12 mai 2017 : décès à Marseille



Daniel Kayser, un enseignant-chercheur au cœur de l'IA

François Lévy¹ et Adeline Nazarenko²



En mai 2017, Daniel Kayser nous a quittés. Il a été emporté par la maladie. C'est une personnalité qui a laissé sa marque dans le paysage universitaire et informatique des cinquante dernières années et qui laisse un souvenir fort à beaucoup de ses collègues.

Nous retraçons ici à grands traits ce qu'ont été son parcours et sa réflexion, en essayant d'expliquer sa pensée et de montrer le rôle qu'il a joué dans la recherche française. Nous mettons l'accent sur ce qui nous semble être l'idée fondatrice de ses travaux de recherche et la manière dont elle les a irrigués.

Un parcours d'universitaire

Titulaire d'une thèse d'état à 29 ans et aussitôt nommé professeur, Daniel Kayser a été un informaticien influent dans la communauté française.

Par son enseignement tout d'abord, et il y tenait beaucoup. Daniel Kayser a dirigé 72 thèses et mémoires de thèse d'état, d'université ou de troisième cycle. Nombre de ses anciens étudiants ont entrepris une carrière académique, en France ou pour certains à l'étranger ; d'autres ont poursuivi dans l'industrie ou la R&D. Il n'est pas

1. Professeur émérite à l'université Paris 13, membre du Laboratoire d'informatique de Paris-Nord.

2. Professeur à l'université Paris 13, membre du Laboratoire d'informatique de Paris-Nord.

rare de rencontrer certains d'entre eux en position de responsabilité, au détour d'un état de l'art, en montant une nouvelle collaboration ou au hasard d'un congrès.

Daniel Kayser a aussi assumé des responsabilités au sein de la communauté universitaire et informatique. Il a été élu au Conseil supérieur des universités (1986-87). En 1985, il a créé le Laboratoire d'informatique de Paris-Nord, le LIPN, dont il a dirigé les débuts jusqu'en 1991. Il a été un des premiers dirigeants du PRC-GDR IA (après Henri Farreny) de 1992 à 1995, quand le CNRS et le ministère utilisaient conjointement cette structure pour soutenir le développement de l'intelligence artificielle (IA) en France. Puis il a participé au Groupe d'experts « Science et technologie de l'information » à la Mission scientifique et technique du ministère de l'Enseignement supérieur et de la Recherche. Il a été membre du CoNRS dans la commission 45, puis président du Comité scientifique disciplinaire n° 1 de l'ANR.

Daniel Kayser a surtout joué un rôle important dans l'animation de la recherche. Il a participé très activement à de nombreux groupes plus ou moins formels qui contribuaient aux échanges et à l'innovation scientifique. Il a contribué en 1981 à la fondation de l'ARC (Association pour la recherche cognitive), dont il a été successivement secrétaire, vice-président, président et qu'il a représentée à l'ECCAI³. Il était membre de l'AFIA (Association française pour l'intelligence artificielle) depuis sa fondation. Il participait au comité de rédaction d'*Intellectica* (revue de l'ARC), il a été rédacteur en chef adjoint de la *Revue d'intelligence artificielle*, il siégeait dans les instances du réseau Paris-Centre du PIR (Programme interdisciplinaire de recherche) Cognosciences, et il a été reconnu comme *ECCAI Fellow* en 2002.

Cette activité d'organisation avait comme pendant les nombreuses collaborations disciplinaires et interdisciplinaires que Daniel Kayser a tissées durant sa vie professionnelle.

La genèse d'une vie de recherches

Dans ses premiers travaux, Daniel Kayser s'intéressait à des dispositifs d'enseignement assisté. Il est venu à la langue par le besoin de traiter sémantiquement les réponses des apprenants et il a expérimenté pour cela des méthodes d'apprentissage.

C'est de cette recherche initiale qu'est venu son intérêt pour un *traitement automatique des langues* (TAL) qui prenne en compte la complexité du langage et ses imperfections. De là aussi le besoin de collaborer avec d'autres disciplines pour accéder à cette compréhension. On peut citer Jean-François le Ny, Jean-François Richard, Anh Nguyen-Xuan, Gérard Vergnaud, Stella Vosniadou, Olivier Houdé pour la psychologie, Pascal Engel ou Didier Andler pour la philosophie, Catherine Fuchs, Jean-Emmanuel Tyvaert en linguistique, mais on pourrait citer d'autres noms et d'autres champs disciplinaires. Daniel Kayser publie par exemple dès 1982 dans « Langage et Compréhension », édité par Jean-François Le Ny.

3. European Association for Artificial Intelligence.

De là aussi est né le souci de rendre informatiquement exploitable le résultat de cette compréhension, ce qui explique la place importante que la *représentation de connaissances* a occupé dans ses travaux. Dans les thèses qu'il dirigeait ou les mémoires de DEA qu'il encadrait, il s'agissait toujours de produire un programme générant et/ou manipulant un contenu sémantique. Daniel Kayser a perçu comme un vrai progrès l'évolution qui a fait passer la représentation des connaissances d'une vision procédurale à une vision logique, à partir des années 80. Il était convaincu qu'il fallait étendre les capacités de représentation au-delà de l'inférence matérielle pour rendre compte du raisonnement humain.

Daniel Kayser a travaillé en ce sens et publié ses travaux dans les congrès et revues d'IA, seul ou avec de nombreux chercheurs en IA. Citons quelques-uns de ces collaborateurs extérieurs à l'université Paris 13 où il a fait une grande partie de sa carrière : Daniel Coulon, Aicha Mokhtari, Adina Magda Florea, Catherine Garbay, Christine Froidevaux, Salem Benferhat, Didier Dubois, Henri Prade...

Une question centrale, celle du sens...

L'expérience avec les apprenants a très tôt conduit Daniel Kayser à prendre conscience que tout le monde ne comprenait pas la même chose à partir d'un même texte, et pas seulement pour des problèmes d'ambiguïté. En 1982, les deux Daniel, Coulon et Kayser, publient ensemble un article intitulé « Les sens uniques sont des impasses ». Cette réflexion sur le sens se poursuivra tout au long de la vie de recherche de Daniel Kayser, et c'en est, de notre point de vue, l'un des aspects les plus riches. Dans [SPS87], Daniel Kayser discute l'existence d'un sens comme objet d'une autre nature que le langage ; puisque la discussion du sens d'une phrase n'utilise en fait que des paraphrases, ce sont ces transformations qu'il faut analyser :

La notion de sens associé à un concept, à un mot, à une phrase, à un texte est peut-être une abstraction commode pour une description très rudimentaire du langage (comme la notion d'*impetus* semblait nécessaire au Moyen Âge, pour une description du mouvement) ; mais cette notion n'est pas véritablement fondée scientifiquement.

Daniel Kayser donne des exemples de transformation du sens au niveau du mot et propose une première ébauche de méthodologie dynamique de recherche du sens lexical. La dynamique s'entend ici comme une alternative à l'utilisation d'une liste de sens prédéfinis et permet de disposer de nouveaux sens en contexte. Dans [SLNL89], il revient à la charge avec une remarque étrangement prémonitoire au regard des tâches sémantiques sur lesquelles les informaticiens testent et évaluent actuellement leurs programmes :

Le sens d'une phrase ou d'un texte ne nous importe pas en tant que tel, mais par les conséquences que nous pouvons en tirer : compréhensibilité d'un

texte, inférences plausibles qui en découlent, jugements d'équivalence entre deux textes.

Il reprend dans [SLI97] sa critique d'une liste préétablie de sens. Au primat de l'inférence comme moyen de comprendre ce que dit un texte, il ajoute le primat de *principe* de l'inférence pour définir et analyser le sens des mots :

Pour s'assurer qu'une personne a compris un texte, on n'a généralement pas de meilleur moyen que de lui poser des questions sur ce texte ; la compréhension inclut donc la compétence, déclenchée par le questionnement, à produire une conclusion — la réponse — en fonction d'une prémisse — le texte. Cette compétence est clairement inférentielle, dans la première acception que nous lui donnons, celle qui se réfère au sens commun.

Avancer que la sémantique lexicale est d'abord inférentielle, c'est conjecturer que l'aptitude à tirer des conclusions ne dérive pas d'une raison plus profonde, par exemple de ce que les mots dénoteraient des objets d'un modèle qu'il suffirait de consulter pour obtenir les conclusions souhaitées ; c'est, tout au contraire, penser que l'aptitude à construire un ou plusieurs modèles résulte des propriétés inférentielles des mots du texte, propriétés vues comme primitives.

Daniel Kayser revient aussi sur la nature même de la communication langagière pour expliciter les raisons de ce renversement :

Ramener la notion de sens à celle de vérité [...] ne laisse pas d'étonner. En effet, on utilise le langage pour produire des effets sur ses interlocuteurs ; leur communiquer des vérités sur un univers de discours en constitue peut-être un des moyens, mais rarement utilisé en tant que tel, et qui présuppose plusieurs hypothèses peu réalistes. [On communique des interprétations de l'univers, pas des vérités.]

La réflexion sur la communication prend alors plus de place dans son argumentation. Il l'explicite par exemple dans [ANLS03] :

Au sein du couple perception-action, la conception traditionnelle du sens, fondée sur la référence, met l'accent sur la perception : les mots représentent les traits communs des entités perçues (même si « arbre » est un nom concret, il n'y a pas dans le monde des entités telles que les arbres, mais seulement des objets que nous avons appris à nommer ainsi, parce qu'ils partagent certaines caractéristiques perceptives ou fonctionnelles). Je veux explorer l'idée opposée : fondamentalement, les mots sont utilisés pour l'action, *i.e.* ils factorisent ce qui déclenche les règles d'inférence qui gouvernent notre comportement (nous dérivons un bon nombre de conclusions semblables quand on nous parle d'un arbre : le mot « arbre » renvoie fondamentalement à ces conclusions et non à l'ensemble des objets que nous nous accordons à appeler « arbre »).

On considère en général que la classe c des entités qui correspondent à un certain mot w est un sous-ensemble d'un univers U . Mais on reconnaît rarement que le but essentiel de la communication est souvent de créer (en partie) U . [...] les théories référentielles ignorent ou sous-estiment le fait qu'un des résultats de la compréhension est la modification de U . [...] aucun critère clair de ce qu'est ou n'est pas le sens littéral n'a été proposé. Les métonymies, les glissements de sens, les métaphores sont considérés comme des phénomènes secondaires que la sémantique peut ignorer dans un premier temps ; ils sont cependant si fréquents, même dans les textes techniques, qu'il est étrange de les laisser hors du cœur de la théorie.

Du sens à la représentation des connaissances

Daniel Kayser s'est aussi attaqué en pratique à cette énigme du sens, en essayant de mettre au point des dispositifs concrets pour construire des systèmes informatiques. Nous en mentionnons quelques-uns.

Il a exploré toute sa vie les raisonnements offerts par la *représentation des connaissances* pour comprendre/reproduire l'usage effectif de la langue. Son livre [RC97] sur le sujet a été une référence pour une large communauté.

Il est plusieurs fois revenu sur la notion de causalité, si intuitivement évidente et pourtant presque impossible à mettre en forme ; il est allé pour cela à la rencontre de la physique naïve, de la philosophie, du droit et des statistiques. Avec Aïcha Mokhtari, il s'est intéressé aux aspects temporels mais aussi épistémiques et normatifs de la causalité (1996-98). Avec quelques collègues du LIPN, Daniel Kayser a été à l'initiative d'un numéro d'*Intellectica* consacré à la comparaison des notions de causalité dans différentes disciplines et il y a lui-même contribué pour les modélisations symboliques (2004). Avec Farid Nouioua, il a proposé de modéliser la causalité comme un écart à un ensemble de normes décrivant le déroulement normal des événements (2004-09). Il a enfin contribué dans un groupe réuni par Henri Prade à la comparaison de différents modèles formels de la causalité (2008).

Daniel Kayser a aussi consacré beaucoup d'efforts à la description du lexique – dans les faits, surtout celui du français ou de l'anglais. Il a expérimenté, d'abord avec Hocine Abir, une approche non monotone du calcul lexical (1995-97). Il s'est efforcé de rendre compte des phénomènes de glissements de sens des noms, des adjectifs, des verbes, mais aussi de la variété d'interprétation de formes grammaticales comme le pluriel (travaux des années 1990 avec Françoise Gayral, François Lévy, Nathalie Pernelle, notamment). Il a ainsi proposé en 2003 [ANLS03] un répertoire de modes de variation qu'il nomme « polysémie systématique », en même temps qu'il esquissait un formalisme de représentation des connaissances lexicales en mesure de supporter ce calcul.

Références

- [SPS87] Daniel Kayser. *Une sémantique qui n'a pas de sens*. Langages 87, pp. 33–45, 1987.
- [SLNL89] Daniel Kayser. *La sémantique des langues naturelles et les logiques*. Annales des Télécommunications, 44, n° 1-2, pp. 34–52, 1989.
- [WKTC88] Daniel Kayser. *What Kind of Thing is a Concept ?* Computational Intelligence 4(4), pp. 158–165, 1988.
- [AILTS91] Daniel Kayser. *Adéquation et inadéquation de la logique au traitement sémantique des langues*. Modèles Linguistiques, tome XII, fasc. 1, pp. 119–136, 1991.
- [GLSL92] Daniel Kayser. *Sur les glissements de sens lexicaux*. Lexique et Inférence(s) (J.E. Tyvaert, ed.), Recherches Linguistiques XVIII, pp. 231–246, 1992.
- [RICL93] Daniel Kayser. *Représentation informatique de la connaissance lexicale*. Cahiers de Linguistique Française, pp. 69-84, 1993.
- [MFSU94] Daniel Kayser. *What kind of models do we need for the simulation of understanding ?* In Continuity in Linguistic Semantics (C. Fuchs & B. Victorri, eds.), pp. 111–126, John Benjamins, 1994.
- [SLI97] Daniel Kayser. *La sémantique lexicale est d'abord inférentielle*. Langue Française 113, pp. 92–106, 1997.
- [QECFR97] Françoise Gayral, Daniel Kayser, François Lévy. *Quelle est la couleur du feu rouge du Boulevard Henri IV ?* In Référence et anaphore, Revue VERBUM, tome XIX, nos 1-2, pp. 177–200, 1997.
- [RC97] Daniel Kayser. *La représentation des connaissances*. ISBN 2-86601-647-5, Hermès, 1997.
- [ANLS03] Daniel Kayser. *Abstraction and Natural Language Semantics*. Philosophical Transactions 358(1435), pp. 1261–1268, 2003.
- [FDAC09] Daniel Kayser, Farid Nouioua. *From the Description of an Accident to its Causes*. Artificial Intelligence 173(12-13), pp. 1154–1193, 2009.
- [PLR10] Daniel Kayser. *The place of Logic in reasoning*. Logica Universalis 4(2), pp. 225–239, 2010.



In memoriam Roger Mohr

Karl Tombre, Long Quan, Radu Horaud,
Patrick Gros, Cordelia Schmid, Peter Sturm

Roger Mohr nous a quittés subitement le 15 juin 2017. Il était l'un des promoteurs et des piliers de la vision par ordinateur en France, mais aussi un enseignant passionné.



Diplômé de l'ENS Cachan en 1969, assistant agrégé de mathématiques à l'IUT d'informatique de Nancy jusqu'en 1973, maître de conférences en informatique à l'université de Nancy jusqu'en 1984 puis professeur à l'INPL- École des Mines, il quitte Nancy pour Grenoble en 1988 et devient professeur à Grenoble INP – Ensimag, dans les laboratoires LIFIA, puis GRAVIR et LJK et au sein du centre Inria de Grenoble, où il a fondé et été responsable jusqu'en 1999 de l'équipe-projet MOVI. De 1999 à 2002, il a dirigé le laboratoire de recherche français de Xerox Research Europe à Meylan, avant de prendre la direction de l'Ensimag de 2003 à 2008.

Les débuts à Nancy : langages et formes

Ancien élève de l'ENS Cachan, titulaire d'une agrégation de mathématiques, Roger obtint son premier poste d'assistant à l'IUT de Nancy. C'était le début de la science informatique en France ; parmi les quelques personnalités qui l'ont fait émerger et s'établir, il y avait Claude Pair, qui après une agrégation de mathématiques

passée à l'ENS en 1956, avait commencé sa carrière comme professeur de mathématiques en classes préparatoires à Metz, puis à Nancy, avant de devenir attaché de recherche CNRS en 1963 puis maître de conférences en mathématiques appliquées à la faculté des sciences de Nancy (1965-1967), époque à laquelle il passe une thèse d'État sur la notion de pile. Avec Jean Legras, professeur de mécanique rationnelle et d'analyse numérique à la même faculté des sciences, Claude Pair établit la discipline informatique à Nancy, devient en 1969 professeur d'informatique à l'université de Nancy, et fonde en 1976 le laboratoire CRIN (devenu ultérieurement le LORIA) en insistant dès le début sur la nécessité de regrouper les forces de recherche en informatique dans le même laboratoire, qu'elles soient en école d'ingénieur, en IUT ou à la faculté. C'est grâce à ces efforts que derrière Grenoble, qui avait une longueur d'avance suite à l'action de Jean Kuntzmann, Nancy a pris place très tôt dans le paysage de l'informatique universitaire, aux côtés de Paris, Toulouse et Nantes-Rennes [Créhange & Haton 2014].

C'est donc dans cet environnement pionnier en informatique que Roger arrive comme jeune assistant. La personnalité de Claude Pair, spécialiste de la théorie des langages, marquait les orientations de nombre de premiers travaux. Si Roger a tout de suite été mis sur la piste du monde de l'image, cela restait dans cette orientation générale « langages ». Il racontait lui-même en 1993, à l'occasion des 20 ans du CRIN : « Claude Pair m'avait demandé d'étudier les langages de description de formes. L'étude des grammaires et des langages était alors active, suite aux premiers succès enregistrés par leur utilisation, notamment en compilation. Mon travail resta théorique (décidabilité de la reconnaissance, capacité descriptive des langages) et fut sanctionné par une thèse. Aucun programme, quelques algorithmes, des démonstrations, des exemples... L'impulsion nécessaire pour approfondir cette étude fut apportée par Jean-Paul Haton, fraîchement arrivé au laboratoire, qui avait pressenti l'intérêt des méthodes d'analyse syntaxique non canoniques pour la reconnaissance du langage parlé. » [Créhange & Haton 2014].

Même si ses intérêts scientifiques se sont de plus en plus focalisés sur la vision et l'image, Roger est resté passionné par tous les aspects liés à la théorie des langages et à l'algorithmique. À la session 1974 de l'École d'été de l'AFCET, il présentait avec plusieurs collègues nancéiens un cours sur quelques aspects de la théorie de la programmation, cours qui a donné lieu en 1978 à la publication d'un livre sous le nom collectif C. Livery [Livery 1978], prônant une approche de l'enseignement des bases de l'informatique que des générations d'étudiants nancéiens ont reçue comme fondement méthodologique [Pair et al. 1988]. Des années plus tard, dans le cadre d'une longue coopération avec Tom Henderson de l'université d'Utah, il a revisité les algorithmes de consistance d'arc et de chemin dans des graphes en proposant l'algorithme de relaxation discrète AC4 [Mohr & Henderson 1986], qui est devenu une référence.

Pour revenir au cœur de son focus scientifique à cette époque-là, la reconnaissance de formes, Roger a souhaité explorer aussi loin que possible le potentiel des langages de description de formes et de l'analyse syntaxique, dans la suite logique de sa thèse de 3^e cycle [Mohr 1975]. Il lance un ambitieux projet d'analyse syntaxique de dessins, baptisé MIRABELLE (Lorraine oblige) [Mohr 1990]. C'est l'objet de sa propre thèse d'État en 1979 [Mohr 1979] et des premières thèses qu'il encadre, dont celle de Gérard Masini [Masini & Mohr 1983]. L'analyse de formes 2D présentait un certain nombre de défis passionnants et Roger a « mis en piste » ou accompagné plusieurs chercheurs qui ont ultérieurement développé leur propre dynamique dans ce domaine, notamment Adnan Amin sur la reconnaissance de l'écriture manuscrite et Karl Tombre sur la reconnaissance de graphiques. Mais son intérêt évoluait vers le 3D et il a lancé le programme TRIDENT, avec le soutien de l'ADI. Il s'agissait de définir une approche de modélisation hiérarchique de formes 3D, permettant l'analyse par des méthodes structurelles, et s'appuyant entre autres sur des algorithmes de propagation de contraintes (d'où son regain d'intérêt à l'époque pour la relaxation discrète).

Mais manipuler des informations 3D nécessite de les acquérir, et Roger a de ce fait opéré une transition assez rapide vers les grandes questions de l'époque en vision par ordinateur : si détecter des plans et des quadriques dans des données 3D avec la transformée de Hough relevait encore dans une large mesure de l'évolution de travaux antérieurs en reconnaissance de formes [Belaïd & Mohr, 1984], les travaux sur la stéréovision par appariement de région qu'il a initiés avec la thèse de Brigitte Wrobel-Dautcourt [Wrobel-Dautcourt 1988] ouvraient le champ des nombreuses questions que pose la perception du 3D à partir de données issues de caméras. Le rêve d'avoir un modèle syntaxique capable d'analyser des scènes 3D un peu complexes (une sorte de « MIRABELLE-3D ») s'estompait peut-être, mais des champs d'investigation passionnants s'ouvraient, notamment le travail avec Long Quan sur l'extraction d'invariants perspectifs à partir de la vision monoculaire [Quan et Mohr 1989] pour apparier avec des modèles 3D, dont la suite allait se développer à Grenoble quand Roger a décidé d'y poursuivre sa carrière.

Roger est néanmoins resté curieux de beaucoup de champs connexes et n'hésitait pas à lancer des sujets de master ou de thèse sur différentes questions d'analyse d'images, qui se sont avérées des briques utiles dans le grand « puzzle » de la vision et qui ont par exemple donné une direction initiale aux travaux scientifiques de Marie-Odile Berger, à qui Roger avait proposé de regarder la notion de contours actifs [Berger & Mohr 1990].

À Grenoble : géométrie projective, reconstruction et indexation

À partir de 1988, Roger a donc poursuivi sa carrière à Grenoble, comme Professeur à Grenoble INP. Il y a rapidement fondé l'équipe MOVI, qui était d'abord

localisée Avenue Félix Viallet dans le centre de Grenoble, avec seulement deux bureaux pour l'équipe ! En 1996, l'équipe a déménagé dans le tout nouveau bâtiment du centre de recherche Inria Grenoble à Montbonnot Saint-Martin, où sont toujours actuellement localisées les trois équipes-projet Perception, Thoth et Morpheo qui sont issues, directement ou indirectement, de MOVI.

Roger a initié des discussions entre son équipe et des mathématiciens de l'Institut Fourier, permettant d'avancer la recherche sur les invariants projectifs et leur utilisation pour la localisation et la reconnaissance d'objets [Mohr & Morin 1991], [Gros & Quan 1992] et débouchant *in fine* sur les travaux en indexation visuelle de l'équipe (voir plus bas). Si ces premiers travaux sur les invariants projectifs concernaient encore surtout la reconnaissance d'objets par vision monoculaire, Roger en a vu rapidement, inspiré par [Tripp 1987], des applications en reconstruction 3D et ce, même quand la caméra n'était pas étalonnée. Ainsi, il était l'un des premiers à montrer que « *It can be done without calibration* » [Mohr & Arbogast 1990] ! Dans ce travail, des invariants projectifs 2D étaient utilisés pour résoudre différentes tâches de vision 3D, principalement en mode monoculaire, dont surtout l'estimation du point de vue de la caméra. La scène était supposée contenir des ensembles de points coplanaires, hypothèse qui fut ensuite rapidement levée lors du passage à plusieurs images, dans les travaux de Roger et son équipe ainsi que ceux de nombreux autres chercheurs de par le monde.

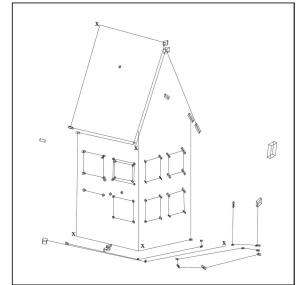
En effet, Roger a eu constamment le souci de nouer des collaborations avec de nombreux chercheurs français, européens et du monde entier et nous ne pouvons pas faire ici une liste exhaustive. Citons les collaborations avec Olivier Faugeras de l'Inria Sophia Antipolis, Hans-Helmut Nagel de l'université de Karlsruhe, Mike Brady et Andrew Zisserman de l'université d'Oxford, Jan-Olof Eklund du KTH, Gunnar Sparr de l'université de Lund, Richard Hartley de General Electric (État de New York, USA) puis de l'université de Melbourne, ou encore Luc Van Gool de l'université de Louvain. Roger a lié des amitiés avec la plupart de ces chercheurs.

C'était le début de l'âge d'or de la vision 3D : durant la majeure partie des années 1990, les recherches fondamentales et algorithmiques sur la reconstruction 3D non étalonnée et l'auto-étalonnage dominaient l'activité de la communauté de vision par ordinateur et l'équipe de Roger était parmi les principaux acteurs de ce domaine, avec les travaux de Long Quan, Radu Horaud, Bill Triggs, Peter Sturm, Boubakeur Boufama, Jérôme Blanc entre autres. Roger n'était nullement satisfait par le seul travail sur les questions théoriques



soulevées et a aussi donné beaucoup d'importance à des préoccupations plus terre-à-terre et notamment, sur la précision de tout le processus de reconstruction 3D, à travers les thèses de Pascal Brand et Gustavo Olague, et sur la validation expérimentale. Quant à ce dernier aspect, des images de la fameuse petite « maison MOVI » ont servi pendant plusieurs années à valider les algorithmes de l'équipe et d'autres chercheurs. Elles ont permis la première reconstruction non calibrée « réelle », passant d'une reconstruction projective à une reconstruction euclidienne ; ces travaux sont résumés dans [Boufama et al. 1993] et [Mohr et al. 1992]. La célébration de la première reconstruction projective de cette maison a eu lieu dans le jardin de la maison, parfaitement réelle et également euclidienne, de Roger.

Motivé entre autres par les questions de précision, Roger était l'un des premiers et rares « visionneux » à aller regarder du côté des photogrammètres et de leur savoir-faire et rigueur en gestion des incertitudes. Fidèle à son habitude, il a réussi à lancer des collaborations et échanges avec des scientifiques de cette communauté, dont Horst Beyer, Wolfgang Förstner (qu'il aimait appeler le « roi des covariances ») ou encore Clive Fraser à Melbourne, où Roger a passé une année sabbatique en 1998/99.



vue d'ensemble

Les travaux en géométrie 3D, et leur application en robotique qui était leur débouché principal au début des années 1990, nécessitaient d'utiliser plusieurs images et d'établir des liens entre ces images : c'est le problème de l'appariement des images, auquel Roger s'est intéressé en corollaire de ses travaux en géométrie. Au sein de l'équipe MOVI, Radu Horaud avait travaillé, en particulier avec son doctorant Thomas Skordas [Horaud et al. 1990], à l'extraction de primitives de type contours, segments et groupements de segments (jonctions) dans les images, et ce sont sur ces primitives qu'on a été initiés les premiers travaux en appariement avec la thèse de Patrick Gros [Gros 1993]. Celle-ci introduisait les concepts de base : invariants et quasi-invariants, contrainte géométrique qui seront repris dans la suite, mais en restait aux primitives de type segment. La thèse de Cordelia Schmid [Schmid 1996] apporta plusieurs changements importants : l'utilisation de détecteurs de points qui n'étaient plus basés sur la détection de contours, et de descripteurs basés directement sur les niveaux de gris des images, qui utilisaient des formules calculés par Luc Florack. Ces invariants différentiels calculés sur les points d'intérêt ouvraient des possibilités bien plus puissantes, tant pour l'appariement que pour l'indexation d'images, la reconnaissance d'objets ou d'images ou le suivi de points. Roger poussa l'étude, avec la thèse de Sylvaine Picard, d'une autre classe de descripteurs, dits non paramétriques, basés en particulier sur une structuration du voisinage des points d'intérêt qui préfigure la méthode proposée par David Lowe pour les descripteurs SIFT

qui connurent un succès mondial avant d'être détrônés par les descripteurs calculés par apprentissage profond.

C'est avec la thèse de Bart Lamiroy que le focus passa de l'appariement de deux images, c'est-à-dire de la comparaison d'une image avec une unique autre image, à l'indexation et à



la reconnaissance d'images, qui impliquait la comparaison d'image avec un ensemble d'autres images. Au départ, il s'agissait simplement d'essayer de paralléliser l'appariement d'une image avec plusieurs autres. Un des principaux résultats de la thèse fut de montrer que ce problème qui paraissait simple était en fait d'une grande complexité combinatoire, malgré la dimension encore réduite des descripteurs utilisés à ce moment là [Lamiroy et Gros, 1997]. Cette malédiction de la dimension qui venait d'être révélée allait être la source de très nombreux travaux par la suite pour essayer de la contourner, d'autant plus que la dimension des descripteurs employés ne faisaient que grandir avec les années : travail sur la représentation des descripteurs, leur stockage, leur indexation, leur comparaison. Mais ce travail se poursuivit sans Roger qui fut happé à cette période par ses charges de direction, d'abord à Xerox puis à l'Ensimag.

Finalement, ces travaux commencés comme un simple à côté technique des travaux de géométrie, se révélèrent bien plus féconds que prévu, et vaudront à Roger de recevoir avec Cordelia Schmid le prix Longuet-Higgins en 2006, prix récompensant l'article de la conférence CVPR, qui s'est tenue 10 ans auparavant, qui a eu le plus d'impact scientifique et qui a résisté à l'épreuve du temps.

Roger Mohr enseignant et mentor



Roger fut un enseignant passionné, qui garda toujours la capacité à s'enthousiasmer pour ce que proposaient ou faisaient ses élèves : « C'est super ! » était certainement une des ses expressions favorites. La direction de l'Ensimag fut le point d'orgue de sa carrière d'enseignant. Il organisa les spécialités de l'école en filières, développa la formation en alternance, créa la spécialité en images, vision et robotique et lança l'enseignement en anglais pour développer l'ouverture à l'international des masters de

l'école. Il travailla aussi beaucoup sur les relations entre l'école et les entreprises.

Tous ceux qui ont côtoyé Roger, qu'ils soient enseignants, chercheurs, doctorants ou étudiants, ont été saisis par sa modestie et son humilité. Il ne cherchait pas la

gloire pour lui-même et, de son propre aveu, il était l'homme le plus heureux du monde quand les chercheurs de son équipe obtenaient des résultats scientifiques de grande qualité et se faisaient remarquer lors de conférences internationales. En sont la preuve, les belles carrières universitaires et industrielles de la majorité de ses doctorants. Vers la fin de sa carrière il s'est confessé à l'un d'entre nous : « ma plus grande contribution scientifique a été le fait d'avoir pu lancer la carrière de plusieurs jeunes chercheurs, ce qui a compté plus pour moi que mes propres résultats. »

Pour nous tous, Roger fut bien plus qu'un collaborateur précieux, il fut un grand ami et un exemple à suivre.

Bibliographie (sélection)

[Belaïd & Mohr 1984] – Yolande Belaïd et Roger Mohr. « Planes and Quadrics Detection Using Hough Transform », Proceedings of 7th International Conference on Pattern Recognition, Montréal (Canada), pp. 1101–1103, 1984.

[Berger & Mohr 1990] – Marie-Odile Berger et Roger Mohr. « Towards Autonomy in Active Contour Models », Proceedings of 10th International Conference on Pattern Recognition, Atlantic City, NJ (USA), pp. 847–851, June 1990.

[Boufama et al. 1993] — Boubakeur Boufama, Roger Mohr et Françoise Veillon. « Euclidean constraints for uncalibrated reconstruction », IEEE International Conference on Computer Vision, pp 466–470, 1993.

[Créhange & Haton 2014] – Marion Créhange et Marie-Christine Haton. « L'informatique universitaire à Nancy : un demi-siècle de développement », 1024 – Bulletin de la société informatique de France, numéro 3, mai 2014.

[Gros 1993] Patrick Gros. « Outils géométriques pour la modélisation et la reconnaissance d'objets polyédriques ». Thèse de l'Institut National Polytechnique de Grenoble, 1993.

[Gros & Quan 1992] – Patrick Gros et Long Quan. « Projective Invariants for Vision », Rapport Technique 90 IMAG - 15 LIFIA, 1992.

[Horaud et al. 1990] – Radu Horaud, Françoise Veillon et Thomas Skordas. « Finding geometric and relational structures in an image ». Proceedings of The First European Conference on Computer Vision, 1990.

[Lamiroy et Gros, 1997] Bart Lamiroy et Patrick Gros. « Object Indexing is a Complex Matter ». 10th Scandinavian Conference on Image Analysis, 1997.

[Livercy 1978] – C. Livercy (nom collectif de JP Finance, M Grandbastien, P Lescanne, P Marchand, R Mohr, A Quéré et JL Rémy). « Théorie des programmes – schémas, preuves, sémantique ». Dunod Informatique, 1978.

[Masini & Mohr 1983] – Gérald Masini et Roger Mohr. « MIRABELLE : A System for Structural Analysis of Drawings », Pattern Recognition, Vol. 16, No. 4, pp. 363–372, 1983.

[Mohr 1975] – Roger Mohr. « Généralisation de la notion de langage à contexte libre – application à l'analyse syntaxique de figures », *Revue française d'Automatique, d'Informatique et de Recherche Opérationnelle*, Vol. 2, pp. 55–88, 1975.

[Mohr 1979] – Roger Mohr. « Descriptions structurées et analyse de formes complexes », Thèse de doctorat d'État de l'Université de Nancy 1, 1979.

[Mohr & Henderson 1986] – Roger Mohr et Thomas C. Henderson. « Arc and Path Consistency Revisited », *Artificial Intelligence*, Vol. 28, No. 2, pp. 225–233, March 1986.

[Mohr 1990] – Roger Mohr. « A General Purpose Line Drawing Analysis System », in H. Bunke and A. Sanfeliu, editors, « Syntactic and Structural Pattern Recognition : Theory and Applications », World Scientific, 1990.

[Mohr & Arbogast 1990] – Roger Mohr et Emmanuel Arbogast. « It can be done without camera calibration », Rapport Technique 805-I-IMAG 106 LIFIA, LIFIA—IMAG. Grenoble, France, 1990. Publié également dans *Pattern Recognition Letters*, Vol. 12, No. 1, pp. 39–43, 1991.

[Mohr et al. 1992] – Roger Mohr, Long Quan, Françoise Veillon et Boubakeur Boufama. « Relative 3D reconstruction using multiple uncalibrated images », *International Journal of Robotics Research*, Vol. 14, No. 6, 1995.

[Mohr & Morin 1991] – Roger Mohr et Luce Morin. « Relative positioning from geometric invariants », *IEEE Conference on Computer Vision and Pattern Recognition*, 1991.

[Pair et al. 1988] – Claude Pair, Roger Mohr et René Schott. « Construire les algorithmes », Dunod, 1988.

[Quan et Mohr 1989] – Long Quan et Roger Mohr. « Determining Perspective Structures Using Hierarchical Hough Transform », *Pattern Recognition Letters*, Vol. 9, No. 4, pp. 279–286, 1989.

[Schmid 1996] Cordelia Schmid. « Appariement d'images par invariants locaux de niveaux de gris. Application à l'indexation d'une base d'objets ». Thèse de l'Institut National Polytechnique de Grenoble, 1996.

[Tripp 1987] Colin Tripp. « Where Is the Camera ? The Use of a Theorem in Projective Geometry to Find from a Photograph the Location of the Camera », *The Mathematical Gazette*, Vol. 71, No. 455, pp. 8–14, 1987.

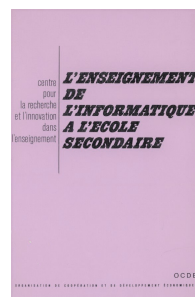
[Wrobel-Dautcourt 1988] – Brigitte Wrobel-Dautcourt. « Perception de la distance par mise en correspondance de régions dans des images stéréoscopiques », thèse de doctorat de l'INPL, 1988.



Le séminaire de Sèvres (mars 1970)

Jacques Baudé¹

Il est communément admis que l'introduction de l'informatique dans l'enseignement général français trouve son origine dans le Séminaire de Sèvres, même si quelques expériences ont eu lieu au cours de la décennie précédente². En effet, du 9 au 14 mars 1970 furent organisées, par l'OCDE-CERI³ avec la collaboration de la Direction de la coopération du MEN, au Centre international d'études pédagogiques de Sèvres, six journées consacrées à « l'enseignement de l'informatique à l'école secondaire »⁴.



On peut lire dans la préface des actes : « *Les travaux du CERI sur l'informatique à l'école secondaire ont pour objet d'aider les pays de l'OCDE à coopérer dans l'exécution des travaux de développement, et de disposer ainsi d'une large gamme*

1. Secrétaire général puis président de l'EPI de 1981 à 1995. Président d'honneur de l'EPI.

2. « *Pour une histoire de l'informatique dans l'enseignement français. Premiers jalons* » par Émilien Péliisset <https://halshs.archives-ouvertes.fr/file/index/docid/276158/filename/h85ep.htm>, et « *Introduction de l'informatique dans l'enseignement secondaire* », étude effectuée d'octobre 1971 à juin 1972 par MM. Pitié et Scherer : http://emmps.net/pdf/1971/etude_pitie-scherer.pdf.

3. L'Organisation de coopération et de développement économiques (OCDE) a été instituée par une convention signée le 14 décembre 1960, à Paris. Le Centre pour la recherche et l'innovation dans l'enseignement (CERI) a été créé en juin 1968 dans le cadre de l'OCDE (voir note 5).

4. « *L'enseignement de l'informatique à l'école secondaire* », publication de l'OCDE-CERI, 1971, 268 pages, archives de l'EPI.

d'informations, de données d'expérience et de résultats de développement, qu'ils n'auraient pu obtenir par leurs propres moyens. Ces travaux visent à inciter certains pays à aborder ce problème plus rapidement et de façon plus systématique qu'ils ne l'auraient fait en d'autres circonstances. »

Dès cette époque, les relations entre l'ordinateur et l'enseignement sont devenues un important sujet de préoccupation.

Comme le déclare James R. Gass, directeur du CERI : *« Tout d'abord, une chose est claire : il est indispensable que les enfants comprennent tous les aspects du monde technologique dans lequel ils sont appelés à vivre, sans quoi ils ne pourront pas échapper plus tard à des étonnements sans fondement ou à des craintes excessives en ce qui concerne les ordinateurs. Il faut prévoir à cette fin, pour la majorité des enfants, un « cours d'évaluation » sous une forme ou une autre, et pour bon nombre d'entre eux, la possibilité d'utiliser tôt ou tard un ordinateur. À cet égard, tout un monde de perspectives s'offre désormais à la coopération entre les écoles, les universités et les établissements scientifiques. »*

En second lieu, la conjoncture fait que l'informatique – c'est-à-dire le système tout entier des concepts de l'information, et notamment les programmes (software) qui constituent un élément plus important que l'ordinateur lui-même – tend à s'imposer comme une discipline à part entière dans les programmes d'enseignement secondaire. De l'avis de certains, les constructions algorithmiques qui sont le fondement de cette science contribuent à introduire une nouvelle forme de raisonnement logique et de solution des problèmes – un « langage » aussi fondamental que les mathématiques ou le langage inné lui-même. »

Dans son allocution⁵, Pierre Billecoq, secrétaire d'État auprès du ministre de l'Éducation nationale française, après avoir déclaré (page 14) : *« que son enseignement [de l'informatique] dans nos écoles ne la constitue pas en une « discipline », une matière du programme supplémentaire qui viendrait s'ajouter à celles qui sont déjà étudiées »*, le secrétaire d'État pose tout une série de questions parmi lesquelles : *« L'informatique vous paraît-elle devoir être – toutes expériences faites et toutes précautions prises – un langage essentiel du monde de demain, et comme telle, enseignée à tous, quelle que soit d'autre part leur orientation ? Si oui, à quelles conditions ? »*.

On trouve déjà là les atermoiements d'une politique que nous retrouverons tout au long du déploiement de l'informatique dans le système éducatif français.

Mais comme le déclare le secrétaire général adjoint de l'OCDE dans son allocution d'ouverture : *« L'informatique doit trouver sa place dans les programmes de l'enseignement secondaire »*.

Reste à voir comment, car tout est à inventer.

5. Allocutions d'ouverture : <http://www.epi.asso.fr/revue/histo/h70-sevres-allocutions-table.htm>

1. Les travaux du séminaire

1.1. L'ordre du jour (résumé des pages 30–31–32 des actes)

Lundi 9 mars

– 10 h : Ouverture des travaux.

– 15 h 00 : Session plénière : « *La signification de l'informatique dans son enseignement à l'école secondaire.* »

Président de la session : M. le professeur Jacques Arzac (France)

Mardi 10 mars

– 9 h 30 : Session plénière : « *Les objectifs et les contenus d'un enseignement de l'informatique dans l'enseignement secondaire.* »

Président : M. T. Crippin (Royaume-Uni)

– 15 h 00 : Session plénière : « *L'enseignement de l'informatique et l'enseignement des autres disciplines.* »

Président : Prof. A. van der Sluis (Pays-Bas)

Mercredi 11 mars

– 9 h 30 : Session plénière : « *Méthodes et moyens d'un enseignement de l'Informatique.* »

Président : Dr. Th. Wachsmann (Allemagne)

– 15 h 00 : Exposition et démonstration de matériels pédagogiques destinés à l'enseignement de l'informatique.

Jeudi 12 mars

– 9 h 30 : Session plénière : « *La formation des enseignants à l'enseignement de l'informatique.* »

Président : Prof. P. Ercoli (Italie) – 15 h 00 : Travail en groupes sur les thèmes suivants : Objectifs et contenus de l'enseignement de l'informatique. Méthodes et moyens. La formation des enseignants.

Vendredi 13 mars

– 9 h 30 : Poursuite du travail en groupes.

– 15 h 00 : Séance plénière : Compte rendu du travail en groupes.

Président : Prof. A. Berger (Autriche)

Samedi 14 mars

– 9 h 30 : Séance plénière : « *Suite à donner au Séminaire et coopération internationale.* »

Président : Prof. W. Atchison (USA)

– 12 h 30 : Fin du Séminaire.

Il est nécessaire en outre de signaler que le secrétariat du CERI avait demandé à plusieurs pays de préparer à l'intention de ce séminaire un document reflétant la situation générale de l'enseignement de l'informatique à l'école secondaire : c'est ainsi que l'Écosse et la Suède mirent à la disposition du secrétariat les documents correspondants. Par la suite, dans le déroulement même du séminaire, certains des participants diffusèrent des documents analogues (Canada, Danemark, Japon, etc.). Enfin, en vue de faciliter la discussion sur les moyens technologiques de cet enseignement d'introduction de l'informatique, le secrétariat du CERI avait demandé à un de ses experts (M. Jacques Perriault) de préparer un document décrivant sous une forme « thématique » les différents matériels d'enseignement de l'informatique, jusqu'ici développés par le secteur privé.

1.2. Les participants (extrait des pages 55 à 60 des actes)

Ils provenaient d'une vingtaine de pays : Allemagne, Autriche, Belgique, Canada, Danemark, États-Unis, Finlande, France, Grèce, Irlande, Italie, Japon, Luxembourg, Norvège, Pays-Bas, Royaume-Uni, Suisse, Suède, Turquie, Yougoslavie.

Pour la France :

- Jacques Arzac, directeur de l'Institut de programmation, Faculté des sciences, Paris
- Jean-Claude Boussard, maître de conférences à la Faculté des sciences de Grenoble
- Jacques Hebenstreit, directeur du Centre de calcul de l'École supérieure d'électricité, maître de conférences à l'Institut de programmation, Faculté des sciences, Paris
- Bernard Jaulin, directeur du Centre de mathématiques appliquées et de calcul, MSH Paris
- Max Lumbroso, chef de la section Enseignement programmé, département de la Recherche pédagogique, Institut pédagogique national, Paris
- André Poly, professeur agrégé à l'École normale supérieure de Saint-Cloud
- Pierre Poulain, chargé de mission à la Direction des enseignements supérieurs

Plus sept délégués nationaux, trois observateurs et un expert : Jacques Perriault, Centre de calcul, MSH Paris.

1.3. Les rapports présentés au cours des cinq sessions (pages 65 à 265 des actes)

Il est impossible de résumer 200 pages de contributions argumentées. Je me contenterai de quelques « flashes » portant essentiellement sur le débat ouvert (et jamais refermé !) : discipline informatique ou informatique dans les disciplines. Informatique « objet » ou informatique « outil » ? Nous avons en France deux hérauts opposés Jacques Arzac et Jacques Hebenstreit dont plusieurs articles sont en ligne sur le site de l'EPI.

Prof. Jacques Arzac (*Directeur de l'Institut de programmation, Faculté des sciences, 9 Quai Saint-Bernard, Paris*) s'attache à définir la finalité de l'enseignement de l'informatique. Or celle-ci ne peut résulter que d'une définition de la science informatique « *science autonome, avec son mode de pensée spécifique, ses lois propres, ses répercussions culturelles qui l'amèneront à être une des dominantes de la société de la fin du XX^e siècle* ».

Prof. W. F. Atchison (*Director of Computer Centre, University of Maryland, College Park 20742*). Pragmatique, l'auteur reconnaît que, pendant plusieurs années encore, les professeurs de mathématique auront à assurer la plupart des cours d'informatique dans le cadre de leur enseignement.

Prof. A. van der Sluis (*Elektronisch Rekencentrum, Universiteit van Utrecht, Boedapestlaan 6, Utrecht*). Pour l'auteur, tous les programmes de l'enseignement secondaire doivent impérativement accorder une place à l'informatique et pour lui la construction d'algorithmes est de toute première importance. Et plus loin : « *Il est très souhaitable que le professeur de mathématiques assure le cours d'introduction à l'informatique, il est très souhaitable également que tous les autres professeurs accordent le plus d'attention possible aux calculateurs dans leurs cours* ».

M. B. Jaulin (*Directeur du Centre de mathématiques appliquées et de calcul, Maison des sciences de l'homme, 54 boulevard Raspail, Paris*) montre « *certaines vertus de l'algorithmique* » dans différentes disciplines, sans se préoccuper, précise-t-il, d'une discussion sur le but d'un enseignement de l'informatique dans le second degré.

M. M. Bloxham (*Head of Mathematics Department Oundle School, Oundle, Peterborough, Royaume Uni*). Tout en reconnaissant les difficultés dues au peu d'enseignants formés, l'auteur affirme clairement que tous les élèves devraient suivre un cours d'introduction à l'informatique et qu'une partie des élèves du secondaire devrait étudier cette discipline. Il fait des propositions de contenus qui vont de la simple sensibilisation jusqu'à l'algorithmique-programmation, contenus précisés dans trois annexes.

Dr. G. C. Bonham (*Ontario Dept. Of Education Computer Science Curriculum, 44 Eglinton Avenue West, Toronto 12, Ontario*) préfère « *Étude* » plutôt que « *Science* » car le terme reflète plus exactement ce qui est enseigné en Ontario. Le programme présenté se différencie des deux autres programmes enseignés : « *Le traitement des données* » et « *La technologie des ordinateurs* ».

M. G. Farrington (*Department of Computer Science, University College of Swansea, Swansea, Royaume-Uni*). L'auteur plaide pour l'informatique « *outil* » pour les non-spécialistes et donne quelques exemples dans les disciplines en insistant sur la simulation.

Doc. J. Zweerus (*Elektronisch Rekencentrum, Universiteit van Utrecht, Boedapestlaan 6, Utrecht*) Les expériences menées dans les écoles néerlandaises ont montré qu'en une année scolaire à raison de deux cours par semaine, il était possible de traiter entièrement les problèmes essentiels que l'on rencontre dans l'utilisation pratique des ordinateurs, jusqu'à construire des algorithmes. Il s'agit maintenant de déterminer comment compléter judicieusement ce cours. La suite pouvant être prise en charge par les différentes disciplines. L'auteur propose des exemples en mathématiques, biologie, physique, géographie...

M. M. D. Mederith (*Senior Consultant, The National Computing Centre Ltd. Quay House, Manchester 3, Royaume-Uni*). Le Royaume-Uni n'est pas uni notamment pour ce qui concerne l'enseignement de l'informatique « *qui peut signifier n'importe quoi, allant de la description d'une règle à calculer... jusqu'à la formation d'opérateurs pour perforatrices à clavier* » (l'article est suivi de trois appendices).

Prof. J. Kuntzmann (*Faculté des sciences de Grenoble*) distingue trois plans : celui de l'éveil au rôle et à l'importance de l'informatique, celui de la préparation mathématique à l'informatique et celui de l'emploi de l'informatique dans l'enseignement.

M. J.-C. Boussard (*Faculté des sciences de Grenoble*) évoque des expériences menées à Grenoble (facultés et classes de troisième et de seconde des lycées). Il conclut : « *L'introduction de l'informatique dans le secondaire, qui semble donc devoir se faire prudemment en tant que discipline propre, devra sans aucun doute avoir lieu rapidement en tant qu'auxiliaire des autres disciplines, et même en tant qu'outil d'enseignement lui-même.* »

Doc. S. Sharp (*Director of Instructional Systems, The School District of Philadelphia*). Son rapport porte sur trois domaines : la formation des enseignants avant leur entrée en fonction, la formation des enseignants en cours de carrière et l'utilisation de l'ordinateur lui-même pour la formation des enseignants.

Prof. J. Hebenstreit (*Directeur du Centre de calcul de l'École supérieure d'électricité, maître de conférences à l'Institut de programmation, Faculté des sciences, 9, quai Saint-Bernard, Paris 5^e*). Pour l'auteur, la meilleure solution consiste non pas à former des professeurs d'informatique mais à recycler ou à former à l'informatique tous les professeurs du secondaire. Cette formation serait organisée de telle sorte que chaque professeur ait une connaissance de l'informatique et de ses applications dans le cadre de sa propre spécialité.

M. F. B. Lovis (*Head of Department of Mathematics, College of Education, City of Leicester, Leicester, Royaume-Uni*). L'auteur propose un programme d'urgence pour la formation des enseignants à l'informatique. Il y va de l'avenir économique (pour ne prendre que cet exemple) de l'Europe. Il n'hésite pas à parler d'apathie

actuelle. Et, de toute évidence, les enseignants ne sont pas suffisamment attirés par l'enseignement de l'informatique.

M. J. Perriault (*Centre de calcul, Maison des sciences de l'homme, 54 boulevard Raspail, Paris*) relate une première tentative de rassemblement des matériels didactiques pour l'enseignement de l'informatique. Étude qui devra être poursuivie compte tenu de l'abondance de la matière et aussi de son degré d'urgence.

1.4. Les rapports des groupes de travail⁶

Le groupe de travail n° 1 (Signification et objectifs) ouvre les deux voies que l'on appellera par la suite « informatique objet » et « informatique outil » : « *L'introduction d'un enseignement de l'informatique – comme discipline séparée ou intégrée à une autre discipline – nécessitera un examen attentif de l'ensemble du programme.* » Ce que les conclusions du séminaire reprendront (cf. ci-dessous).

De plus, le groupe de travail n° 2 (Méthodes et moyens) estime que, pour des raisons pédagogiques, les cours d'initiation à l'informatique devront inclure des travaux pratiques « *comprenant des passages sur ordinateur de programmes établis par les élèves* ». Des langages de programmation adaptés à l'enseignement sont donc nécessaires. Et le groupe de travail n° 3 (Formation des enseignants) prévoit « *la formation des enseignants qui auront pour tâche d'introduire cet enseignement de l'informatique* ».

Il est donc inexact de dire, comme l'ont prétendu certains, que le séminaire de Sèvres n'avait pas reconnu l'informatique comme discipline autonome.

Le deuxième volet, « l'informatique outil », est évidemment présent. Le groupe de travail n° 1 déclare : « *L'avènement de l'ordinateur influence de nombreuses disciplines. Les applications de l'ordinateur à toute discipline où elles sont justifiées devraient être développées de préférence dans le cadre même de cette discipline. Les contacts entre les enseignants chargés de cette dernière et ceux compétents dans le domaine de l'informatique devraient être facilités.* » (C'est ce qui se passera d'ailleurs en France, au cours des décennies 80 et 90, dans les très nombreux lycées où cohabitaient : pratique de l'EAO et option informatique.)

Le groupe n° 3 demande une formation générale pour les enseignants de toutes les disciplines. Ce groupe pousse un cri d'alarme : « *La situation actuelle de l'enseignement de l'informatique à l'école secondaire est critique... De tous les éléments qui contribuent à cette situation, la formation des enseignants est la plus importante.* » Nous pouvons malheureusement dire la même chose près d'un demi-siècle plus tard ! Les conclusions du séminaire reprennent l'essentiel.

6. Recommandations finales : <http://www.epi.asso.fr/revue/histo/h70-sevres-groupe-travail.htm>

1.5. Les conclusions du séminaire (pages 33 à 40)

Curieusement, dans les actes, les conclusions sont présentées avant les recommandations des groupes de travail. Elles méritent d'être lues dans leur intégralité⁷.

Je reproduis quelques passages importants qui donnent une bonne idée de l'ensemble.

Le séminaire insiste sur la « démarche informatique » :

« [...] sans méconnaître l'intérêt des autres cycles de formation, le Séminaire s'est tout spécialement intéressé à ce qui constitue le premier niveau de cet enseignement, à savoir cette introduction à la fin du premier cycle de l'enseignement secondaire. [...] L'accord a été général parmi les participants au Séminaire pour affirmer que ce qui était important dans cette introduction était, non pas l'ordinateur, mais bien la démarche informatique que l'on peut caractériser comme algorithmique, opérationnelle, organisationnelle. »

et ne ferme aucune voie :

« Envisagé comme enseignement ayant son propre statut, ou intégré dans une autre discipline d'enseignement, l'informatique est avant tout un langage, un système de signes qui permet de communiquer au même titre que d'autres langages, tels que les mathématiques ou les langues. Elle possède, atout majeur, mais aussi contrainte formatrice, la rigueur nécessaire à une approche scientifique. »

Les participants au séminaire ont bien vu l'importance cruciale de la formation des enseignants car l'enseignement de l'informatique, sous tous ses aspects, est une nécessité à laquelle les responsables de l'éducation n'échapperont pas :

« Cette formation des enseignants à l'informatique devrait s'effectuer à plusieurs niveaux :

(i) formation des enseignants qui auront pour tâche d'introduire cet enseignement de l'informatique,

(ii) formation plus spécialisée et plus différenciée pour les enseignants qui auront à développer cet enseignement dans l'enseignement technique, en particulier l'enseignement économique,

(iii) formation générale pour les enseignants de toutes les disciplines. qui peuvent être intéressés par l'informatique.

En l'état actuel du développement de cet enseignement de l'informatique à l'école secondaire, il semble bien que les objectifs (i) et (ii) soient prioritaires, étant bien entendu que très vite se posera le problème important évoqué en (iii), si l'on ne veut pas isoler l'enseignement de l'informatique des autres disciplines d'enseignement... »

Sans oublier les équipements, car : *« Dans la mesure où le contenu de cet enseignement d'introduction prévoit une initiation à la programmation conduisant à la*

7. Les conclusions du séminaire : <https://edutice.archives-ouvertes.fr/file/index/docid/276158/filename/h70ocde.htm> et <http://www.epi.asso.fr/revue/histo/h70ocde.htm>

rédaction de programmes par les élèves, il est nécessaire, au moins pour des raisons pédagogiques, que ces derniers voient le résultat de leur travail et prennent ainsi conscience de ce qu'un ordinateur est en mesure de faire, et aussi de ce qu'il exige comme rigueur et exactitude. »

« Il importe que sur le strict plan des équipements, les autorités responsables pratiquent une politique d'investissements d'autant plus soigneusement étudiée que le coût en est élevé. »

Et tant le problème est perçu comme fondamental, le séminaire revient, dans « *Les contraintes d'une politique d'enseignement de l'informatique* », sur la formation des enseignants :

« Il faut enfin souligner fortement que la pénurie générale en enseignants ne fait que rendre plus difficile la formation d'enseignants d'informatique. Si l'on ne veut pas se contenter d'une formation au rabais, avec comme conséquence une dégradation presque immédiate de cet enseignement de l'informatique au niveau des élèves, ce seront des ressources importantes qui devront être consacrées à cet aspect du problème sans doute prioritaire, tant en argent qu'en hommes. Or, il y a là un goulot d'étranglement dans la mesure où déjà les structures normales d'enseignement, y compris l'université, ne sont pas en mesure de former un nombre suffisant de techniciens de l'informatique pour répondre aux besoins du marché du travail. »

L'EPI, fondée l'année suivante, reprend les orientations du séminaire. On peut relire à ce propos l'éditorial du Bulletin n° 1 de décembre 1971⁸.

Tout cela était dit clairement il y a presque un demi-siècle !

2. La réponse française

Elle se manifeste dès mai 1970 par la création, au ministère de l'Éducation nationale, d'une Mission à l'informatique dont le responsable est Wladimir Mercouroff.

Je reprends ici l'analyse d'Émilien Pélisset, président d'honneur de l'EPI (voir note 2) :

La réponse française se trouve dans la circulaire ministérielle 70-232 du 21 mai 1970 (BOEN n° 22 du 28 mai) qui dessine d'ambitieux objectifs : « *L'informatique est un phénomène qui est en train de bouleverser profondément les pays industrialisés... [ses applications] en font un outil scientifique, technique et intellectuel unique. L'enseignement secondaire tout entier et dès la classe de 4^e (sic) ne peut rester à l'écart de cette révolution. Il doit préparer au monde de demain dans lequel ceux qui ignoreront tout de l'informatique seront infirmes. Il doit apprendre la portée de cet outil pour éviter les enthousiasmes excessifs et les scepticismes étroits. Il doit profiter de la valeur formatrice de l'enseignement de l'informatique, de la rigueur et*

8. Éditorial du premier Bulletin de l'EPI (1971), <http://www.epi.asso.fr/revue/01/b01p001.htm>

de la logique qu'elle impose. Il doit faire apparaître la portée économique du phénomène et faire savoir ce que l'informatique peut apporter à la vie professionnelle. Enfin, il doit préparer les consciences à affronter les responsabilités nouvelles créées par sa généralisation. »

2.1. La formation des enseignants

L'originalité de l'expérience française tient surtout à l'attitude des responsables, le chargé de mission à l'informatique et le comité pédagogique qu'il anime au Ministère de l'Éducation nationale : les problèmes posés par l'introduction de l'informatique sont pédagogiques, leur solution est affaire d'enseignants ; l'expérience débutera donc par une sérieuse formation informatique d'enseignants plongés dans la réalité industrielle des constructeurs d'ordinateurs. Il sera demandé à ces cobayes d'étudier comment utiliser cette formation et cette informatique dans leur enseignement. ... C'est ainsi que, dans la région parisienne, à la rentrée 1970-1971, sélectionnés sans qu'on sût comment parmi 1024 candidats (sic), 80 enseignants, venus de partout, se retrouvent, pour une année scolaire, répartis dans les trois centres de formation des principaux constructeurs d'ordinateurs (40 chez I.B.M., 20 à la C.I.I., 20 chez Honeywell-Bull).

Les leçons des premiers stages tirées, la formation approfondie dite « lourde » se déroulera, à partir de l'année 1971-1972, en milieu universitaire dans quatre centres : IUT de Nancy, de Toulouse, IMAG de Grenoble et ENS de St Cloud (ici à mi-temps). Comme l'avait souhaité l'EPI, l'encadrement fut « panaché » (enseignants du supérieur et enseignants du secondaire, du terrain), le recrutement des stagiaires resta pluri-catégoriel et pluridisciplinaire. L'ouverture du centre de Rennes en 1973 n'augmenta pas l'effectif annuel qui passa de 90 à 80. Ainsi de 1970 à 1976, 528 collègues furent formés (140 « mathématiciens », environ 200 « littéraires », plus de quarante « économistes »...). Dans le même temps, le CNTE de Vanves diffusa un cours par correspondance rédigé par un groupe d'enseignants et complété par deux fois deux jours d'applications sur ordinateur. Cette formation dite « légère » concerna plus de 5000 collègues. Enfin, nombreux furent ceux qui se formèrent « sur le tas » et/ou en dehors de l'expérience.

Pour la première, et la dernière fois, la formation des enseignants précéda les machines.

Ce que je résume ici brièvement a été développé dans mon article « L'expérience des 58 lycées » paru dans 1024 n° 4 d'octobre 2014⁹.

9. « *L'expérience des 58 lycées* » : <http://www.societe-informatique-de-france.fr/wp-content/uploads/2014/10/1024-4-baude.pdf>

2.2 L'orientation retenue : « l'outil » informatique dans les disciplines

Alors que la circulaire du 21 mai 1970 envisage un enseignement de l'informatique par les professeurs formés en 1970-1971 « soit à l'occasion de leurs cours traditionnels, soit dans des cours spéciaux pour volontaires, organisés à l'intérieur de l'horaire », le Comité pédagogique donne la place prépondérante à l'outil pédagogique pour les disciplines générales et à la pratique de la démarche informatique (modélisante, algorithmique et organisationnelle)¹⁰.

2.3. Les logiciels

Pour cette orientation il fallait des logiciels pédagogiques. La Section Informatique et Enseignement (SIE), créée dès 1971 à l'INRDP, a regroupé un certain nombre d'enseignants formés dans des groupes disciplinaires. Ces enseignants recevaient des décharges de service pour leur donner le temps nécessaire à leur recherche de terrain et à la conception-réalisation de logiciels¹¹.

Je cite Georges-Louis Baron, expert dans ce domaine : « Dès le lancement en 1970 de l'expérience des 58 lycées, des enseignants du second degré ont été associés aux recherches sur les applications de l'informatique à l'éducation, notamment par l'INRP qui a assuré le pilotage puis l'évaluation de la première expérimentation nationale, et également par des IREM. On trouve trace de leurs activités dans un nombre très important d'articles, de brochures, d'ouvrages, de revues, et notamment dans la revue de l'association EPI (*Enseignement Public et Informatique*) qui est une des références du domaine¹². »

2.4. Les matériels

De 1973 à 1976 furent mis en place dans 58 lycées, les mini-ordinateurs Mitra 15 (CII puis SEMS) et T1600 (Télémechanique) équipés du système L.S.E. Je n'y reviens pas¹³.

10. Comme le souligne Jacques Arsac : « La France n'a pas suivi à l'époque ces recommandations [celles du séminaire], préférant, sous la conduite de Wladimir Mercouroff et Jacques Hebenstreit, développer l'utilisation de l'informatique dans l'enseignement des autres disciplines (expérience dite « des 58 lycées »). » « *Des ordinateurs à l'informatique* », in Colloque sur l'Histoire de l'Informatique en France, Actes édités par Philippe Chatelin, 2 volumes (461+428 pp.), tome 1, pp. 31-43, Grenoble, mars 1988, ISBN 2-9502887-0-7, <http://jacques-andre.fr/chi/index.html#CHI1988>.

11. « *Dix ans d'informatique dans l'enseignement secondaire (1970-1980)* », Recherches pédagogiques n° 113, INRP, 4^e trimestre 1981, 182 pages. Archivé sur LARA, http://lara.inist.fr/bitstream/2332/1250/2/INRP_RP_81_113op.pdf

12. « *L'informatique en éducation – Le cas de la France* », Georges-Louis Baron, Revue française de pédagogie. Année 1990, Volume 92, n° 1, pp. 57-77, http://www.persee.fr/doc/rfp_0556-7807_1990_num_92_1_2474

13. « *Le système LSE* » : <http://www.societe-informatique-de-france.fr/wp-content/uploads/2015/12/1024-no7-Baude.pdf>

3. Conclusion

Ainsi, les bonnes questions furent posées au cours de ce séminaire, les diagnostics furent faits. Des décisions opportunes furent prises, ainsi celle concernant la formation des enseignants qui devait précéder les matériels. Malheureusement, les formations « lourdes » seront interrompues quelques années plus tard, elles repriront en 1981 (cf. le rapport Pair-Le Corre ¹⁴) pour s'arrêter définitivement. Nous en avons payé les conséquences au cours des décennies suivantes.

Pour ce qui est de l'informatique « objet » (objet d'enseignement), au cours des années 70 les clubs informatiques permettront un certain apprentissage de la programmation, mais il faudra attendre les années 80 pour voir apparaître l'option informatique des lycées à l'initiative de Jacques Arsac ¹⁵. Elle fut supprimée, puis rétablie, puis supprimée à nouveau. Nous connaissons tous la suite et les efforts de l'EPI, de la SIF, d'Inria, de Pascaline et de bien d'autres, pour le rétablissement d'un enseignement de l'informatique dans le secondaire, encore loin d'être acquis !

La « complémentarité » des approches, proposée par le séminaire de Sèvres dès 1970, n'en finit pas de s'imposer. Les idées et surtout les pratiques mettent plus de temps à s'imposer que les équipements ! Dans l'hypothèse optimiste – il n'est pas interdit de rêver – où l'enseignement de l'informatique et l'utilisation des outils numériques cohabiteraient harmonieusement dans l'enseignement général, il aura fallu un demi-siècle pour résoudre un problème loin d'être insoluble.

Bibliographie

– *L'informatique en éducation – Le cas de la France*, Georges-Louis Baron, Revue française de pédagogie. Année 1990. Volume 92. Numéro 1. pp. 57-77.

http://www.persee.fr/doc/rfp_0556-7807_1990_num_92_1_2474

– *L'expérience des « 58 lycées »*, Jacques Baudé (pour compléter le chapitre « La réponse de la France »), Bulletin 1024 n° 4, octobre 2014.

<http://www.societe-informatique-de-france.fr/wp-content/uploads/2014/10/1024-4-baude.pdf>

– *Dix ans d'informatique dans l'enseignement secondaire (1970-1980)*, collection Recherches pédagogiques, Recherches pédagogiques n° 113, Collectif, INRP, 4^e trimestre 1981, 182 pages. Archivé sur LARA (compléments pour les logiciels pédagogiques des années 70).

http://lara.inist.fr/bitstream/2332/1250/2/INRP_RP_81_113op.pdf

14. « *L'introduction de l'informatique dans l'éducation nationale* », rapport de MM. Claude Pair et Yves Le Corre, remis à Alain Savary le 15 octobre 1981, http://www.epi.asso.fr/revue/histo/h81_Pair-Le-Corre.htm

15. J. Baudé, « *L'option informatique des lycées dans les années 80 et 90. Première partie – La naissance d'une option* », http://www.epi.asso.fr/revue/histo/h10oi_jb1.htm

– *Quelques points de repère dans une histoire de 40 ans : L'association Enseignement Public et Informatique (EPI) de février 1971 à février 2011*, Jacques Baudé, EpiNet n° 132, février 2011.

http://www.epi.asso.fr/revue/histo/h11epi_jb.htm

– *L'option informatique des lycées*, Jacques Baudé, Bulletin 1024 n° 2, janvier 2014.

<http://www.societe-informatique-de-france.fr/wp-content/uploads/2014/02/1024-2-baude.pdf>

– *L'informatique dans les écoles, collèges et lycées français. Plus de 40 années de présence active de l'EPI. Les actes électroniques du colloque « Vers un musée de l'Informatique et de la Société numérique en France ? »*, Jacques Baudé, CNAM.

http://minf.cnam.fr/Papiers-Verifies/5.5_informatique_pedagogique_%20Baude%CC%81.pdf

– *Le système LSE*, Jacques Baudé (compléments pour les matériels et logiciels pédagogiques des années 70), Bulletin 1024 n° 7, novembre 2015.

<http://www.societe-informatique-de-france.fr/wp-content/uploads/2015/12/1024-no7-Baude.pdf>



Pour tout le monde : Tim Berners-Lee, lauréat du prix Turing 2016 pour avoir inventé... le Web

Fabien Gandon¹

Introduction

« *Je suis moi et mes circonstances.* »

– Jose Ortega

Quel pourrait être un point commun entre se renseigner à propos d'un concert, effectuer un virement depuis son compte bancaire, publier une base de données génomiques, échanger avec ses enfants à l'autre bout du monde et accéder aux données de sa voiture [1] ? Le fait de pouvoir le faire à travers le Web. Il est en effet difficile de trouver une activité humaine qui n'ait pas été impactée par le Web et, alors que j'écris cet article en avril 2017, on estime que le Web compte plus de trois milliards d'utilisateurs directs de par le monde. Ce même mois, le britannique Sir Timothy John Berners-Lee est lauréat du prix Turing 2016 pour avoir inventé ce World Wide Web, le premier navigateur Web et les protocoles et algorithmes permettant le passage à l'échelle du Web [2]. Sir Tim, comme il est coutume de l'appeler, est actuellement professeur au MIT et à l'université d'Oxford. Ce prix est le dernier en date d'une longue liste de distinctions qu'il a reçues. Mais le prix Turing est considéré comme le prix Nobel de l'informatique et nous étions nombreux à attendre la nomination de Tim pour son invention du Web, cette invention qui a transformé notre société depuis sa création en 1989. C'est donc pour nous l'occasion dans cet article de revenir sur

1. Directeur de Recherche, Inria Sophia Antipolis. Responsable de l'équipe Wimmics (Université Côte d'Azur, Inria, CNRS, I3S). Représentant Inria au W3C (World Wide Web Consortium). Fabien.Gandon@inria.fr – <http://fabien.info>

son histoire en essayant de montrer un grand nombre des influences et courants qui se sont mêlés pour permettre une telle invention. Ce sera aussi l'occasion, pour nous, de démêler certaines notions et d'en réintégrer d'autres dans un effort de mise en cohérence des nombreuses influences qui ont tissé la toile.

De Turing à Berners-Lee : une brève préhistoire du Web

« [...] juchés sur des épaules de géants,
de telle sorte que nous puissions voir plus de choses
et de plus éloignées que n'en voyaient ces derniers. »

– Bernard de Chartres, XII^e siècle

Les travaux de Turing sont évidemment omniprésents en informatique mais lorsqu'il s'agit du Web, il y a un cas singulier qui est celui du CAPTCHA : ces insupportables tests que le Web nous fait passer à outrance au prétexte de vérifier que nous sommes humains. L'acronyme (« *Completely Automated Public Turing test to tell Computers and Humans Apart* ») signifie littéralement qu'il s'agit d'un cas particulier du jeu de l'imitation de Turing, complètement automatique et ayant pour but de différencier les humains des machines [26]. Accessoirement, c'est aussi, notamment pour certains géants du numérique, un moyen de se procurer du temps de cerveau disponible gratuitement [27] et de « *Web-sourcer* » (externaliser sur le Web) l'étiquetage massif de bases d'entraînement d'algorithmes d'apprentissage (*machine learning*), par exemple. Au-delà de cette anecdote établissant un lien direct entre Turing et le Web, on peut identifier quelques grandes influences plus fondamentales dans l'invention du Web et ceci dès les contemporains de Turing.

La recherche de techniques d'organisation et de moyens d'accès efficaces aux masses d'informations que nous collectons a été une motivation omniprésente dans la préhistoire du Web. En juillet 1945, Vannevar Bush (MIT) écrit l'article « *As we may think* » (« Tel que nous pourrions penser ») [28] où il s'inquiète de ce que la capacité des scientifiques à assimiler les publications pertinentes pour leurs travaux soit dépassée par le volume de celles-ci. Vannevar pose alors comme un défi scientifique en soi le problème d'améliorer les moyens d'accès aux connaissances. Il propose comme premier élément de réponse un système (MEMEX [28]) utilisant des codes d'indexation mnémotechniques pour pointer et accéder rapidement à n'importe quelle partie de l'un des documents qui nous importent. Ces points d'accroche doivent aussi permettre de créer des liens entre deux éléments, externalisant ainsi le lien d'association. On trouve même littéralement la notion d'une toile d'idées (« web ») dans cet article historique lorsque Vannevar parle des liens d'association en ces termes : « *the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain* » [28].

Techniquement, pour réaliser un tel système, il faudra attendre vingt ans, l'apparition des ordinateurs et les débuts de leur utilisation pour l'édition de texte. Alors, dans un article à la conférence de l'ACM en 1965, Ted Nelson propose « une structure de fichier pour l'information complexe, changeante et indéterminée » [29] et il forge un néologisme pour nommer cet ensemble d'écrits interconnectés : l'hypertexte, et même l'hypermédia. Les années suivantes verront la réalisation des premiers éditeurs d'hypertexte utilisant, en particulier, une autre invention de cette décennie : la souris avec les interfaces et interactions qu'elle permet. On les doit notamment à Douglas Engelbart du Stanford Research Institute (SRI) qui recevra lui aussi le Prix Turing en 1997 pour sa vision pionnière et inspirante des interactions homme-machine et les technologies clefs qui ont permis son développement. En particulier, son système NLS (« *oN-Line System* ») combinera dans les années 1960, hypertexte, interface d'édition et souris. Le système sera plus tard renommé *Augment* en référence au programme de recherche d'Engelbart qui, en regard des travaux pionniers de cette époque en intelligence artificielle, propose de travailler sur l'intelligence augmentée [30].

Si le concept d'hypertexte est né, il reste, pendant des années, essentiellement limité à des applications s'exécutant localement à un ordinateur. La communication par commutation de paquets et inter-réseaux (*inter-networking* en anglais) se développe entre 1972 et 1975 avec les travaux de Louis Pouzin (IRIA puis Inria, [31]), de Vinton Cerf (SRI) et de Robert Kahn (DARPA, [32]) qui culminent en 1978 avec les protocoles standards (TCP/IP) et les débuts de l'Internet. Les applications de communication sur les réseaux se multiplient alors : le courrier électronique (SMTP), les listes de diffusion, le transfert de fichier (FTP), la connexion distante (Telnet), les forums de discussion, etc. Là encore nous trouvons deux récipiendaires du Prix Turing en 2004 : Vinton Cerf et Robert E. Kahn.

Cette trop courte généalogie intellectuelle place déjà Timothy Berners-Lee sur des épaules de géants lorsqu'il arrive en 1980 comme consultant pour le CERN (Centre européen de recherche nucléaire), alors jeune diplômé en physique (Queen's College, Oxford) et autodidacte de l'informatique. Devant la quantité d'informations et de documentations à gérer dans son travail et son équipe au CERN, Tim écrit un programme (*Enquire*) pour stocker des informations et les lier à volonté et notamment pour documenter les logiciels, ressources et personnes avec lesquels il travaille. Une différence importante d'*Enquire* avec d'autres systèmes hypertextes de l'époque est qu'il s'exécute sur un système d'exploitation multi-utilisateur et permet ainsi à plusieurs personnes d'accéder et de contribuer aux mêmes données [6]. Comparé au système de documentation CERNDoc basé sur une structure hiérarchique contraignante, Tim retient que les liens arbitraires et bidirectionnels de *Enquire* sont plus flexibles et évolutifs. De 1981 à 1983, Tim retourne en entreprise où il travaille sur l'appel de procédures à distance en temps réel (« *real-time remote procedure call* »),

donc dans le domaine des réseaux et de la programmation en réseau, avant de revenir au CERN en 1984.

Vague mais passionnant : les ruptures du filet et la naissance d'une toile

« *The World Wide Web was precisely what we were trying to PREVENT – ever-breaking links, links going outward only, quotes you can't follow to their origins, no version management, no rights management.* »

– Ted Nelson

En réalité, la première motivation de Tim pour créer le Web sera qu'il en avait lui-même besoin pour son travail au CERN, un campus où plusieurs milliers de personnes se croisent avec de multiples spécialités et instruments [13]. Tim est convaincu qu'il y a un besoin pour un système d'hypertexte global au centre de recherche du CERN. En mars 1989, pour convaincre la direction du CERN, Tim écrit une proposition de projet intitulée « *Information Management : A Proposal* » [6]. Tim s'y fixe comme objectif de construire un espace où mettre toute information ou référence que l'on juge importante et où fournir le moyen de les retrouver ensuite [6]. À défaut d'un meilleur nom, il y décrit un système qu'il appelle « Mesh » (Filet) où il suggère l'utilisation d'un hypertexte distribué pour la gestion d'information au CERN avec notamment la notion d'ancrage (*hotspot*, équivalent aux entrées de Ted Nelson) permettant de déclarer un morceau de texte ou une icône comme le départ d'un lien activable à la souris. Cette proposition est un formidable exercice d'équilibre entre intégration et rupture avec les paradigmes existants et émergents à l'époque.

Dans son livre [4], Robert Caillaud revient sur certains points importants du contexte général au CERN qui expliquent plusieurs choix faits dans la proposition de Tim. Commençons par quelques résultats existants à cette époque et que Tim intègre au cœur de l'architecture du Web.

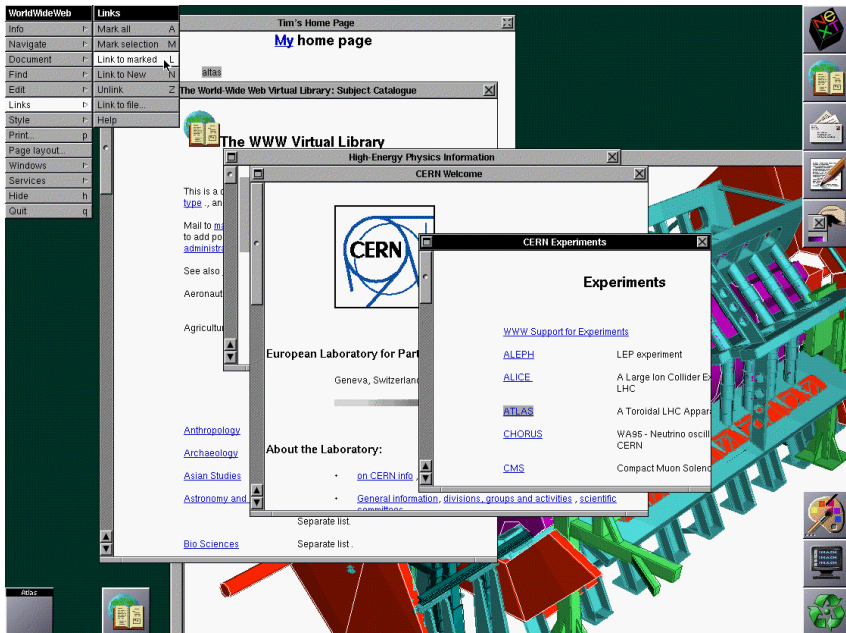
En 1989, le CERN est le plus grand nœud Internet d'Europe. Dans ce contexte, Tim a l'idée d'étendre les références des liens aux adresses réseau des documents afin de tisser ce « filet » (« *mesh* ») entre des documents mis à disposition par différentes machines. Conceptuellement, il reprend des principes des hypertextes et leur intègre TCP et le DNS. Plus précisément, Tim et l'équipe où il travaille s'investissent beaucoup pour l'acceptation et le déploiement au CERN de RPC (*Remote Procedure Call*, littéralement « appel de procédures à distance ») permettant à un programme s'exécutant sur une machine d'invoquer des procédures appartenant à des programmes s'exécutant sur d'autres machines. On peut voir RPC comme le chaînon manquant dans la distribution d'un hypertexte où le click sur un lien devient conceptuellement un appel de procédures à distance et le Web se conçoit alors moins comme un réseau

de documents mais comme un réseau de procédures. Des années plus tard, la thèse de Roy Fielding [33] introduira d'ailleurs le style d'architecture REST (*Representational State Transfer*) pour caractériser le système d'hypertexte distribué qu'est le Web. L'époque où Tim formule sa proposition est aussi celle où le système d'exploitation UNIX est très populaire et intègre nativement des fonctionnalités comme le support des protocoles Internet et les environnements multi-utilisateur.

Autre élément de contexte, en regard de la problématique de la documentation partagée notamment entre différentes disciplines au CERN, il règne déjà dans ce centre une culture de l'approche programmatique du document (p. ex. LaTeX, SGML) et ceci influence directement l'idée de s'orienter vers un langage de balisage simplifié. Une autre caractéristique assez peu connue du Web à sa conception, elle aussi héritée des systèmes d'édition d'hypertextes, est le fait qu'il était conceptuellement ouvert en écriture : tout dans son architecture permet non seulement de consulter (HTTP GET [34]) mais aussi de modifier le Web (HTTP PUT, POST, DELETE [34]). Le premier prototype de client imaginé par Tim est un navigateur-éditeur WYSIWYG [20] nommé W3. Cette fonction d'édition sera aussi longtemps portée et maintenue à l'Inria avec les travaux de Vincent Quint, Irène Vatton et leur équipe sur des bibliothèques et éditeurs Thot, Grif et Amaya [17]. Vincent avait par ailleurs travaillé avec Louis Pouzin sur le réseau Cyclades avant de s'intéresser aux documents structurés et en réseaux et, plus tard, Vincent deviendra le représentant d'Inria au consortium de standardisation de l'architecture du Web (W3C) avant de me passer le flambeau en 2012. Mais pour diverses raisons (sécurité, facilité, etc.) et par manque d'éditeurs [20], les navigateurs et serveurs les plus populaires ne vont pas mettre en œuvre cette possibilité de modifier les pages Web. Elle sera d'une certaine façon ré-introduite avec l'invention des Wikis (WikiWikiWeb) en 1994 par Ward Cunningham [44].

Enfin, la proposition de Tim relevant de la R&D informatique, elle ne peut être un projet du CERN en lui-même car elle est jugée hors du cœur de recherche de l'institut. La proposition est donc présentée comme un projet pour tester un nouvel ordinateur reçu au CERN [13], conçu et livré par la société NeXT dirigée alors par Steve Jobs, cofondateur d'Apple Computer, après sa démission forcée. Ce NeXT Cube était une station de travail haut de gamme, et elle va être utilisée par Tim comme premier serveur Web, ainsi que pour la conception du premier navigateur. Son système d'exploitation NeXTSTEP deviendra OPENSTEP, puis, à la suite du rachat de NeXT par Apple en 1996, Apple Rhapsody et enfin Mac OS et iOS. C'est aussi un environnement de programmation objet et graphique performant qui verra le développement de jeux mythiques comme Wolfenstein 3D, Doom et Quake. Pour ses premiers développements, trône donc sur le bureau de Tim un Cube NeXT qui matérialise la convergence des différentes influences que nous venons de mentionner et lui permet de s'appuyer sur l'expérience des hypertextes, l'avènement d'Internet et la mise

en réseau massive des ordinateurs (TCP/IP et Internet), les langages de programmation de haut niveau, les progrès des interactions homme-machine et notamment les interfaces graphiques et le multi-fenêtrage. Ce contexte de développement sera un accélérateur de la réalisation du premier serveur et du premier navigateur Web au-dessus d'une plateforme fournissant un grand nombre de briques de constructions de haut niveau. Ainsi fin 1990, le premier serveur et le premier navigateur sont testés à travers une connexion Internet. Le navigateur s'appelle World Wide Web ou « la toile d'envergure mondiale », qui deviendra le nom de l'hypertexte qu'il va engendrer et enverra aux oubliettes le « Filet » (« Mesh »).



Le premier site Web au monde est mis en ligne au CERN le 20 décembre 1990 (Photothèque CERN).

Mais, nous l'avons dit, Tim Berners-Lee choisit aussi de rompre avec un certain nombre de caractéristiques des solutions existant à l'époque de sa proposition. Une raison est qu'il soutient l'idée qu'il faut travailler à un système d'information universel, dans lequel la généralité et la portabilité sont plus importantes que d'autres extensions.

La proposition est conçue pour couvrir des ressources publiques comme privées et leurs liens. La solution doit permettre des associations aléatoires entre des objets arbitraires et une contribution incrémentale et triviale pour les objets comme pour les liens par différents contributeurs, justifiant à nouveau une approche décentralisée

et non hiérarchique [5]. Dès le départ, il s'agit aussi de sortir des silos des applications existantes de prise de notes, publication, édition, documentation, d'aide, etc. et de rechercher l'indépendance au domaine, à la tâche, au matériel et aux systèmes d'exploitation [13]. On peut encore voir cette spécificité du Web dans le choix d'un couplage faible et notamment l'indépendance entre le serveur et le client. En effet, quel que soit le navigateur utilisé et quel que soit le serveur interrogé la communication fonctionne du moment que les standards sont respectés. Le navigateur masquera même l'utilisation de différents protocoles (HTTP, FTP, etc.) [20]. Notons que l'on retrouve cette préoccupation de briser les silos jusque dans les combats actuels de Tim Berners-Lee sur le maintien de la décentralisation du Web. La diversité des plateformes utilisées au CERN l'incitera non seulement à rechercher l'indépendance par rapport aux systèmes d'exploitation mais aussi, par rebond, à s'abstraire des chemins et systèmes de fichiers [5].



Tim Berners-Lee faisant une démonstration du World Wide Web à la conférence Hypertext 1991, San Antonio, Texas (Photothèque CERN).

Enfin dans la proposition de Tim, l'hypertexte s'affranchit d'un serveur central : les données et les liens sont décentralisés sur Internet. De plus, les liens sont unidirectionnels et ne sont plus forcément maintenus : l'erreur 404 vient de naître. Cette rupture avec les fondamentaux des hypertextes est ce à quoi Ted Nelson fait référence dans la citation en début de cette section et explique aussi que la soumission de Tim à la troisième conférence ACM Hypertext en 1991 ne sera acceptée que comme une démo. Cependant ces ruptures avec l'existant sont autant de conditions

nécessaires au passage à l'échelle et à la viralité du Web. De plus, un certain nombre des ruptures du Web avec les hypertextes seront à l'origine de la création de services et d'applications Web majeurs. Par exemple, l'absence d'un index central et de liens bidirectionnels motivera la création d'annuaires, de répertoires (p. ex. Yahoo ! en 1994) et de moteurs de recherche (p. ex. Altavista, Google) pour (re)trouver des pages et des liens. Et, à bien y regarder, les pages comme celles générées par les résultats d'un moteur de recherche tissent maintenant à la demande une partie importante de la toile du Web.

In fine, on peut retenir trois notions fondamentales qui sont dès l'origine au cœur de l'architecture du Web :

— La première et la plus importante est celle des identifiants. Le prérequis majeur pour tisser une toile est d'avoir des points d'attache (*anchors*). La notion d'identification passera par les UDI, les URL, les URI et les IRI [13]. Les URL (*Uniform Resource Locator*) ou URI (*Universal Resource Identifier*) sont des formats d'adresses et d'identifiants permettant de localiser ou simplement nommer et faire référence sur le Web à n'importe quelle ressource. Si un identifiant (URI) donne un chemin d'accès pour obtenir une représentation de cette ressource sur le Web, alors on parle d'URL. Tout le monde connaît maintenant ces adresses, même si à l'origine elles n'étaient pas destinées à être manipulées directement par les usagers du Web. Par exemple, « <https://www.inria.fr/> » est l'URL de la page d'accueil d'Inria. Les URIs sont qualifiés d'universels au sens où ils doivent permettre d'identifier tout membre de l'ensemble universel des adresses réseau, dès l'instant où le protocole a une notion d'objet (la ressource). Il s'agit donc ici d'une vision orientée objet du réseau [20].

— La deuxième notion fondamentale de l'architecture du Web est celle de la communication et du transfert des données. Pour cela, le protocole HTTP permet notamment, à partir d'une adresse URL, de demander une représentation de la ressource identifiée et localisée par cet URL et d'obtenir en retour soit les données de cette représentation, soit des codes d'erreur indiquant un problème rencontré : par exemple, la célèbre erreur 404, qui indique que la ressource demandée n'a pas été trouvée.

— La troisième notion fondamentale est celle de la représentation des données échangées. Le Web étant initialement motivé par la représentation et l'échange de documents, le langage HTML sera le premier proposé pour représenter, stocker et communiquer les pages Web. Dès 1990, la documentation du Web est elle-même en HTML et le Web commence à s'auto-documenter ouvrant la possibilité à chacun d'augmenter cette documentation et d'apprendre à tisser en tissant. Cette forme de réflexivité confère au Web le statut d'un espace conçu pour qu'humains et machines y co-évoluent et y collaborent. Elle supporte la co-compréhension, co-documentation et la co-conception de tous les

sujets qui viennent s’y inscrire, à commencer par le Web lui-même.

Notons enfin que si URI, HTTP et HTML sont des inventions, la façon dont elles se combinent en une architecture simple et élégante l’est encore plus [13] et cela contribuera grandement à la viralité du Web.

Vague mais passionnant : début d’une viralité

« *Build small, but viral.* »

– Tim Berners-Lee

C’est avec la formule « vague mais passionnant » (« *vague but exciting* ») que fut acceptée la proposition d’hypertexte distribué de Tim Berners-Lee par son responsable Mike Sendall [35]. À l’heure où beaucoup de scientifiques s’interrogent sur l’efficacité de la gestion de la recherche par appels à projets, ce retour succinct mais assumant la prise de risque peut laisser songeur.

Nous l’avons dit, fin 1990, le premier serveur et le premier navigateur sont testés à travers une connexion Internet. Le 6 août 1991, Tim poste un résumé du projet World Wide Web sur plusieurs groupes de discussions en ligne (*newsgroups*), dont le forum dédié au sujet des hypertextes `alt.hypertext`. C’est le moment où le projet Web devient une application publiquement disponible sur Internet [18]. Le premier site Web est mis en ligne le même jour à l’adresse `http://info.cern.ch`. C’est aussi la source de documentation du Web lui-même, la graine de départ, le noyau auto-générateur en quelque sorte dont on trouve encore une archive en ligne [16]. Toujours en 1991, le premier serveur Web est installé hors d’Europe au Stanford Linear Accelerator Center. De ce point de départ, le Web va multiplier par dix le nombre de ses serveurs chaque année les deux premières années. Début 1992, on recense une dizaine de serveurs Web et de nouveaux navigateurs apparaissent au cours de l’année (Erwise, ViolaWWW, MidasWWW, Samba pour Macintosh, etc.). En 1993, les dirigeants du CERN annoncent officiellement que la technologie du Web sera gratuite et libre de droits [18], et en début d’année, on dénombre une cinquantaine de serveurs. De nouveaux navigateurs apparaissent (Lynx, Cello, Arena) mais le plus important est Mosaic, alors disponible sous Unix, Windows et Mac OS. Il permet notamment de visualiser les images directement dans le texte d’une page. Avec le navigateur Mosaic, le Web va réellement se répandre mondialement, laissant derrière lui ses ancêtres Gopher, WAIS et FTP. À Mosaic succéderont, dans l’arbre généalogique des navigateurs, Netscape puis Mozilla et enfin FireFox. En 1994, plus de 600 serveurs sont en ligne. L’année suivante, plus de 10 000 serveurs Web sont disponibles et Microsoft lance Internet Explorer, qui s’imposera comme le navigateur sous Windows, avec lequel il est diffusé. L’année 1995 voit aussi la naissance de JavaScript poussé par Netscape. En 1996 on passe la barre des 100 000 serveurs

et, en 1998 celle du million. Au début des années 2000 on en dénombre 26 millions et en 2004 les 46 millions ont été largement dépassés.

Un autre indicateur de cette viralité est donné par Tim dans un article de 1994 où il explique que la charge des accès à la documentation du Web sur le serveur `info.cern.ch` double tous les quatre mois entre avril 1991 et avril 1994 [20]. Cette viralité s'explique par plusieurs choix faits lors de la conception et la réalisation du Web. Outre les choix architecturaux déjà décrits, plusieurs éléments vont être décisifs pour permettre et maintenir la viralité du Web. Il y a des éléments techniques à commencer par la proposition de Tim de mettre l'accent sur l'importance de la généralité, de la portabilité et de l'extensibilité, plus importantes selon lui que la satisfaction d'utiliser les dernières capacités des ordinateurs (p. ex., le graphisme). Tim aurait pu intégrer des technologies plus complexes ou viser de multiples fonctionnalités supplémentaires mais si le Web est aujourd'hui une contribution d'importance technique durable et majeure à la communauté informatique c'est notamment en raison de sa simplicité, de son élégance et de son extensibilité [2].

Dans sa proposition, Tim prévient aussi que le résultat doit être suffisamment attrayant à l'usage pour que les informations contenues dépassent un seuil critique et qu'à son tour cette masse d'information encourage toujours plus l'utilisation et la contribution. Pour cela, il propose dès le début de prévoir une liaison entre les bases de données existantes et les nouvelles [6], et de rendre le Web systématiquement compatible avec l'existant. Dès sa création, le Web intègre des serveurs « *gateway* » pour importer des ressources légataires et donner accès à d'autres applications à travers des techniques comme CGI (*Common Gateway Interface*). Le Web va ainsi intégrer, interopérer et finalement absorber des systèmes existants, notamment WAIS et Gopher. Cette approche facilite le basculement complet de communautés d'utilisateurs d'anciens systèmes vers le Web. Tim conçoit aussi la rétrocompatibilité ou compatibilité descendante avec les protocoles précédents (FTP, Gopher, WAIS, etc.) comme une contrainte d'interopérabilité, une preuve de flexibilité et surtout une assurance d'évolutivité pour le futur [5, 20]. De plus, avec des approches comme les CGIs, la génération automatique de pages dynamiques et la possibilité de les référencer et de les lier joue immédiatement un rôle vital dans le tissage d'un graphe du Web qui soit un minimum fourni (nombreux nœuds), connecté et dense (nombreux liens) [20].

Tim a donc reconnu que la simplicité était nécessaire pour une adoption généralisée, en particulier dans la communauté scientifique qu'il visait. Ses simplifications de protocole, y compris son insistance sur l'absence d'états dans le protocole HTTP, ont rendu la conception facile à mettre en œuvre. De même, l'utilisation de scripts lisibles par l'humain rendait le système compréhensible, facile à déboguer, et virtuellement réutilisable par copier-coller-adapté [2]. Tim s'inspire systématiquement de l'existant aussi pour encourager l'adoption. Ainsi, HTTP s'inspire de SMTP et NNTP et les en-têtes utilisées dans les échanges HTTP (*headers*) sont une extension

des MIME (*Multipurpose Internet Mail Extensions*) afin d'ouvrir la porte à l'intégration des hypermédia, des mails, des news, etc. [20]. Du côté des URI, on peut noter l'intégration du DNS et des slash (/) pour leur structure [13]. De même, la volontaire proximité de HTML à SGML va encourager l'adoption de HTML par la communauté de la documentation [5]. Enfin, en termes d'interaction, le couplage de l'hypermédia et des formulaires textuels se révèle suffisamment simple et puissant pour couvrir de nombreux besoins tout en restant facile à mettre en œuvre et utiliser [20].

Mais au-delà des choix techniques d'autres choix économiques, légaux et sociaux vont aussi faire la différence et peser dans l'acceptation du Web. En premier lieu, l'architecture et les fondations technologiques du Web sont *open source* (« code source ouvert »), libres de droits et gratuites. Là encore, le contexte historique est important : la « Fondation pour le logiciel libre » (FSF) a été fondée par Richard Stallman en 1985 pour promouvoir le logiciel libre et elle sera à l'origine du projet GNU et des licences GPL qui, dès 1989, fixent les conditions légales de distribution d'un logiciel libre. Le début des années 90 verra aussi les négociations et le passage d'Unix (BSD) à l'*open source*. En rendant le Web gratuit et son code ouvert, le CERN fait littéralement don du Web au monde. En particulier, en 1992, Tim Berners-Lee et Jean-François Groff travaillent à une version en C qui aboutit à la création en code libre de la célèbre bibliothèque logicielle `libwww` qui sera utilisée par la suite dans de nombreuses implémentations [20]. Tim lui-même n'a jamais cherché à monétiser son travail et défend dès le départ un Web en logiciel libre et un code ouvert. Ce tournant important va permettre la pénétration virale des technologies Web dans toutes les organisations et dans leurs applications.

Une autre initiative décisive est l'établissement en 1994 du World Wide Web Consortium (W3C) qui va jouer un rôle primordial dans la normalisation des évolutions de l'architecture du Web, lui permettant de grandir sans perdre l'interopérabilité standard qui lui donne son universalité. Parce que le CERN est conduit à mettre toutes ses ressources dans la construction du LHC (*Large Hadron Collider*), il annonce fin 1994 qu'il ne peut plus continuer à s'investir dans le projet du Web (appelé projet WebCore). Avec le soutien de la Commission européenne, l'activité Web du CERN est transférée au W3C avec comme membres fondateurs : le MIT aux États-Unis, l'Inria en France et l'Université de Keio au Japon. Jean-François Abramatic, alors directeur des relations industrielles à l'Inria, va jouer un rôle crucial dans ce transfert et il deviendra quelques temps après président du W3C. Par la suite, Inria transférera le rôle d'hôte européen du W3C à l'ERCIM (European Research Consortium for Informatics and Mathematics) que l'institut a aussi contribué à créer avec des partenaires européens en 1989. Avant le W3C, les standards du Web étaient publiés sous forme de RFC (*Request For Comments*). Ils seront dès lors publiés comme des recommandations du W3C. Le consortium n'a cessé de chercher à améliorer son action depuis sa création en affinant et affichant sa mission [21], son

organisation [23], son fonctionnement [24] et son éthique [25]. Les travaux du W3C ont été, et le seront encore, décisifs pour permettre au Web de traverser des crises majeures, qu'elles fussent techniques (p. ex. la guerre des navigateurs et les incompatibilités), commerciales (p. ex. les extensions propriétaires), politiques (p. ex. le standard PICS pour répondre aux inquiétudes quant aux contenus inappropriés), etc. Le W3C ouvrira ainsi très vite des groupes de travail sur de multiples évolutions et facettes du Web comme l'accessibilité ou l'internationalisation. Maintenant, de nombreux groupes de standardisation ou simplement de discussion existent sur une grande variété de sujets.

Évolutions du Web : tisser une toile mondiale de ressources

*« When I took office, only high energy physicists
had ever heard of what is called the World Wide Web...
Now even my cat has its own page. »*

– Bill Clinton, 1996

La relecture que fait Tim de l'architecture du Web en 1996 [5] change un peu ses trois piliers en mettant l'accent sur l'adressage, le protocole et la négociation de contenu. Cette négociation de contenu est un mécanisme natif du protocole HTTP qui offre la possibilité de proposer, pour un même URI, différentes versions d'une même ressource et qui est directement inspiré du mécanisme de négociation de format du « System 33 » de Steve Putz au Xerox PARC. Le langage HTML n'est alors plus considéré que comme l'un des formats disponibles. Ainsi, dès 1994, Tim souligne lui-même que HTTP est peut-être mal nommé puisqu'il n'est pas tant limité au transfert de HTML que destiné à échanger des données arbitraires efficacement dans le contexte de liens et sauts dans un hypermédia distribué [20]. D'une certaine façon le mécanisme de négociation de contenu déclassé HTML en donnant la possibilité de négocier tout type de format dès le départ. Les URI eux permettent d'identifier tous les types de ressources. HTML redevient donc juste un prérequis pour un navigateur [20]. Il aura permis dans un premier temps de fournir un format uniforme de documents hypertextuels et de documentariser le réseau de ressources que devient le Web, avant d'évoluer vers un langage de programmation des applications Web. Inversement, le Web affirme ainsi son indépendance à un modèle ou une structure de données.

En 1994, Tim notait aussi déjà un besoin de faire évoluer le Web vers plus de fonctionnalités temps-réel (p. ex. téléconférence, réalité virtuelle) et d'envisager un support au commerce sur le Web, notamment pour le paiement en ligne [20]. Tim identifie de plus dans cet article de 1996 [5] des directions d'évolution qu'il considère importantes pour le Web et qui ouvrent encore actuellement de nombreuses

directions de recherche et développement. Il cite comme premier axe de travail l'infrastructure et les performances du Web ainsi que le Web comme un espace non seulement social mais aussi comme un espace pour les machines. Il suggère même que cet aspect, que l'on pourrait appeler maintenant des communautés hybrides [36], est vital et il explique que les machines sur le Web sont une nécessité du passage à l'échelle des interactions humaines. Il rappelle aussi un objectif initial toujours non atteint : la possibilité de lier facilement et sûrement les différentes échelles et sphères des documents privés et publics, allant des systèmes d'informations personnelles (PIM) aux systèmes de discussion globale, et d'offrir des outils de groupe à tous les niveaux (personnel, organisationnel, public) en préservant la capacité à lier à travers ces niveaux. Tim déplore aussi qu'en 1996 l'infrastructure de sécurité soit toujours absente du Web, car il est vrai qu'à part l'arrivée de SSL et HTTPS avec Netscape en 1994, peu de progrès ont été réalisés. Il faudra même attendre la RFC2818 en 2000 pour une première spécification formelle officielle de HTTPS. Enfin, Tim insiste encore et toujours sur l'importance de systématiquement rechercher plus de décentralisation : chercher l'ouverture à l'écriture et la contribution de contenus, l'externalisation des liens, l'externalisation des annotations, et l'externalisation des vérifications.

La standardisation de PICS comme l'une des premières recommandations du W3C en 1996 permet de filtrer les contenus inappropriés notamment pour les enfants. C'est aussi un exemple de cette décentralisation voulue : les filtres capturant les préférences des utilisateurs sont créés et stockés dans les clients (navigateur) ; les descripteurs sont stockés sur les sites consultés mais générés par des sites d'autorités tiers. Incidemment, PICS ouvre donc l'idée d'aller plus systématiquement dans l'architecture du Web vers une approche générique pour un problème spécifique, réutilisable pour d'autres scénarios d'usage. Ainsi, l'idée d'étiqueter par et en référence à un site tiers, propose plus génériquement de casser le lien binaire (navigateur/page)-(serveur/site) pour aller vers des liens plus complexes et aussi de ne pas limiter l'étiquetage aux problèmes d'acceptabilité du contenu. On s'ouvre alors à la notion de métadonnées en général sur le Web et on rejoint une autre évolution vers un Web de documents et données structurés.

Dans cette évolution, CSS est une étape importante qui marque le début de la séparation du fond et de la forme sur le Web (CSS vs. HTML). La notion de feuille de style permet de sortir et séparer la mise en forme de la structure du document et aussi de d'utiliser une même mise en forme pour plusieurs documents ou inversement de faire varier la mise en forme d'un même document. Cette étape va permettre aux contenus et aux présentations de se multiplier et de se diversifier de façon indépendante et créative. CSS va aussi marquer le passage du Web à des capacités de présentation professionnelles et très avancées rivalisant à terme avec les productions permises par les meilleurs outils d'édition de documents électroniques. Là encore la notion de feuille de style était présente dans le premier prototype de Tim, mais il

ne publie pas cette fonctionnalité ni la syntaxe qu'il utilise. Cette capacité est donc essentiellement perdue dans les premiers navigateurs qui suivront comme Mosaic. Håkon Wium Lie et Bert Bos seront les premiers à travailler sur une standardisation des feuilles de style qui débouchera en décembre 1996 sur la recommandation CSS level 1. Le succès est aussi dû à Thomas Reardon et Chris Wilson de Microsoft qui dès 1995 assurent qu'ils supporteront CSS. Et en effet, Microsoft Internet Explorer 3 sera la première implémentation de CSS alors que sa spécification n'est encore qu'un brouillon. L'histoire de CSS [45] est aussi l'un des premiers exemples d'activité qui demandera le développement de suites de tests dédiées et de démonstrateurs (p. ex. CSS Zen Garden) pour accélérer l'implémentation et l'adoption d'une recommandation et d'une évolution de l'architecture du Web.

Après la séparation du fond et de la forme, le contenu du Web va pouvoir évoluer en permettant de créer et gérer ses propres structures de documents et données. C'est la standardisation de XML en 1998 et, dans son sillage, de plusieurs langages permettant sa validation (DTD, XML Schemas), son interrogation (XPath, XQuery), sa transformation (XSLT), sa gestion (XProc), etc. À titre personnel, je retiens en particulier à cette époque la feuille de route pour le Web sémantique que Tim publie en 1998 [37]. Cette feuille est dans la continuité de sa présentation de 1994 [19] et aussi de son article de 1994 où l'on peut lire qu'il souhaite une évolution des objets du Web, qui sont à l'époque essentiellement des documents destinés aux humains, vers des ressources avec une sémantique plus orientée vers les machines pour permettre des traitements plus automatisés [20]. La feuille de route de 1998 [37] ouvrira tous les travaux sur le Web de Données et le Web Sémantique (RDF, RDFS, SPARQL, OWL, etc.). En France, Rose Dieng-Kuntz, Olivier Corby et leur équipe à l'Inria identifient immédiatement cette évolution et lancent des recherches sur le sujet [10, 11]. Cette feuille de route sera décisive pour moi qui commence, au sein de cette équipe, en 1999, une thèse sur le couplage des architectures logicielles de l'intelligence artificielle distribuée (modèles multi-agents) et des modèles formels de connaissances distribuées (modèles du Web Sémantique) [12]. À ce jour, Tim est toujours un défenseur passionné des évolutions du Web comme le Web Sémantique [2].

En parallèle de cette évolution des contenus, a aussi lieu la naissance de la programmation Web et des applications Web qui vont devenir un pilier important de la toile et ses usages. Au départ, à la naissance du Web, le principe était très simple : une page générée à la demande suite à un clic sur un simple lien ou en réponse à la soumission d'un formulaire. À chaque interaction, la page était entièrement rechargée. Sur le serveur, des codes écrits en C, C++, Shell ou autres langages utilisent la méthode CGI (*Common Gateway Interface*) pour traiter ces requêtes et générer des pages en retour. Avec les *frames* vers 1996, on commence à pouvoir découper la page en cadres et par la suite ne recharger qu'une partie de ce qui est affiché. La même année, JavaScript fait son apparition et on peut commencer alors à utiliser dans les applications Web, même conjointement, de la programmation côté serveur et de

la programmation côté navigateur. Commence alors une généalogie de langages et techniques de programmation pour le Web, certains plutôt côté serveur (ASP, PHP, C#, Python, Ruby, Perl, JSP, etc.) d'autres plutôt client (p. ex., Plugins, ActiveX, CSS+HTML) et aussi d'autres utilisables des deux côtés (Java Servlet et Applet, JavaScript). Enfin, avec le composant *XMLHttpRequest* proposé par Microsoft en 1998 puis ajouté à JavaScript et rapidement adopté par la plupart des navigateurs, on peut échanger des données entre une page et son serveur, et, grâce au DOM de celle-ci, modifier la page affichée sans nécessairement la recharger. Cette technique sera nommée en 2005 AJAX et massivement adoptée dans les applications Web pour conjuguer du code s'exécutant sur le client et du code s'exécutant sur le serveur et permettre des interactions fluides avec l'utilisateur. En parallèle, l'architecture du Web est de plus en plus étudiée et formalisée comme avec la thèse de Roy Fielding qui introduit l'architecture REST pour la caractériser [34]. Avec le début des années 2000, c'est la proposition d'évoluer vers un Web de services (standards SOAP et WSDL) qui ouvre une nouvelle direction de travail pour l'utilisation du Web comme une plateforme programmatique et qui se réalise plus actuellement à travers les API et les langages liés comme JSON. Certains vont jusqu'à parler du Web comme un système d'exploitation au-dessus de la collection mondiale de services qui s'offrent sur Internet, et indépendant des ordinateurs et objets individuels qui s'y connectent. Ils positionnent le Web comme l'environnement de programmation et d'exécution par excellence des applications de l'Internet. Dans cette mouvance, l'une des prochaines évolutions actuellement étudiée est de faire du Web la plateforme applicative universelle de l'Internet des objets que l'on nomme le Web des Objets (*Web of Things* [38]). De nos jours, on envisage littéralement de faire une toile de tout. Mais si l'on revient cependant à cette influence initiale que RPC a eu sur la conception d'un hypertexte distribué, on peut en fait voir cette évolution comme un retour aux sources. De plus, en se représentant le Web comme une toile d'appels potentiels de procédures que l'on invoque à chaque lien que l'on suit, on comprend mieux pourquoi des ambitions comme celles de moissonner (*crawler*), indexer ou archiver le Web sont compliquées voire paradoxales.

Toujours en parallèle, dès 1996, des compagnies comme Nokia en Finlande vont s'intéresser à proposer un accès au Web sur les téléphones mobiles. En 1999, se sera NTT DoCoMo au Japon avec le i-mode. Cette même année, le code QR, créé par l'entreprise japonaise Denso-Wave en 1994, passe sous licence libre ce qui va contribuer à sa diffusion et sa mise à disposition dans des applications de reconnaissance sur les téléphones mobiles notamment pour glaner des URL affichés autour de nous. À cette époque aussi, WAP (*Wireless Application Protocol*) et WML (*Wireless Markup Language*) seront proposés pour adapter l'accès et les contenus Web aux contraintes des téléphones portables et de leur connectivité. Ils seront par la suite abandonnés lorsque les téléphones et réseaux atteindront des performances leur permettant de directement accéder au Web et à l'Internet classiques. Ces premiers essais,

ainsi que ceux faits avec des PDA ayant une carte réseau sans fil, devront attendre le milieu des années 2000 et l'avènement des smartphones pour trouver des plateformes au-dessus desquelles réaliser leur plein potentiel. C'est en 2004 que le W3C en lancera l'initiative Web mobile (MWI). En 2005, le nom de domaine de premier niveau « .mobi » est proposé mais sera critiqué, notamment par Tim, comme une solution qui casse l'indépendance du Web vis-à-vis du terminal. En 2007, dans une conférence invitée au congrès GSM, Tim se positionne aussi contre les prés carrés créés par les plateformes propriétaires et défend la plateforme ouverte qu'offre le Web [46]. Dans un premier temps, les problématiques auront donc été de pallier les limitations d'une connexion mobile (écran, bande passante, interactions limitées, puissance de calcul limitée, coût de connexion, etc.). Puis, soit avec la résolution de ces problèmes (p. ex. compression), soit avec leur disparition (p. ex. montée en puissance des terminaux et réseaux), les problématiques de cette évolution du Web vont progressivement passer de l'accès mobile simple à l'adaptation plus profonde à une utilisation en situation de mobilité (p. ex. géolocalisation, adaptation des interactions et interfaces, accès aux données personnelles, contextualisation, interaction audio et vocale, réalité augmentée, couplage de plusieurs terminaux, etc.). Nous avons maintenant de façon courante des applications Web mobiles, et beaucoup d'applications mobiles natives sont dans les faits grandement développées avec des langages et standards du Web. De plus, les prix d'entrée de gamme et la démocratisation des smartphones font que dans certaines régions du monde, le mobile est maintenant la première façon d'accéder au Web et le moyen le plus répandu d'avoir un premier contact avec le Web [39].

Entre Web de données, Web d'applications, Web de services, Web mobile, mais aussi Web multimédia, accessible, internationalisé, etc., le Web commence donc très tôt sa transformation vers une architecture de programmation et d'interaction hypermédia générique et surtout sa généralisation à une toile liant potentiellement tous types de ressources computationnelles ou non. Le Web peut toucher à tout puisque tout peut être identifié par un URI. Les principes du Web étant extensibles et génériques, ils nous ont permis de passer d'une vision documentaire de bibliothèque mondiale à un réseau de ressources protéiformes. L'une des plus grandes forces du Web est dans son universalité mais nous verrons qu'elle demande une constante attention pour être préservée.

Catégories et déclinaisons : une toile à facettes

*« As soon as you externalize an idea
you see facets of it that weren't clear
when it was just floating around in your head. »*

– Brian Eno

Pour parler des temps forts des évolutions du Web on trouve maintenant les appellations Web 1.0, Web 2.0, Web 3.0, etc., par lesquelles je ne suis pas convaincu car elles laissent penser qu'elles correspondent à des évolutions logicielles majeures du Web alors qu'elles sont plus souvent des évolutions des pratiques ou même de la compréhension que nous avons du Web et qu'elles ne rendent pas compte des multiples directions dans lesquelles le Web évolue en parallèle. Le Web 1.0 correspond essentiellement à la vision initialement documentaire et d'hypermédia distribué du Web. Le Web 2.0 est aussi appelé Web social et rend compte à la fois de la réouverture en écriture du Web, de l'approche AJAX pour l'interaction, et de la contribution et des échanges sociaux massifs qu'ils permettent avec les wiki, les blogs, les forums, les médias sociaux, etc., avec l'impact que l'on connaît. Le Web 3.0 recouvre en général l'intégration des pratiques du Web sémantique et du Web de données au Web 2.0, par exemple avec RDFa dans le protocole OGP de Facebook ou dans l'utilisation de `Schema.org` dans de nombreux sites intégrant des fonctionnalités sociales (p. ex. votes, critiques, etc.). Cependant, comme nous l'avons vu dans la section précédente, plus que des évolutions par sauts de versions, le Web vit en permanence un bouquet d'évolutions concurrentes qui demandent un travail constant et considérable au W3C pour rester compatibles, mais qui en même temps, dans une approche évolutionniste, lui permettent de lancer de multiples sondes à la recherche de ses prochaines mutations et de leurs croisements.

Un autre jeu de facettes que l'on reconnaît au Web sont : le Web de surface, le Web profond et le Web obscur. Le Web de surface est le Web indexable et parcouru par les moteurs de recherche (services appelés *crawlers*). Il est public et forme la partie la plus émergée du Web. Il compte aussi de nombreuses pages d'entrée de portails et applications qui eux ouvrent sur le Web profond (*deep Web*) avec des modalités d'accès, de parcours et de recherche dédiées. Le Web profond est aussi injustement qualifié de caché ou invisible alors que nous le voyons tous les jours. Il s'agit du Web essentiellement généré dynamiquement comme les pages de résultats de recherche qui nous sont bien visibles et accessibles mais qui ne sont pas indexées par les moteurs pour différentes raisons : contenus ou liens générés dynamiquement, contenus accessibles après une authentification ou tout autre formulaire ou interaction complexe allant au-delà du suivi d'un lien, contenus non liées au reste du Web, contenus avec une politique d'exclusion des robots d'indexation (fichier robots.txt), etc. Cette partie dite profonde du Web représente en fait la majeure partie des contenus de la toile.

Le terme de *deep Web* est parfois confondu à tort avec les termes *darknet* et *dark-web* qui ont pourtant un sens bien différent. Le Web obscur (*dark Web*), Web sombre ou Web noir par équivalence au terme de matière noire, est un Web de l'ombre utilisé dans des activités recherchant l'anonymat que ce soit, par exemple, un opposant politique cherchant à échapper à l'oppression de son pays ou une organisation criminelle cherchant à utiliser la puissance des outils du Web pour son activité. Les

toiles de ce Web sont volontairement déconnectées du Web classique pour ne pas être trouvées et indexées. Ces toiles sont tissées sur l'Internet obscur (*darknet*) qui utilise des techniques d'anonymisation, de cryptographie, de réseaux (p. ex. pair-à-pair) et de sécurité en général pour masquer les identités et échanges des usagers. En combinant ces techniques avec l'architecture Web, on peut tisser et naviguer sur des toiles cachées (p. ex. le navigateur TOR). Les termes *dark Web* et *darknet* sont souvent utilisés de façon interchangeable alors que la notion de *darknets* désignant des réseaux volontairement isolés existait déjà à l'époque d'Arpanet dans les années 1970, donc bien avant le Web, et leur utilisation applicative s'étend à bien d'autres services (mail, IRC, forum, etc.).

Enfin, on peut aussi mentionner les IntraWeb (l'utilisation du Web en intranet) qui, derrière des VPN et pare-feux (*firewall*) d'entreprise, utilisent au sein de nos organisations les solutions du Web pour créer des toiles réservées à leurs membres. Là encore, des techniques de sécurité sont utilisées pour contrôler l'accès à ces composantes du graphe du Web. On ne cherche pas forcément à éviter les liens entre ces toiles et le Web public, mais on en sécurise l'accès.

Les évolutions de la section précédente ou les facettes que nous venons de mentionner relèvent cependant d'une seule et même architecture : l'architecture standardisée du Web. Comme insiste la description de la mission du W3C (*One Web*), il ne s'agit pas de différents Webs mais de différentes facettes d'un et un seul Web, d'une et une seule architecture [21]. Au sein même du W3C, il existe d'ailleurs un groupe spécial appelé le TAG (Technical Architecture Group) [22] où siège notamment Tim au titre de Directeur du W3C et qui veille, notamment, à ce que toutes les évolutions du Web restent compatibles et continuent à respecter et se combiner en une architecture cohérente et universelle. Ces quelques exemples des évolutions et facettes du Web nous amènent tout de même à un nouveau besoin : celui de se doter de moyens d'étudier le Web et ses évolutions.

L'inconnu du Web

« Web science – what makes the Web what it is, how it evolves and will evolve, what are the scenarios that could kill it or change it in ways that would be detrimental to its use. »

– Dame Wendy Hall

Aujourd'hui encore, il est frappant de voir à quel point le Web est à la fois très connu et à la fois mal connu comme en témoigne la confusion tenace entre les termes Web et Internet que l'on rencontre encore bien trop souvent. Malgré le fait que leurs inventeurs respectifs aient reçu deux prix Turing bien distincts, respectivement en 2004 et en 2016 pour deux inventions bien différentes, Internet et Web sont encore

trop souvent utilisés de façon interchangeable. Redisons le : Internet permet l'interconnexion des réseaux d'ordinateurs et objets connectés en général. Il fournit une infrastructure de communication qui supporte au-dessus d'elle de nombreuses applications comme : la messagerie électronique (*mail*), la téléphonie et la vidéophonie... et le Web, cet hypermédia distribué qui devient l'architecture logicielle majoritaire des applications sur Internet.

Outre cette confusion, le terme Web est aussi souvent utilisé de façon indifférenciée pour se référer à la fois aux principes fondateurs de cette architecture logicielle et à l'objet qui en émerge, i.e. la toile tissée par des milliards d'utilisateurs. Dès 1994, Tim note que le terme de World Wide Web a très rapidement recouvert plusieurs choses et notamment : d'un côté une architecture et de l'autre un jeu de données mises à disposition sur Internet selon cette architecture [20]. L'architecture et l'objet qui en émerge ont deux histoires liées mais portent sur des aspects différents. Chacune des deux facettes exhibe cependant de façon différente une complexité qui nécessite à la fois recherche et développement.

En effet, l'architecture du Web repose sur des protocoles, des modèles, des langages et des algorithmes qui nécessitent d'être spécifiés, conçus, caractérisés et validés avec de plus, systématiquement, des contraintes comme celles du passage à l'échelle, de l'efficacité en temps et en mémoire, d'interopérabilité et d'internationalisation. Pour cela, l'architecture du Web et ses extensions sont des objets de recherche notamment dans les sciences du numérique. Au sein de l'informatique et des sciences du numérique, et nous verrons dans la section suivante que c'est valable pour de nombreux autres domaines aussi, le Web s'est répandu à la fois comme un nouvel outil de travail mais aussi comme un nouveau sujet apportant à la fois des solutions et des nouveaux problèmes et besoins.

Quant à l'objet qui émerge de l'usage de cette architecture, la complexité de ses usages, l'hétérogénéité et les volumes de ses contenus, services et données, la dynamique de certains de ses flots ou les cycles de vie de ses ressources et communautés sont autant de sources de complexité qui à nouveau requièrent une approche scientifique et des recherches théoriques, appliquées, expérimentales et multidisciplinaires. Une proposition, parmi d'autres, illustrant cette idée d'étudier le Web comme un objet complexe est, par exemple, d'avoir des observatoires du Web [40] avec des instruments d'observation et des méthodes des sciences expérimentales pour étudier le Web ; se doter en quelque sorte de microscopes ou télescopes du Web et de la méthode scientifique pour « tourner ce télescope vers les toiles ».

Et plus le Web grandit en complexité architecturale et en ressources liées dans sa toile, plus il appelle des recherches et des développements transdisciplinaires [41]. Tim ira même jusqu'à dire qu'*in fine* le Web est plus une création sociale qu'une création technique [3].

Historiquement, le Web est rapidement devenu un objet de recherche et, là encore sous l'impulsion de Robert Cailliau et Tim Berners-Lee, commence au CERN dès

1994 le cycle des conférences WWW, « *The Web Conference* », devenu le rendez-vous annuel de la recherche, du développement et des industries et acteurs majeurs du Web. Pour avoir accepté d'être deux fois co-président de cette conférence, en 2012 et en 2018, je peux attester de l'importance et du niveau des recherches scientifiques que le Web suscite. Et cette communauté de recherche grandit et se diversifie avec des initiatives résolument multidisciplinaires comme « Web Science » ou des conférences plus spécialisées comme ISWC, WI, WebIST, etc. Pour conclure ce volet, notons que Tim est devenu Directeur en 2006 de la Web Science Research Initiative (WSRI), un programme de recherche scientifique portant sur les aspects techniques et sociaux qui sous-tendent les évolutions du Web [14] et qu'il est aussi fondateur du Web Science Trust [15], une organisation britannique à but non lucratif consacrée à l'étude interdisciplinaire du Web et de ses effets sur la société [2].



Extrait central du Panel final à la conférence WWW 1994. De gauche à droite : Dr. Joseph Hardin, Robert Cailliau, Tim Berners-Lee et Dan Connolly (Photothèque CERN).

Où l'on bascule vers un Web-Wide World

*« We can only see a short distance ahead,
but we can see plenty there that needs to be done. »*

– Alan Turing

« *This is for everyone* », c'est avec ce message que Sir Tim Berners-Lee présente le Web lors de la cérémonie d'ouverture des Jeux olympiques d'été de 2012 à Londres. Ce tweet devenu célèbre a inspiré le titre volontairement ambivalent de cet article (« pour tout le monde ») car désormais, le Web s'adresse et touche tout le monde et ce faisant le Web s'intègre à tout autour de nous et se déploie dans les moindre recoins de notre monde. Du Web, pour tous, partout et pour tout. Cette toile d'envergure mondiale, appellation qui au départ a pu être perçue comme immodeste, se révèle en quelques années être une prophétie auto-réalisatrice où le fait de concevoir pour l'universalité aura effectivement permis de tendre à l'universel.

Le Web a modifié notre rapport avec le temps et l'espace en nous donnant la possibilité d'interagir avec des objets ou des personnes éloignées, en nous donnant accès à des informations non disponibles localement, en enregistrant les traces de nos actions, en documentarisant une variété d'activités et en nous permettant ainsi de nous y replonger et d'y naviguer *a posteriori* et à distance. On ne note plus l'adresse du dentiste, on la retrouve sur le Web. On ne programme plus son magnétoscope ou sa box, on cherche un enregistrement en ligne. On ne passe plus à la gare acheter ses billets mais on télécharge son billet électronique. L'omniprésence et l'hypermnésie du Web sont acquises au point où l'on ne supporte plus quand il n'est pas là pour nous répondre et où la page « *no results found* » des moteurs de recherche n'est presque plus jamais vue.

Le Web est aussi un formidable outil d'intégration et d'interopérabilité, un espace d'échange entre applications. Les lignes que vous lisez actuellement ont été partiellement écrites à travers l'interface Web d'une application d'édition collaborative de documents et partiellement dictées à travers une application mobile connectée au travers du Web à cette même application. L'orthographe et la grammaire ont été grandement corrigées à la volée par des outils ayant statistiquement appris du Web. Se connecter à de telles ressources, les intégrer, les synchroniser et en assurer l'accessibilité quels que soient l'application et le terminal utilisés, tout ceci est assuré par la standardisation du Web.

À l'échelle de l'histoire de l'informatique, le Web n'a pas uniquement fait ses preuves d'une architecture qui passe le test du temps, il a défini un nouveau temps : il existe une époque avant et une époque après le Web. Une époque où le problème est d'avoir des informations ou l'accès à un service, puis une époque où il faut pouvoir se retrouver dans leur masse. Une époque de fragmentation puis une époque d'hyper-intégration voire de sur-intégration avec ses risques.

Le Web a instauré un certain nombres de nouveaux objets et concepts maintenant courants comme : le serveur Web, le navigateur Web, la page Web, la page perso, le site Web, l'adresse Web ou l'application Web. Certains comme le « site Web » restent encore informellement définis. Le Web a aussi hybridé beaucoup d'objets et d'activités existantes : web documentaire, web series, webisode, web tv, web radio, webinar, web sourcing, web publishing, web ID, web mail, web commerce, web publicité,

webcam, webcast, weblog, web conf, web journal, webOS, etc. Le plus effrayant à écrire ces termes c'est que mon correcteur orthographique les reconnaît déjà. Dès qu'une ressource Web s'instancie dans un type d'objet, dès lors que ces objets s'hybrident et entrent en contact avec les principes et pratiques du Web, ils deviennent de nouveaux objets appelant de nouvelles pratiques. Et le phénomène est fractal : si une technologie Web se développe en restant au plus proche de son architecture et de ses principes universels (p. ex. wiki), elle acquiert sa viralité et le potentiel de se décliner de multiples façons (p. ex. wikipedia, wiktionary, wiki travel, wikileaks, etc.). Sorti de la chrysalide initiale de la métaphore d'une bibliothèque universelle, le Web comme outil de partage est maintenant tellement vulgarisé qu'il fournit même des métaphores pour d'autres domaines, comme par exemple les réseaux mycorhiziens appelés par certains biologistes des *Wood-Wide Web* [9].

Comme on peut le voir au travers des références et allusions de cet article, Tim continue à défendre le Web et à s'investir sans compter pour amener le Web à son plein potentiel². Ce prix Turing récompense non seulement le fait qu'il ait inventé le Web mais aussi le fait qu'il ait travaillé toute sa vie à le défendre. Car la défense du Web reste un enjeu. Il est universellement utile et utilisé mais reste fragile et son idéal de départ pourrait n'être qu'une parenthèse historique si l'on ne veille pas en permanence à sa préservation.

Tim Berners-Lee se bat encore actuellement contre toute forme de recentralisation, par exemple la centralisation applicative induite par des monopoles de certaines firmes sur certains services du Web, et pour la neutralité des réseaux en général et du Web en particulier. Les enjeux (neutralité, décentralisation, démocratisation, etc.), les dangers (recentraliser, classes d'accès, Web à plusieurs vitesses, etc.), les limitations (besoins en infrastructure, énergies, coûts, etc.) font du Web un interminable projet plus qu'une réalisation acquise. La plus grande crainte de Tim reste qu'une entité politique ou commerciale prenne le contrôle du Web ce qui, pour lui, signerait l'arrêt de mort du Web. Cela motive son engagement en faveur, entre autres, de la neutralité du net, de la redécentralisation du Web, de l'éclatement des bulles de filtrage [42] et de la réappropriation des données par les usagers [13]. L'architecture du Web est et doit rester robuste, même en milieu hostile, neutre, même si les couches basses venaient à être compromises, résiliente, même si les infrastructures venaient à manquer ou être limitées, etc. Pour le protéger, nous devons concevoir les araignées de l'architecture du Web comme des animaux extrêmophiles.

À l'inverse, le Web a pu lui-même être perçu comme un danger. Dès 1996, Tim écrit que la force de diversification qu'est la géographie est affaiblie par le Web [5]. Il le fait à une époque où, alors que je donnais mes premiers cours de Web à mes co-promotionnaires du Génie Mathématique de l'INSA de Rouen, on entendait en France certains s'inquiéter de voir le Web devenir un outil d'hégémonie de la langue

2. Le slogan du W3C : « *leading the Web to its full potential.* »

anglaise et de la culture anglophone. Cette année-là, Tim attire déjà l'attention de façon générale sur l'éthique et les enjeux sociétaux du Web : il rappelle l'impact des choix architecturaux du Web sur les formes de sociétés dans lesquelles nous vivons ; la nécessité de revisiter la notion de copyright dans un espace où la copie peut prendre bien des formes (p. ex. la mise en cache d'une copie d'un contenu et le statut juridique de cette copie) ; les problèmes que posent au respect de la vie privée les multiples opportunités de capter des données ; l'impact en termes d'information du Web sur un public de votants ; la nécessité de travailler main dans la main avec les systèmes de législation ; etc. [5].

À mon avis, ces sujets préfigurent en 1996 des besoins souvent encore plus impérieux deux décennies après, et la nécessité de chercher activement l'interdisciplinarité dans l'étude du Web et de ses évolutions. Nous reconnaissons tous que le World Wide Web a eu un impact social énorme [2]. Le Web est un objet de recherche, de développement, d'activité économique et commerciale (géants du numérique), un vecteur d'action et de structuration sociale, un sujet et un outil de politique, un nouvel objet et espace juridique, et même de questionnement philosophique [8]. Il est donc important pour comprendre le Web de façon holistique, dans toute sa complexité, d'encourager le mouvement « Web Science » vers la transdisciplinarité. Pour moi, les trois W du World Wide Web appellent les trois M d'une Méthode Massivement Multidisciplinaire [41].

De plus, le Web doit aussi résolument devenir un sujet d'éducation et de formation en lui-même. Son utilisation (notions de base de navigation, de recherche, etc.), les bonnes pratiques (lecture critique, validation croisée, contribution active, etc.), la prévention (protection de la vie privée, protection des enfants, etc.), sont autant de sujets auxquels toute génération devrait être formée à l'école comme un élément important dans l'égalité des chances que l'on doit. En 2008, Tim est ainsi un des fondateurs de la World Wide Web Foundation, organisation à but non lucratif promouvant l'accès au Web pour tous [13] et un Web ouvert comme un bien public et un droit élémentaire [2].

L'universalité de l'approche et le pouvoir de lier tout ce qui est concevable et partout dans le monde étaient difficiles à comprendre au début, et peu de personnes voyait la différence avec les systèmes hypertextes existants. Si à l'origine, le problème était d'imaginer un monde avec le Web avant que celui-ci n'advienne, nous sommes maintenant dans le cas inverse où les gens oublient ou n'imaginent plus ce que serait un monde sans le Web [13]. Nous sommes à un point de bascule où l'on parle de réalité augmentée par le Web [43] et où la perception s'inverse tant et si bien que c'est le monde qui paraît s'inscrire dans le Web, plus ce dernier s'y déploie. Cette idée est pour moi résumée dans l'inversion des termes *Web-Wide World* où le Web dépasse le monde physique en 3D et ce dernier y rentre pour s'étendre et s'augmenter dans un nombre ouvert de dimensions. Et dans cette perspective, les échos de

la phrase de Tim sur le fait que nos choix dans l'architecture du Web ont un impact sur nos sociétés ne cesse de s'amplifier avec les évolutions du Web.

Le prix Turing récompense donc Sir Tim Berners-Lee non seulement pour l'invention du Web en 1989 mais aussi pour le fait que, depuis, Tim n'a jamais cessé de remettre la toile du Web sur de nombreux métiers.

« *The Web as I envisaged it, we have not seen it yet.
The future is still so much bigger than the past.* »

– Sir Tim Berners-Lee, WWW Conference 2009



De gauche à droite : Fabien Gandon, Sir Tim Berners-Lee, Louis Pouzin, Jean-François Abramatic, pour les 25 ans du Web à Futur en Seine, Paris, 2014.

Remerciements. Je tiens à remercier mes collègues et amis du W3C, de la SIF, d'Inria, d'UCA, du CNRS et le l'IS3S pour leur relecture de cet article.

Références

- [1] W3C Automotive Working Group Charter, accédé le 10 mai 2017, <https://www.w3.org/2014/automotive/charter.html>
- [2] ACM Turing Award Sir Tim Berners-Lee, http://amturing.acm.org/award_winners/berners-lee_8087960.cfm
- [3] Weaving the Web : The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor, Tim Berners-Lee, HarperCollins, 1999, ISBN 0062515861, 9780062515865
- [4] How the Web was Born : The Story of the World Wide Web, James Gillies, Robert Cailliau, January 15, 2000, Oxford University Press, ISBN-10 : 0192862073 ISBN-13 : 978-0192862075
- [5] The World Wide Web : Past, Present and Future, Tim Berners-Lee, August 1996, <https://www.w3.org/People/Berners-Lee/1996/ppf.html>

- [6] “Information Management : A Proposal” (March 1989), the original proposal for the software project at CERN that became the World Wide Web. <https://www.w3.org/History/1989/proposal.html>
- [7] Sergey Brin, Lawrence Page, The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks* 30(1–7): 107–117 (1998)
- [8] Philosophical Engineering and Ownership of URIs, Tim Berners-Lee, an interview in “Philosophical Engineering”, collected set of writings from the PhiloWeb workshops, <https://www.w3.org/DesignIssues/PhilosophicalEngineering.html>
- [9] *La Vie secrète des arbres*, Peter Wohlleben, Editions Les Arènes, 2017, 2352045932
- [10] Cédric Hébert, Modèle de traitement de RDF basé sur les graphes conceptuels, Rapport de stage de DEA, I3S, université de Nice Sophia-Antipolis, 1999.
- [11] Olivier Corby, Rose Dieng, Cédric Hébert. A Conceptual Graph Model for W3C Resource Description Framework. International Conference on Conceptual Structures, Aug 2000, Darmstadt, Germany. pp. 468–482, 2000
- [12] Fabien Gandon, Distributed Artificial Intelligence And Knowledge Management : Ontologies And Multi-Agent Systems For A Corporate Semantic Web, Thèse, Université Nice Sophia Antipolis, 2002.
- [13] Neil Savage, Weaving the Web, *Communications of the ACM*, Vol. 60 No. 6, Pages 20–22, 10.1145/3077334, <https://cacm.acm.org/magazines/2017/6/217732-weaving-the-web/fulltext>
- [14] MIT and University of Southampton launch World Wide Web research collaboration Initiative will analyze and shape web’s evolution, MIT News, November 2, 2006, <http://news.mit.edu/2006/wsri>
- [15] Web Science Trust (WST), accédé le 7 juin 2017, <http://www.webscience.org/>
- [16] “World Wide Web—Archive of world’s first website”. World Wide Web Consortium. <https://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html>
- [17] Welcome to Amaya, <https://www.w3.org/Amaya/>
- [18] <https://timeline.web.cern.ch/timelines/The-birth-of-the-World-Wide-Web>
- [19] Tim Berners-Lee, Plenary Talk extracted slides, First WWW Conference, Geneva 94, <https://www.w3.org/Talks/WWW94Tim/>
- [20] Berners-Lee, Tim and Cailliau, Robert and Luotonen, Ari and Nielsen, Henrik Frystyk and Secret, Arthur, The World-Wide Web, *Commun. ACM*, Aug. 1994, Vol. 37, n° 8, 0001-0782, pp. 76–82, 10.1145/179606.179671, ACM, New York, NY, USA
- [21] W3C mission, accédé le 16 juin 2017, <https://www.w3.org/Consortium/mission>
- [22] The W3C Technical Architecture Group (TAG), accédé le 16 juin 2017, <https://www.w3.org/2001/tag/>
- [23] W3C organisation, accédé le 16 juin 2017, <http://www.w3.org/Consortium/facts>
- [24] W3C Process, accédé le 16 juin 2017, <http://www.w3.org/Consortium/Process/>
- [25] W3C Ethics, accédé le 16 juin 2017, <http://www.w3.org/Consortium/cepc/>
- [26] Luis Von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford. 2003. CAPTCHA : using hard AI problems for security. In Proceedings of the 22nd international conference on Theory and applications of cryptographic techniques (EUROCRYPT’03), Eli Biham (Ed.). Springer-Verlag, Berlin, Heidelberg, pp. 294–311.
- [27] Von Ahn, Luis, et al. “recaptcha : Human-based character recognition via web security measures”, *Science* 321.5895 :1465–1468 (2008).
- [28] Bush, Vannevar. “As we may think”. *The atlantic monthly* 176.1:101–108 (1945).

- [29] T. H. Nelson. 1965. Complex information processing : a file structure for the complex, the changing and the indeterminate. In Proceedings of the 1965 20th national conference (ACM '65), Lewis Winner (Ed.). ACM, New York, NY, USA, pp. 84–100. <http://dx.doi.org/10.1145/800197.806036>
- [30] C. Engelbart, and William K. English, AFIPS Conference Proceedings of the 1968 Fall Joint Computer Conference, San Francisco, CA, December 1968, Vol. 33, pp. 395–410 (AUGMENT,3954,).
- [31] Pouzin, L., Presentation and major design aspects of the Cyclades Computer Network. IN : Proc. 3rd Data Communications Symposium, Tampa, Fla., Nov. 1973, pp. 80–85.
- [32] Cerf, V. G. and Kahn, R. E., A protocol for packet network intercommunication. IEEE Trans. Commun., Vol. COM-22, No. 5, pp. 637–648 (May 1974).
- [33] Roy T. Fielding and Richard N. Taylor. 2000. Principled design of the modern Web architecture. In Proceedings of the 22nd international conference on Software engineering (ICSE '00). ACM, New York, NY, USA, pp. 407–416, <http://dx.doi.org/10.1145/337180.337228>
- [34] R. Fielding, J. Reschke, Hypertext Transfer Protocol (HTTP/1.1) : Semantics and Content, RFC7231, Internet Engineering Task Force (IETF) , 2014
- [35] Tim Berners-Lee's proposal, CERN, March 1989, <http://info.cern.ch/Proposal.html>
- [36] Fabien Gandon, Michel Buffa, Elena Cabrio, Olivier Corby, Catherine Faron-Zucker, et al.. Challenges in Bridging Social Semantics and Formal Semantics on the Web. Hammoudi, S. and Cordeiro, J. and Maciaszek, L.A. and Filipe, J. 5h International Conference, ICEIS 2013, Jul 2013, Angers, France. Springer, 190, pp. 3–15, 2014, Lecture Notes in Business Information Processing. <https://hal.inria.fr/hal-01059273>
- [37] Berners-Lee, Tim, “Semantic web road map”, <https://www.w3.org/DesignIssues/Semantic.html> (1998).
- [38] W3C Begins Standards Work on Web of Things to Reduce IoT Fragmentation, <https://www.w3.org/WoT/>
- [39] The Mobile Web, World Wide Web Foundation, <http://webfoundation.org/about/vision/the-mobile-web/>
- [40] Tiropanis, Thanassis, et al. “The web science observatory”. IEEE Intelligent Systems 28.2 (2013), pp. 100–104.
- [41] Fabien Gandon. The three ‘W’ of the World Wide Web callfor the three ‘M’ of a Massively Multidisciplinary Methodology. Valérie Monfort; Karl-Heinz Krempels. 10th International Conference, WEBIST 2014, Apr 2014, Barcelona, Spain. Springer International Publishing, 226, Web Information Systems and Technologies. <http://www.springer.com/fr/book/9783319270296>, <https://hal.inria.fr/hal-01223236>
- [42] Pariser, Eli. The filter bubble : What the Internet is hiding from you. Penguin UK, 2011.
- [43] Fabien Gandon, Alain Giboin. Paving the WAI : Defining Web-Augmented Interactions. Web Science 2017 (WebSci17), Jun 2017, Troy, NY, United States. pp. 381–382, 2017, <https://hal.inria.fr/hal-01560180>
- [44] Michel Buffa, Du Web aux wikis : une histoire des outils collaboratifs, Interstices, 23 mai 2008, https://interstices.info/jcms/c_37151/du-web-aux-wikis-une-histoire-des-outils-collaboratifs
- [45] A brief history of CSS until 2016, Bert Bos, Style Activity Lead, W3C, 17 December 2016, <https://www.w3.org/Style/CSS20/history.html>
- [46] The Mobile Web, Tim Berners-Lee, keynote, GSM congress, Barcelone, 2007



Les chapeaux de l'icosaèdre

Jean-Paul Delahaye¹

La rubrique « Récréation informatique » propose une petite énigme algorithmique ou sur un thème de mathématiques discrètes susceptible d'intéresser un lecteur de 1024. La solution est donnée dans le numéro suivant.

Rappel et solution du problème précédent

CARRÉ MAGIQUE IMPOSSIBLE ?

Dans le dessin de ce célèbre carré magique, chaque alignement de trois pastilles contenant des nombres donne un total de 15. Il y a huit alignements possibles : trois horizontaux, trois verticaux, et deux diagonaux. L'énigme paradoxale inventée par Le Sal-lows est la suivante : repositionner les pastilles, en plaçant toujours une pastille dans chaque case du tableau 3×3 , de manière à ce qu'il y ait toujours huit alignements de trois pastilles, chacun donnant un total que cette fois on veut égal à 16.

⑧	①	⑥
③	⑤	⑦
④	⑨	②

La chose paraît impossible, puisqu'il semble que si c'est faisable avec une somme de 15, on ne peut certainement pas

1. Université de Lille 1, Sciences et Technologies, Centre de recherche en informatique signal et automatique de Lille (CRISAL), UMR 9189 CNRS, Bât M3-ext, 59655 Villeneuve d'Ascq Cedex.
E-mail : jean-paul.delahaye@univ-lille1.fr.

le faire avec une somme plus grande. Il y a pourtant une solution... et même plusieurs. Il n'y a aucune entourloupe, mais il faut remarquer qu'on ne dit pas que les pastilles doivent se trouver au centre de chaque case.

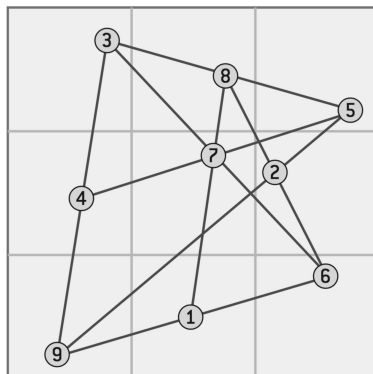
SOLUTION.

Déplacer le centre des pastilles était suggéré. La solution de Lee Sallows est donnée ci-contre.

François Lemaire, du laboratoire CRISTAL de l'université de Lille (<http://www.lifl.fr/~lemaire/homepage/index.php>), a étudié le problème. Il a trouvé 23 solutions. Je cite maintenant le petit article où il décrit rapidement sa méthode qui est numérique.

« Les cases font chacune une largeur de 1. Si les pions sont à une distance de 0.00001 au moins de chaque bord, alors on trouve 23 solutions. Si la distance est de 0.001, on trouve 20 solutions. Si la distance est de 0.04, on trouve 18 solutions. Il n'y a aucune garantie d'avoir trouvé toutes les solutions. Après avoir retiré toutes les symétries, et les cas impossibles, il en reste au plus 82. C'est sur ces 82 qu'on en trouve 23. La méthode est numérique, et consiste à aligner chaque triplet de nombres par un mécanisme de type ressort : chaque triplet s'aligne progressivement au cours du temps (un peu comme l'intégration d'une équation différentielle), et on s'arrête quand les alignements sont quasi parfaits. Important : il faudrait vérifier que chaque solution est bien correcte en passant par les rationnels. »

Je tiens à la disposition de ceux qui le souhaite son document. Pour avoir une réponse définitive sur le nombre de solutions, il faudrait encore approfondir le sujet.



Nouveau problème

LES CHAPEAUX DE L'ICOSAÈDRE

L'icosaèdre est un des cinq polyèdres réguliers de Platon (pour chacun d'eux, chaque face est un même polygone régulier, et tous les sommets ont exactement la même forme). L'icosaèdre est le polyèdre composé de 20 triangles équilatéraux assemblés comme indiqué page suivante.



À un sommet de l'icosaèdre se rejoignent cinq triangles équilatéraux. Si on considère uniquement ces cinq triangles, cela constitue une sorte de chapeau. La question posée est :

Combien faut-il de tels chapeaux (chacun placé pour s'ajuster à cinq faces de l'icosaèdre) au minimum pour recouvrir l'icosaèdre ?

Comme il y a 20 faces et que chaque chapeau est fait de cinq triangles équilatéraux, il faut au moins quatre chapeaux pour recouvrir l'icosaèdre. Mais est-ce que quatre chapeaux peuvent réussir ce recouvrement ? Si quatre ne suffisent pas, est-ce que cinq suffisent ? Si cinq ne suffisent pas est-ce que six suffisent ? etc.

On peut bien sûr concevoir un programme qui énumère tous les choix possibles avec quatre, puis cinq, puis six chapeaux, etc., jusqu'à trouver une solution. Mais comment raisonner pour éviter d'avoir à utiliser son ordinateur ?

Envoyez vos réponses à jean-paul.delahaye@univ-lille1.fr. Le nom des premiers lecteurs à me donner la bonne réponse (et à la justifier) seront mentionnés dans le prochain numéro de 1024.



1024
B U L L E T I N
de la société informatique
de France



Institut Henri Poincaré,
11 rue Pierre et Marie Curie,
75231 Paris Cedex