

PLAN DE COURS

PHI-7706 : Sujets spéciaux (philosophie)

NRC 91303 | Automne 2020

Mode d'enseignement : Comodal

Temps consacré : 3-0-6	Crédit(s) : 3
------------------------	---------------

Il s'agit de cours proposés, de façon ponctuelle, à l'occasion de la présence à l'Université Laval de professeurs invités ou encore en raison de thématiques particulières.

La formation comodale combine de façon simultanée les modes de formation en présentiel et à distance synchrone. Chaque séance peut être suivie aussi bien en classe qu'à distance ce qui permet à l'étudiant de choisir sur une base hebdomadaire le mode de diffusion qui lui convient, en fonction de ses besoins ou de ses préférences. Les enregistrements des présentations, diffusés en direct et en différé sur le site Web du cours, sont effectués chaque semaine, à l'horaire indiqué, avec des étudiants présents (en classe virtuelle ou, si possible, en classe physique) qui peuvent ainsi poser des questions et intervenir. Veuillez prendre note que, en fonction des directives de la santé publique, l'accès à l'enseignement en présentiel peut être limité au courant de la session. De plus, si des examens sous surveillance peuvent être réalisés, ceux-ci ont lieu en soirée ou la fin de semaine et peuvent donc se dérouler à un autre moment que la plage prévue pour les séances en classe et synchrones. Plus de détails seront fournis ultérieurement.

Plage horaire

Cours en classe		
jeudi	12h30 à 15h20	Du 31 août 2020 au 11 déc. 2020 (14h55)
Classe virtuelle synchrone		
jeudi	12h30 à 15h20	Du 31 août 2020 au 11 déc. 2020

Il se peut que l'horaire du cours ait été modifié depuis la dernière synchronisation avec Capsule. [Vérifier l'horaire dans Capsule](#)

Site de cours

<https://sitescours.monportail.ulaval.ca/ena/site/accueil?idSite=119861>

Coordonnées et disponibilités

Jocelyn Maclure

Enseignant

jocelyn.maclure@fp.ulaval.ca

Soutien technique

CSTIP - Centre de services en TI et pédagogie

Pavillon Charles-De Koninck, local 0248

aide@cstip.ulaval.ca

Téléphone : 418-656-2131 poste 405222

Informations techniques importantes

[Accéder à l'horaire du CSTIP](#) ↗

[Liste des navigateurs supportés](#) ↗

[Avis important concernant les appareils mobiles](#) ↗ (iOS, Android, Chrome OS)

[TeamViewer](#) ↗ est l'application utilisée par le CSTIP pour faire du soutien à distance. En exécutant cette application, vous permettez à un agent de soutien informatique de se connecter à votre poste pour vous aider.

Plan de cours non officiel

29 juil. 2020 (14h55)

Sommaire

Description du cours	4
Introduction	4
But du cours	4
Objectifs	4
Contenu	5
Formule pédagogique	5
Mode et critères d'évaluation	5
Contenu et activités	6
Évaluations et résultats	6
Modes et critères d'évaluation	6
Informations détaillées sur les évaluations sommatives	7
8 brefs commentaires critiques des lectures obligatoires (une page ou moins)	7
Présentation orale en classe	7
Plan détaillé du travail de recherche	7
Travail de recherche	8
Participation en classe	8
Échelle de notation	8
Gestion des délais	8
Politique sur le plagiat	8
Politique du français	9
Appréciation de l'enseignement	9
Prévenir et contrer le harcèlement à l'Université Laval	9
Évaluation de l'enseignement	9
Matériel didactique	9
Spécifications technologiques	9
Bibliographie	10
Bibliographie	10

Description du cours

Introduction

L'intelligence artificielle (IA) est présentement la technoscience qui suscite le plus d'intérêt à travers le monde. Après quelques décennies d'attentes déçues et de stagnation relative, ce qui a été appelé l'« hiver » de l'IA semble maintenant terminé. Des ordinateurs sont parvenus à vaincre les meilleurs joueurs humains à des jeux comme les échecs, Jeopardy et le jeu de go. Des algorithmes d'apprentissage automatique (*machine learning*) sont plus efficaces que des médecins spécialistes pour poser certains diagnostics. Des véhicules complètement autonomes pourraient rouler sur nos routes dans la prochaine décennie. Plusieurs avancent que la capacité croissante des systèmes d'IA à percevoir correctement le monde extérieur, à traiter les langues naturelles et à établir des relations significatives entre des données nombreuses et variées (*Big Data*) fera en sorte que des tâches exécutées par des humains seront de plus en plus confiées à des machines, ce qui pourrait engendrer un « chômage technologique » important, ainsi qu'une remise en question de la place du travail rémunéré dans la vie humaine.

Les progrès récents en IA ont relancé les spéculations sur l'émergence d'IA « fortes » et « générales » dotées d'une intelligence supérieure à celle des êtres humains. Certaines des meilleures fictions actuelles—*Her*, *Ex Machina*, *Black Mirror*, *Westworld*, le roman d'Ian McEwan *Machines Like Me*—explorent à nouveaux frais le thème de la relation entre l'être humain et des IA capables de réussir aisément le « jeu de l'imitation » (ou « test de Turing »). Des philosophes, chercheurs et personnalités publiques influentes comme Nick Bostrom, David Chalmers, Max Tegmark, Ray Kurzweil, Stephan Hawking, Stuart Russell, Bill Gates et Elon Musk soutiennent (avec plus ou moins de nuance selon les cas) que l'émergence de « superintelligences » artificielles est possible et que ces IA pourraient poser un « risque existentiel » eu égard à la survie même du genre humain.

But du cours

Les discours inflationnistes ne manquent dans l'univers de l'IA : « Superintelligence », « singularité », « risque existentiel », création d'un « cerveau complet artificiel », émergence d'une conscience « sans substrat biologique », statut moral et droits des robots, « fin du travail », « Quatrième Révolution industrielle », « obsolescence de l'être humain ». Sans faire preuve de dogmatisme, ces perspectives seront passées au crible d'une approche analytique, critique et déflationniste dans le cadre du séminaire.

29 juil. 2020 (14h55)

Cela étant dit, le sain scepticisme au sujet des perspectives inflationnistes portées par le battage médiatique actuel (*hype*) ne doit pas nous faire perdre de vue les enjeux philosophiques, éthiques et politiques complexes soulevés par les avancées actuelles de l'IA. Nous aborderons des questions comme :

- Une IA pourrait-elle être consciente? Une intelligence générale et multidimensionnelle comme celle démontrée par les humains peut-elle être créée artificiellement? Le corps et l'enchaînement dans un « monde vécu » sont-ils, comme le veut la tradition phénoménologique, nécessaire à la cognition véritable? Quelles sont les forces et limites de l'approche inductive, probabiliste et corrélationniste du nouveau paradigme scientifique en IA (apprentissage automatique, réseaux de neurones artificiels)?
- Quel devrait être le statut moral et juridique des agents artificiels? Devrait-on leur accorder des droits? Doivent-ils être tenus responsables des conséquences de leurs erreurs ou défaillances (pensons à une information erronée donnée par un assistant vocal ou à un accident causé par un véhicule autonome)? Les processus décisionnels des algorithmes d'apprentissage profond sont, contrairement aux approches classiques en IA basées sur la logique formelle, opaques. Doit-on exiger qu'ils puissent expliquer leurs décisions et justifier leurs jugements?
- Quelles sont les conséquences de la gouvernance algorithmique de nos vies individuelles et collectives? Des algorithmes d'aide à la décision sont de plus en plus utilisés pour déterminer qui sera reçu en entrevue d'embauche, admis dans un programme d'étude, admissible à une libération conditionnelle, etc. Les machines peuvent-elles prendre des décisions discriminatoires? Pourquoi protège-t-on la vie privée et comment l'IA la menace-t-elle? Les algorithmes utilisés par les réseaux sociaux contribuent-ils à la création de « bulles informationnelles », à la détérioration de la qualité épistémique du débat public et à la polarisation sociale? L'automatisation engendrée par l'IA dans les milieux de travail fournit-elle un argument supplémentaire aux défenseurs de l'idée d'un revenu minimum garanti inconditionnel?

Objectifs

Objectifs de connaissance et de compréhension

1. Comprendre les théories et thèses des auteur-e-s;
2. Comprendre ce qu'est l'IA, les raisons de sa renaissance, ainsi que les forces et limites du nouveau paradigme dominant;
3. Cerner les enjeux philosophiques, éthiques et politiques soulevés par l'intégration de systèmes d'IA dans les différentes sphères de la vie humaine.

Objectifs d'habiletés intellectuelles

1. Être capable de mettre en relation et de faire dialoguer des théories contradictoires ou complémentaires;
2. Être capable de distinguer et de mettre en relation des problématiques qui relèvent de différents champs de la philosophie (philosophie de l'esprit, sciences cognitives, éthique, philosophie politique);
3. Développer la capacité de réfléchir de façon normative à tous les niveaux de généralité (de l'éthique normative fondamentale à l'éthique normative appliquée);
4. Développer sa capacité d'exprimer sa pensée de façon claire, de débattre et participer à la recherche conjointe de jugements rigoureux sur les questions abordées dans le séminaire;
5. Permettre à l'étudiant de clarifier et nuancer ses positions personnelles face aux diverses questions liées à l'IA et, ce faisant, de mieux comprendre l'esprit humain et le monde dans lequel il vit.

Contenu

La liste des lectures hebdomadaires sera remise à la première séance du séminaire.

Formule pédagogique

La formule retenue sera celle du *séminaire de recherche*. Le but de cette formule est de préparer l'étudiant.e à la recherche, à la rédaction et à la présentation orale des idées. L'étudiant.e aura à présenter en classe une synthèse des lectures obligatoires, ainsi qu'à réagir aux exposés de ses collègues. Le professeur fera de brèves introductions en début de séance et interviendra ponctuellement afin d'expliquer certaines thèses, préciser des faits, présenter des arguments et recadrer la discussion. L'objectif ultime est qu'étudiant.e.s et professeur enrichissent leur compréhension des théories et des phénomènes étudiés par la médiation des lectures et des discussions hebdomadaires. La lecture attentive des textes est une condition essentielle au bon déroulement du séminaire. Les étudiants devront déposer de courts commentaires sur les lectures obligatoires sur le forum ENA.

Mode et critères d'évaluation

- 1) 8 brefs commentaires critiques des lectures obligatoires (une page ou moins). Le commentaire critique doit cerner, de façon concise, le ou les principaux problèmes soulevés par la thèse de l'auteur étudié et poser quelques questions sur lesquelles nous pourrions revenir dans la période de discussion. *Les commentaires doivent être déposés sur le forum le jour précédant le séminaire.* 20% de la note finale.
- 2) Une présentation de 20 minutes sur Zoom. Les présentations portent sur les lectures obligatoires et ont pour but de lancer les discussions de groupe sur les textes à l'étude. L'étudiant.e doit présenter l'approche de l'auteur étudié, son but, expliquer (lorsque pertinent) les concepts utilisés, faire une synthèse de l'argumentaire et soulever des pistes de réflexion critique. 20% de la note finale.

3) Un plan détaillé du travail de recherche, incluant une présentation de la problématique et du but du texte, la structure provisoire et une courte bibliographie (livres, articles publiés dans des revues universitaires et chapitres de livres). 10% de la note finale. Date de remise : 4 décembre.

4) Travail de recherche de 10 pages (double interligne) sur un thème pertinent dans le cadre du séminaire. Critères de correction : compréhension de la problématique, qualité de la recherche, clarté et rigueur de l'argumentation et qualité de la langue. 40% de la note finale. Date de remise : 18 décembre.

5) Participation aux discussions. 10% de la note finale.

Contenu et activités

Le tableau ci-dessous présente les semaines d'activités prévues dans le cadre du cours.

Titre	Date
Séance 1 : Présentation et introduction au séminaire	3 sept. 2020
Séance 2 : La naissance de l'IA : l'approche symbolique	10 sept. 2020
Séance 3 : Qu'est-ce que la conscience?	17 sept. 2020
Séance 4 : L'IA forte et faible : la chambre chinoise	24 sept. 2020
Séance 5 : L'approche connexionniste et l'apprentissage profond	1 oct. 2020
Séance 6 : Superintelligence, singularité et risque existentiel	8 oct. 2020
Séance 7 : Le statut moral des robots: personnes morales ou esclaves?	15 oct. 2020
Séance 8 : L'éducation morale des machines	22 oct. 2020
Séance 9 : Semaine de lecture	29 oct. 2020
Séance 10 : L'IA comme boîte noire : le problème de l'explicabilité	5 nov. 2020
Séance 11 : Des IA discriminatoires?	12 nov. 2020
Séance 12 : IA et vie privée	19 nov. 2020
Séance 13 : Automatisation et justice sociale : l'IA va-t-elle exacerber les injustices?	26 nov. 2020
Séance 14 : À déterminer	3 déc. 2020

Note : Veuillez vous référer à la section *Contenu et activités* de votre site de cours pour de plus amples détails.

Évaluations et résultats

Modes et critères d'évaluation

Sommatives

Titre	Date	Mode de travail	Pondération
8 brefs commentaires critiques des lectures obligatoires (une page ou moins)	À déterminer	Individuel	20 %
Présentation orale en classe	À déterminer	Individuel	20 %
Plan détaillé du travail de recherche	À déterminer	Individuel	10 %
Travail de recherche	À déterminer	Individuel	40 %
Participation en classe	À déterminer	Individuel	10 %

Formatives

Titre	Date	Mode de travail
Cette liste ne contient aucun élément.		

Informations détaillées sur les évaluations sommatives

8 brefs commentaires critiques des lectures obligatoires (une page ou moins)

Date de remise : À déterminer

Mode de travail : Individuel

Pondération : 20 %

Remise de l'évaluation : Forums du portail du cours

Directives de l'évaluation :

Le commentaire critique doit cerner, de façon concise, le ou les principaux problèmes soulevés par la thèse de l'auteur étudié et poser quelques questions sur lesquelles nous pourrions revenir dans la période de discussion. Les commentaires doivent être déposés sur le forum du séminaire le jour précédant le séminaire.

Présentation orale en classe

Date de remise : À déterminer

Mode de travail : Individuel

Pondération : 20 %

Remise de l'évaluation : En classe

Directives de l'évaluation :

Une présentation de 20-25 minutes en classe. Les présentations portent sur les lectures obligatoires et ont pour but de lancer les discussions de groupe sur les textes à l'étude. L'étudiant-e doit présenter l'approche de l'auteur étudié, son but, expliquer (lorsque pertinent) les concepts utilisés, faire une synthèse de l'argumentaire et soulever des pistes de réflexion critique.

Plan détaillé du travail de recherche

Date de remise : À déterminer
Mode de travail : Individuel
Pondération : 10 %
Remise de l'évaluation : jocelyn.maclure@fp.ulaval.ca
Directives de l'évaluation : Un plan détaillé du travail de recherche, incluant une présentation de la problématique et du but du texte, la structure provisoire et une courte bibliographie (livres, articles publiés dans des revues universitaires et chapitres de livres).

Travail de recherche

Date de remise : À déterminer
Mode de travail : Individuel
Pondération : 40 %
Remise de l'évaluation : jocelyn.maclure@fp.ulaval.ca
Directives de l'évaluation : Travail de recherche de 10 pages (double interligne) sur un thème pertinent dans le cadre du séminaire. Critères de correction : compréhension de la problématique, qualité de la recherche, clarté et rigueur de l'argumentation et qualité de la langue.

Participation en classe

Date de remise : À déterminer
Mode de travail : Individuel
Pondération : 10 %
Remise de l'évaluation : En classe

Plan de cours non officiel

29 juil. 2020 (14h55)

Échelle de notation

Cote	% minimum	% maximum
A+	95	100
A	90	94,99
A-	85	89,99
B+	80	84,99
B	76	79,99
B-	72	75,99

Cote	% minimum	% maximum
C+	68	71,99
C	64	67,99
E	0	63,99

Gestion des délais

Cette sous-section n'a pas encore été complétée par votre enseignant.

Politique sur le plagiat

L'étudiant inscrit à ce cours est assujéti au *Règlement disciplinaire à l'intention des étudiants de l'Université Laval* qui interdit toute forme de plagiat. Tout étudiant qui commet une infraction à ce règlement dans le cadre du présent cours est passible de sanctions. Pour plus d'informations concernant ce qui est interdit et ce qui est considéré comme une forme de plagiat, consultez le document à l'adresse suivante :

<http://ulaval.ca/reglement-disciplinaire> 

Politique du français

Le Conseil de la Faculté de philosophie a adopté une politique du français qui est en vigueur depuis la session d'hiver 2009. Attendu que la qualité de l'expression de la pensée est une composante essentielle de l'excellence en philosophie, le Comité des programmes de premier cycle en philosophie recommande au minimum que la qualité du français soit prise explicitement en considération dans l'évaluation de chaque travail ou examen, qu'il soit fait en classe ou à la maison. L'entièreté de la politique du français en vigueur à la Faculté de philosophie peut être consultée à l'adresse suivante :

<http://www.fp.ulaval.ca/fileadmin/philo/documents/formation/information-documentation/guides-politiques/politique-francais.pdf>

Appréciation de l'enseignement

En conformité avec la Politique de valorisation de l'enseignement et Dispositions relatives à l'évaluation de l'enseignement à l'Université Laval, il est possible que ce cours fasse l'objet d'une appréciation par les étudiants. Votre opinion est très importante. Les commentaires constructifs, positifs comme négatifs, aident à distinguer les aspects à améliorer et ceux à consolider. Vous ne profiterez peut-être pas vous-mêmes des bonifications apportées, mais vous contribuerez à l'amélioration continue de ce cours dont bénéficieront les futures cohortes, tout comme les précédentes l'ont fait pour vous. Nous comptons donc grandement sur votre collaboration.

Prévenir et contrer le harcèlement à l'Université Laval

Ainsi qu'en dispose le [Règlement pour prévenir et contrer le harcèlement à l'Université Laval](#) dans son préambule, « l'Université reconnaît sa responsabilité de créer et de maintenir, pour tous les membres de la communauté et pour les tiers, un environnement propice à la réalisation de ses missions d'enseignement et de recherche, de même qu'un milieu de travail, d'étude et de prestation de services harmonieux, exempt de harcèlement, où chacun a droit au respect, à la dignité et à la protection de son intégrité physique et psychologique. »

Au sein de l'Université, n'hésitez pas à consulter les ressources développées par le [Centre de prévention et d'intervention en matière de harcèlement](#), ainsi qu'à avoir recours au soutien et aux conseils offerts.

Évaluation de l'enseignement

« Il est possible que le cours que vous suivez soit évalué. Si tel est le cas, vous recevrez une invitation à remplir un questionnaire d'appréciation en ligne [ou transmis par la poste selon le cas]. Votre opinion est très importante car elle permettra d'améliorer la qualité de ce cours. Nous comptons donc grandement sur votre collaboration. »

Matériel didactique

Spécifications technologiques

Afin d'accéder sans difficulté au contenu de votre cours sur monPortail, vous devez vous référer aux informations ci-dessous.

Liste des navigateurs supportés :

- [Google Chrome](#) 

- [Mozilla Firefox](#) ↗
- Safari
- Internet Explorer 11 et plus
- [Opera](#) ↗

Appareils mobiles

Il est important de préciser que l'utilisation des périphériques mobiles n'est pas officiellement supportée pour les sites de cours. La consultation des contenus est fonctionnelle pour la plupart des appareils sous Android et iOS, mais il existe certaines limites :

- les appareils mobiles sous Windows RT, BlackBerry et les iPhone 4 et 4S ne sont pas supportés;
- le contenu Flash n'est pas supporté, ce qui empêche la consultation des blocs de contenu audio-vidéo, de contenu Flash et de certaines présentations multimédias;
- la fonctionnalité de glisser-déposer n'est pas supportée, ce qui empêche la réalisation de certains questionnaires.

Team Viewer

TeamViewer est l'application utilisée par le CSTIP pour faire du soutien à distance. En exécutant cette application, vous permettrez à un agent de soutien informatique de se connecter à votre poste pour vous aider.

- [Accédez à TeamViewer pour Windows](#) ↗
- [Accédez à TeamViewer pour Mac](#) ↗

Bibliographie

Bibliographie

Bibliographie complémentaire.

Boden, M. A. (2016). *AI. Its Nature and Future*. New York : Oxford University Press.

Canguilhem, G. (1993). « Machine et organisme », dans *La Connaissance de la vie*. Paris : Vrin, pp. 129-164.

Crawford, K., Miltner, K. et Gray, M. (2014). « Critiquing Big Data: Politics, Ethics, Epistemology », *International Journal of Communications*, 8, pp. 1663-1672.

en ligne : <http://ijoc.org/index.php/ijoc/article/view/2167/1164> ↗

Crawford, K. et Calo, R. (2016, 13 octobre). « There is a blind spot in AI research ». *Nature*.

en ligne : <http://www.nature.com/news/there-is-a-blind-spot-in-ai-research-1.20805> ↗

Dreyfus, H. (1992). *Intelligence artificielle. Mythes et limites*. Paris : Flammarion.

Floridi, L. (2016, 28 décembre). « Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions ». *Phil. Trans. R. Soc. A* 374: 20160112.

en ligne : <https://doi.org/10.1098/rsta.2016.0112>

Floridi, L. (2017). « Robots, Jobs, Taxes, and Responsibilities ». *Springer Science+Business Media Dordrecht*.

en ligne : https://www.researchgate.net/publication/314953596_Robots_Jobs_Taxes_and_Responsibilities ↗

Ford, M. (dir.). (2018). *Architects of Intelligence: The truth about AI from the people building it*. Birmingham : Packt Publishing.

Fry, H. (2018). *Hello World: Being Human in the Age of Algorithms*. New York : W. W. Norton & Company.

Kaplan, J. (2016). *Artificial Intelligence. What Everyone Needs to Know*. New York : Oxford University Press.

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G. & Yu, H. (2017). « Accountable Algorithms ». *University of Pennsylvania Law Review*, 165(3), pp. 633-705.

en ligne : <https://pdfs.semanticscholar.org/c72a/19f4a7edf4c688549f5572c10796763613c5.pdf> ↗

O'Neill, C. (2017). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York : Broadway Books.

en ligne : <http://www.raison-publique.fr/article776.html> ↗

Stilgoe, J. (2017, 21 novembre). « Machine learning, social learning and the governance of self-driving cars ». *Social Studies Science*, 48(1), pp. 25-56.

en ligne : <https://doi.org/10.1177/0306312717741687> ↗

Sunstein, C. (2016). « Délibération, nouvelles technologies et extrémisme », traduit de l'anglais par Solange Chavel, *Raison publique*.

Vermeys, N. (2018, octobre). « La responsabilité civile du fait des agents autonomes », *Les Cahiers de propriété intellectuelle*, 30(3), pp. 851-880.

en ligne : <https://www.lescpi.ca/s/3747> ↗

Warner, R. & Sloan, R. H. (2017, 17 juillet). « The Ethics of the Algorithm: Autonomous Systems and the Wrapper of Human Control ». 48 *CUMBERLAND LAW REVIEW* 101.

en ligne : <http://dx.doi.org/10.2139/ssrn.3004016> ↗

Zubov, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York : PublicAffairs.

Plan de cours non officiel

29 juil. 2020 (14h55)