

English Follow

LUNDI 19 octobre 2026

<u>Lyse Langlois - Gouverner l'IA : Normes, savoir et pouvoir</u>

Lyse Langlois est directrice générale de l'Observatoire international sur les impacts sociétaux de l'IA et du numérique (Obvia) depuis 2018 et professeure titulaire au Département des relations industrielles de l'Université Laval. Elle a dirigé pendant huit ans l'Institut d'éthique appliquée (IDÉA) de cette même université et est chercheuse régulière au Centre de recherche interuniversitaire sur la mondialisation du travail (CRIMT) ainsi qu'à l'Institut intelligence et données (IID) de l'Université Laval. En 2025-2026, elle a été professeure invitée à la Korea University avec qui est poursuit plusieurs projets de collaboration. Ses recherches ont donné lieu à de nombreuses publications scientifiques, ouvrages, articles et rapports sur l'éthique et les transformations du travail liées à la technologie. Sa plus récente contribution porte sur la géopolitique de la science et la diplomatie scientifique, dans un chapitre publié chez Palgrave Macmillan dans Polycentric Science Diplomacy in an Age of Disruption, sous la direction de J. Lüdert et M. Wählisch.

Résumé : La gouvernance de l'intelligence artificielle s'appuie sur un ensemble croissant de principes éthiques, de normes techniques et de cadres réglementaires. S'y ajoutent des instruments conçus pour traduire ces normes en pratiques et en évaluer la mise en œuvre, comme la méthodologie d'évaluation de l'état de préparation (RAM) et l'évaluation de l'impact éthique (EIA) de l'UNESCO, le NIST AI RMF ou la norme ISO/IEC 42001.

Gouverner l'IA ne se limite pas à fixer ce qui est acceptable. Il faut aussi décider de ce qui mérite d'être connu, mesuré et rendu visible, et donc de ce qui restera hors champ. Cette conférence examine les rapports entre normes, savoir et pouvoir. Trois questions guident la

réflexion. Par quels processus certaines valeurs deviennent-elles des normes de gouvernance, et quels savoirs leur donnent leur légitimité ? Que laissent ces normes dans l'ombre, et enfin, à quelles conditions une gouvernance peut-elle soumettre ses propres normes à la critique ? Je défendrai l'idée qu'une gouvernance crédible de l'IA doit accepter d'être contestée. Cela suppose des mécanismes d'évaluation indépendants et pluralistes.

Sarah Brown - Réinterprétations relationnelles de l'IA

La Dre **Sarah M. Brown** est professeure adjointe en informatique à l'Université du Rhode Island. Ses recherches portent sur la manière de rendre les systèmes d'intelligence artificielle meilleurs pour toutes et tous, en étudiant les interactions entre la technologie, les personnes qui la développent et ses impacts sociaux. Ses travaux placent les personnes au centre à toutes les échelles : en développant des interventions qui soutiennent les scientifiques des données dans le développement de systèmes utilisés dans le monde réel, et en développant directement des techniques qui intègrent des valeurs sociales dans les systèmes d'intelligence artificielle. Elle est titulaire d'un baccalauréat, d'une maîtrise et d'un doctorat en génie électrique de la Northeastern University. La Dre Brown est récipiendaire de la Graduate Research Fellowship et du prix CAREER de la National Science Foundation (NSF).

Résumé : De nombreuses valeurs généralement attribuées à l'IA trouvent leur origine dans le langage, et certaines peuvent être façonnées au moyen de stratégies d'alignement post-entraînement. Dans d'autres cas, certaines valeurs sont rattachées à des contraintes techniques. Cependant, souvent, ces valeurs ne sont pas fondamentales à l'objet mathématique ou computationnel lui-même, mais découlent d'une philosophie des sciences dominante. La valeur est imposée dans l'interprétation de l'objet, puis lui reste associée lors de la mise en œuvre de technologies réelles. Ces valeurs sont souvent présentées comme inévitables dans la poursuite du développement technologique, en partie parce que la philosophie des sciences n'occupe pas une place centrale dans la formation dans les domaines computationnels. En réalité, presque toutes les idées deviennent contestables lorsqu'elles sont examinées en profondeur. Dans cette présentation, je mettrai en évidence quelques exemples de ce phénomène. Je proposerai ensuite un fondement philosophique alternatif : le réalisme relationnel. Je justifierai son utilité et sa pertinence pour l'IA afin d'en présenter les principes fondamentaux, puis je conclurai par des exemples de manières de réinterpréter, depuis cette perspective, les contraintes techniques liées à l'IA, permettant ainsi d'utiliser l'IA de façons qui favorisent des valeurs souvent sous-représentées dans l'IA.

Marc-Antoine Dilhac - Regard sur l'évolution des cadres de gouvernance

Marc-Antoine Dilhac est professeur d'éthique et de philosophie politique à l'Université de Montréal, où ses travaux portent sur l'éthique et la gouvernance de l'intelligence artificielle. En 2017, il a lancé et dirigé le processus de co-construction de la Déclaration de Montréal pour un

développement responsable de l'intelligence artificielle. Il est directeur de la gouvernance et de la collaboration internationale à l'OBVIA et membre du Mila , où il a animé une chaire Canada-CIFAR en éthique de l'IA. Il intervient en tant qu'expert de l'UNESCO au sein du réseau « Experts Without Borders » sur l'éthique de l'IA et a dirigé la mise en œuvre de la méthodologie d'évaluation de l'état de préparation (RAM) de l'UNESCO en Ouzbékistan et au Québec. Il a auparavant siégé au Conseil consultatif sur l'IA du gouvernement du Canada et a été l'expert chargé du processus de consultation du Conseil de l'Europe sur la Convention-cadre relative à l'intelligence artificielle, aux droits de l'homme, à la démocratie et à l'État de droit.

Résumé : Dès les premières déclarations éthiques sur l'intelligence artificielle, leur mise en œuvre concrète, tant par des procédures d'entreprise que des politiques publiques, a été identifiée comme un impératif prioritaire. Valeurs et principes éthiques devaient être « opérationnalisés » pour orienter effectivement la conception, le déploiement et l'encadrement des systèmes d'IA. Cependant cet impératif s'est avéré plus difficile à satisfaire que prévu. Près de dix ans après le lancement de l'initiative de la Déclaration de Montréal pour un développement responsable de l'IA, cette conférence propose un bilan de l'évolution de la gouvernance de l'IA. Comment les cadres de gouvernance se sont-ils transformés ? Qu'a-t-on appris de leur application ? Que reste-t-il à faire pour s'assurer que nous avons toujours la main sur le développement de l'IA ? Pour répondre à ces questions, on examinera des cadres non contraignants influents comme la Recommandation de l'UNESCO sur l'éthique de l'IA ou le Code de conduite du Hiroshima AI Process (G7/OCDE) qui est accompagné d'une procédure d'autodéclaration pour les entreprises, ainsi que des cadres contraignants comme la Convention-cadre du Conseil de l'Europe ou le Règlement européen sur l'IA. On soutiendra que les valeurs et les principes ne sont effectivement mis en œuvre que lorsqu'on dote les cadres de gouvernance de méthodes de preuve, qui permettent de démontrer, et pas seulement de déclarer, que les normes éthiques sont respectés.

Alison Gillwald – Remédier à la répartition inégale des préjudices et des opportunités liés à l'IA en donnant la priorité aux droits de deuxième et de troisième génération et à la valeur de l'« ubuntu »

Alison Gillwald (PhD) est la directrice fondatrice de Research ICT Africa (RIA) et professeure associée à la Nelson Mandela School of Public Governance de l'Université du Cap, où elle a dirigé un programme doctoral panafricain sur les politiques numériques.

Alors que RIA était partenaire de connaissances de la présidence sud-africaine du G20, elle a agi à titre de conseillère technique auprès du Groupe de travail sur l'économie numérique et de la Task Force de haut niveau sur l'intelligence artificielle, la gouvernance des données et l'innovation pour le développement durable. Elle a également coprésidé la Task Force du T20 sur la transformation numérique. Elle est membre du groupe de travail T7 de la présidence française du G20 sur l'IA et les puissances moyennes et siège au sein de l'organe consultatif académique de l'Union internationale des télécommunications sur les technologies avancées. Elle a été active au sein du Partenariat mondial sur l'intelligence artificielle (GPAI), codirigeant les projets sur la transparence algorithmique dans le secteur public et sur la justice des données.

Alison a dirigé l'assistance technique de RIA auprès de la Commission de l'Union africaine, en facilitant les travaux de l'équipe internationale et en rédigeant le Cadre de politique des données de l'UA et son plan de mise en œuvre. Actuellement, elle est chercheuse principale du projet de la Banque africaine de développement sur les infrastructures publiques numériques pour l'Afrique, qui examine la réduction des risques liés aux investissements bancaires à travers le prisme de l'économie politique.

Ses publications les plus récentes comprennent le document de cadrage pour le Dialogue du Club de Madrid sur la gouvernance des biens publics mondiaux dans un monde fragmenté : biens publics numériques et infrastructures, ainsi que Gillwald, A., et Fuller, J. (2026), Transformation numérique et limites du G20 : enseignements tirés des présidences du G20 des économies émergentes.

MARDI 20 octobre 2026

Alexandra Bensamoun – La transparence, valeur ou instrument ? L'AI Act face aux enjeux culturels

Alexandra Bensamoun est professeure de droit à l'Université Paris-Saclay, spécialiste de propriété intellectuelle et de régulation du numérique, directrice du Master 2/LLM Propriété intellectuelle fondamentale et technologies numériques, en codiplômation avec l'Université Laval (Québec) et l'Université Autonome de Madrid (Espagne). Experte pour des autorités internationales (notamment l'UNESCO) et nationales, personnalité qualifiée au CSPLA (ministère de la Culture français), elle a siégé en France à la Commission interministérielle de l'intelligence artificielle, qui a rendu un rapport sur la stratégie française au Président de la République, IA : notre ambition pour la France (mars 2024). Ses rapports au ministère de la Culture portent notamment sur la rémunération des contenus culturels utilisés par les IA (à l'origine de la proposition de loi n° 220 votée à l'unanimité au Sénat en avril 2026) ou encore sur le statut des productions de l'IA. Elle est l'auteur de nombreuses recherches individuelles ou collectives, notamment un Traité collectif sur le Droit de l'intelligence artificielle (LGDJ/Lextenso, 3e éd., 2026).

Résumé : La transparence est souvent présentée comme l'une des valeurs structurantes de la régulation européenne de l'IA. Dans les secteurs culturels, elle prend une tournure particulière, puisqu'elle conditionne à la fois l'effectivité des droits des créateurs et la confiance du public dans les contenus qui circulent. A partir d'une actualité transversale, celle de la signature et de l'authenticité des œuvres, les deux temps de la transparence dans le règlement (UE) 2024/1689 seront évalués. En amont, l'articulation avec l'exception de fouille de textes et de données et son mécanisme d'opposition, le résumé suffisamment détaillé des contenus d'entraînement et le modèle publié par la Commission, ainsi que la présomption d'utilisation proposée en droit français, qui révèle les limites de l'information. En aval, le marquage des contenus générés et l'information du public, avec le régime aménagé prévu pour les œuvres artistiques. Il s'agira aussi de mesurer si cette transparence, pensée comme une obligation d'information, peut devenir un véritable outil de préservation des valeurs culturelles.

Maurice Jones – ArtIA : Vers un nouveau commun numérique pour l'IA créative

Dr Maurice Jones est chercheur, artiste et stratège culturel basé à Tiohtià/Montréal. Il a récemment soutenu sa thèse de doctorat interdisciplinaire en sciences humaines à l'Université Concordia. Ses domaines de recherche portent sur les festivals et les infrastructures de la culture numérique, l'IA culturelle et les communs, la gouvernance transnationale de l'IA, ainsi que la recherche-création numérique.

Résumé : S'appuyant sur « ArtIA », une consultation de 18 mois menée auprès de plus de 200 artistes, institutions culturelles, technologues et travailleur·euse·s du secteur créatif à travers le Canada, cette présentation explore les possibilités d'approches fondées sur les communs pour l'IA créative. Dr. Maurice Jones examine comment la propriété collective, les infrastructures partagées, les littératies critiques en IA et la réappropriation créative peuvent offrir des alternatives aux modèles dominants de développement et d'adoption de l'IA pilotés par l'industrie, et pose la question de ce que signifierait la construction d'écosystèmes d'IA créative ancrés dans les besoins culturels et communautaires plutôt que dans les logiques extractives des plateformes.

Antoine Henry - Opérationnaliser des valeurs dans les systèmes de recommandations culturels : la création d'un Fairness Score

Antoine Henry est maître de conférences en sciences de l'information et de la communication et membre de l'axe 4 de GERIICO dont la thématique est la circulation de l'information et l'organisation des connaissances.

Ses travaux de recherche portent, principalement, sur les questions relatives au numérique (IA, communs numériques, éthique des algorithmes), à la transformation des organisations et à l'intelligence collective.

Il est aussi :

- membre du conseil d'administration du chapitre français de l'association savante ISKO,

- membre du groupe de recherche du Centre Internet et Société (UPR 2000 du CNRS) au sein duquel il co-anime le groupe de travail IA, art et créativité.

Résumé : L'équité algorithmique représente un défi majeur dans le développement des systèmes de recommandation, particulièrement dans le secteur musical. Cette intervention propose une approche novatrice aux cadres actuels d'équité algorithmique en introduisant un nouveau modèle pour évaluer l'équité. Nous développons une méthode permettant de synthétiser neuf dimensions complexes pour influencer la boucle de rétroaction des algorithmes. À travers cette proposition du Fairness Score, nous visons à opérationnaliser l'équité et à offrir aux utilisateurs un contrôle significatif sur leur expérience et leur pratique culturelle.

Christophe Abrassard – La sobriété numérique, un nouveau paradigme pour l'IA dans un monde incertain.

Christophe Abrassart est professeur agrégé à l'École de design de la Faculté de l'aménagement de l'Université de Montréal. Spécialiste d'éco-design et de prospective stratégique, il est co-directeur du Lab Ville Prospective de la Faculté de l'aménagement et co-responsable de l'Axe de recherche sur la Sobriété numérique et la transition socioécologique de l'OBVIA. Il conçoit et anime depuis 15 ans des ateliers de codesign et de prospective participative sur la transition écologique et énergétique au Québec avec différentes organisations (ex. Ville de Montréal, Hydro-Québec). Il a notamment coordonné un exercice de prospective à l'OBVIA sur la sobriété numérique au Québec à l'horizon 2050.

Résumé : Le monde actuel est marqué par une profonde antinomie. D'un côté, les scientifiques de l'environnement sonnent l'alarme sur l'urgence climatique, et l'impératif de retrouver au plus vite un "espace de fonctionnement sécuritaire pour l'humanité" en deça des limites planétaires (Rockström, 2009). De l'autre, les multinationales du numérique accélèrent leurs investissements dans des centres de données toujours plus grands et gourmands en matériaux, eau et énergie. A tel point que les transitions écologiques et numériques, jadis présentées comme complémentaires, sont devenues contradictoires. Par exemple, entre l'électrification des transports, l'adaptation du bâti aux périodes de canicule et l'électrification des centres de données, tout ne pourra pas être réalisé sans tensions majeures dans le système énergétique au cours des prochaines décennies (Shift Project, 2025).

Face à cette contradiction, une première approche est celle de l'écoconception pour optimiser la performance énergétique des centres de données. Mais ce secteur est caractérisé par des effets rebonds continus. Une autre approche est celle de la sobriété systémique qui consiste à interroger de manière critique les usages de l'IA et à opérer un discernement entre les projets prioritaires et les autres (ex. entre des IA pour la santé et des IA pour le divertissement ou le marqueting expérienciel), ou à réfléchir à la juste place du numérique dans la société (ex. à l'école ou dans les organisations). Pour explorer les potentiels de cette approche et imaginer des innovations sociotechniques auxquelles elle pourrait conduire, cette interventiion s'appuiera sur quatre scénarios de prospective sur la sobriété numérique au Québec en 2050, issus d'un projet coordonné par l'intervenant à l'OBVIA.

Jocelyn Maclure - L'éthique et la gouvernance de l'IA: Pourquoi l'approche de l'alignement sur les valeurs est mal avisée

Jocelyn Maclure est professeur de philosophie et titulaire de la Chaire Jarislowsky sur la nature humaine et la technologie à l'Université McGill. Ses champs de recherche et d'enseignement sont l'éthique, la philosophie politique, la philosophie de l'IA et l'épistémologie sociale. Ses travaux récents ont été publiés dans des revues comme Minds & Machines, AI & Ethics, Digital Society et AI & Society. Il a récemment été professeur invité aux universités de Bonn et de

Grenoble, et a présidé la Commission de l'éthique en science et en technologie du Québec de 2017 à 2024.

Le « problème de l'alignement sur valeurs humaines » (PAV) constitue sans doute l'approche la plus influente chez les philosophes et les chercheurs en intelligence artificielle dans les champs de l'éthique et de la gouvernance de l'IA. Le PAV est centré sur la question de savoir comment concevoir des systèmes d'IA avancés dont les objectifs et les résultats sont alignés sur les valeurs humaines. S'il n'y a évidemment rien de problématique à vouloir ancrer le développement des IA de pointe dans des valeurs humaines partagées, cette approche laisse de côté les aspects les plus difficiles de l'éthique pratique et de l'élaboration des politiques publiques. Après avoir examiné les principaux problèmes auxquels se heurte le PAV (les problèmes du pluralisme, de la sous-détermination et de la motivation), je proposerai les grandes lignes d'une approche alternative que j'appelle l'« intégrité normative ».

MERCREDI 21 octobre 2026

Jean-Gabriel Ganascia - IA agentique et/ou garantie humaine?

Ingénieur et philosophe de formation initiale, **Jean-Gabriel Ganascia** s'est très tôt orienté vers l'informatique et l'intelligence artificielle. Il a soutenu en 1983, à l'université Paris-Saclay, une thèse de doctorat sur les systèmes experts puis en 1987, toujours à l'université Paris-Saclay, une thèse d'État sur l'apprentissage machine. Présentement, professeur émérite à la faculté des sciences de Sorbonne Université, membre honoraire de l'institut universitaire de France, EurAI fellow et chercheur au LIP6, il poursuit des travaux sur l'apprentissage machine, les humanités numériques, la détection d'Infox et l'éthique du numérique. Dans le passé, il a coordonné la recherche française en sciences cognitives en créant puis en animant le Groupement d'Intérêt Scientifique « Sciences de la cognition ». Il est aussi à l'origine du Labex OBVIL (Observatoire de la vie littéraire) qui coordonnait de 2011 à 2021 les efforts de son équipe au sein du LIP6 et des laboratoires de littérature de Sorbonne Université sur le versant littéraire des humanités numériques. Outre ses activités de recherche, il préside le comité d'éthique de France Travail, le comité d'éthique de Docaposte et le comité d'orientation du CHEC (Cycle des Hautes Études de la Culture). Il est aussi membre du COMEST (Commission mondiale d'éthique des connaissances scientifiques et des technologies de l'UNESCO) Enfin, il a présidé le comité d'éthique du CNRS (COMETS) de 2016 à 2021 et l'AFAS (Association Française pour l'Avancement des Sciences) de 2022 à 2026. Au cours de sa carrière, il a publié plus de 600 articles dans des actes de conférences, des livres et des revues scientifiques. Il est aussi l'auteur d'une douzaine d'ouvrages destinés au grand public et il tient depuis 2017 plusieurs chroniques éthiques régulières, l'une dans la revue La Recherche, l'autre dans la revue Sciences et avenir.

Résumé : Suite au boum de l'IA générative, nous entrons, depuis la fin 2025, dans l'âge de l'IA agentique. L'idée est ancienne: un agent capte de l'information, texte, image ou son, l'interprète, prend seul des décisions sur la foi de ses entrées, et agit (d'où le nom d'agent) en vue de réaliser les buts qu'on lui a fixés. Comment ne pas s'en inquiéter: l'IA agentique risque de nous mettre sous tutelle en nous court-circuitant. Dans l'éthique du soin, on a introduit la

notion de « garantie humaine » pour s'assurer qu'aucune décision ne se prenne sans qu'un personnel de santé ne l'ait préalablement approuvée. Pourquoi ne pas toujours procéder ainsi? Parce qu'on n'a pas le temps et que l'humain n'est pas une garantie! Errare humanum est: nous nous trompons parfois plus souvent que les automates. L'étude rétrospective d'accidents d'avions ou de centrales nucléaires le montre. Comment résoudre en amont les potentiels les conflits d'autorité entre humains et machines et décider de celui qui doit l'emporter? On choisit parfois, sur la base d'arguments rationnels, de laisser la main à l'agent artificiel. Mais, cette délégation de responsabilité ne vaut que pour un temps et doit rester révocable à tout moment. Du moins, l'espère-t-on… Ce sont toutes ces questions qu'il s'agira d'aborder ici afin de se demander ce qui, dans chaque situation, reste admissible et ce qu'il faut condamner.

Maroussia Lévesque - Les droits de la personne, boussole de la gouvernance de l'IA en pratique

Maroussia Lévesque a rejoint la Faculté de droit de l'Université Queen's en tant que professeure adjointe en droit et IA en 2026. Ses recherches portent sur la gouvernance de l'IA, en examinant la conception de la réglementation, les déséquilibres de pouvoir et la concentration du marché dans l'industrie de l'IA, ainsi que les rôles respectifs des acteurs publics et privés dans l'écosystème de gouvernance. Ses projets actuels analysent la manière dont les structures organisationnelles des laboratoires d'IA de pointe façonnent leur comportement dans la pratique, ainsi que les implications environnementales et les usages judiciaires de l'IA.

Elle est affiliée au Berkman Klein Center for Internet & Society, chercheuse principale au Centre pour l'innovation dans la gouvernance internationale (CIGI) et collaboratrice du projet Abundant Intelligences, dirigé par des Autochtones.

Avant de rejoindre l'Université Queen's, la professeure Lévesque a obtenu son doctorat en sciences juridiques (SJD) à la Faculté de droit de Harvard, où elle a enseigné la gouvernance comparée de l'IA (avec la professeure Martha Minow) et a été chercheuse au Weatherhead Center for International Affairs. Auparavant, elle a dirigé des travaux sur l'IA et les droits de la personne à Affaires mondiales Canada, a été auxiliaire juridique auprès du juge en chef de la Cour d'appel du Québec, a collaboré avec l'Algorithmic Accountability Lab d'Amnistie internationale et a contribué à la mise en place du Partenariat mondial sur l'intelligence artificielle. Elle est membre du Barreau du Québec.

Les travaux de recherche de la professeure Lévesque ont notamment été publiés dans le NYU Journal of Legislation & Public Policy, le University of Illinois Journal of Law, Technology & Policy et les actes de NeurIPS. Elle a écrit pour The Guardian, le Toronto Star, La Presse, The Walrus et Tech Policy Press.

Résumé : Les droits de la personne sont une avenue prometteuse pour guider la gouvernance de l'IA. Leur vocabulaire établi et légitimité internationale sont appuyés par une architecture institutionnelle bien rodée. Par ailleurs, les Principes directeurs des Nations Unies relatifs aux entreprises et aux droits de l'homme étendent ce vocabulaire aux acteurs privés, articulant une

responsabilité des entreprises de respecter les droits de la personne distincte de l'obligation contraignante des États de les protéger. Cela dit, demander sur quelles valeurs la gouvernance de l'IA devrait reposer est plus facile que de cerner ce qui rend ces valeurs utiles en pratique.

Cette présentation examinera le potentiel et les limites des droits de la personne comme guide de la gouvernance de l'IA. Au début de l'ascension des modèles d'apprentissage automatique en 2018, la Déclaration de Toronto a traduit l'exigence de diligence raisonnable des Principes directeurs en étapes concrètes, offrant des balises opérationnelles pour cerner les risques, les prévenir et les atténuer, suivre les réponses et offrir des mécanismes de reddition de comptes. Plus récemment, le Conseil de surveillance de Meta (Oversight Board) s'est décrit comme un mécanisme non judiciaire indépendant au sens des Principes directeurs, invoquant les droits de la personne dans son examen de la modération automatisée des contenus. Ces exemples démontrent que les droits de la personne peuvent être mis en œuvre de manière à faire progresser la gouvernance de l'IA.

Par contre, les droits de la personne comportent également des limites. Des engagements non contraignants peuvent conférer un dividende de légitimité inapproprié, les entreprises drapant leurs démarches dans le langage de normes malléables tout en esquivant leur mise en application. Les cadres volontaires permettent également de faire perdurer le statu quo, soit l'absence de réglementation contraignante, normalisant des pratiques problématiques tandis que les attentes sociales s'ajustent. Enfin, les engagements souples cèdent sous la pression concurrentielle, comme le suggèrent les récents reculs à l'égard d'engagements en matière de droits de la personne.

Du côté des développements en cours, les règles de protection des investisseurs, qui exigent une divulgation rigoureuse des risques, pourraient faire progresser la diligence raisonnable à mesure que les laboratoires d'IA de pointe entrent en bourse. Le projet de traité contraignant sur les entreprises et les droits de la personne obligerait par ailleurs les États à réglementer, plutôt que de demander aux entreprises de se restreindre elles-mêmes. En définitive, énoncer des valeurs importe moins que la mise en place d'incitatifs et d'obligations contraignantes qui leur donnent prise.

Richard Khoury - L'IA Générative et la guerre contre la diversité linguistique

Richard Khoury a obtenu son baccalauréat et sa maîtrise en Génie Électrique et Génie Informatique de l'Université Laval (Québec, QC) en 2002 et 2004 respectivement, et son doctorat en Génie Électrique et Génie Informatique de l'Université de Waterloo (Waterloo, ON) en 2007. De 2008 à 2016, il a été professeur adjoint, puis professeur agrégé, au département de Génie Logiciel à l'Université Lakehead. En 2016, il s'est joint à l'Université Laval à titre de professeur agrégé, avant d'être nommé professeur titulaire en 2026. De 2021 à 2023 il a également été président de l'Association pour l'intelligence artificielle au Canada. Ses intérêts de recherche sont en forage de données et en traitement du langage naturel, et incluent de plus la gestion des connaissances, l'apprentissage automatique, et l'intelligence artificielle.

Résumé : Les outils d'IA Génératives sont aujourd'hui disponibles dans plus d'une centaine de langues différentes. À première vue, ceci semble être une victoire pour la diversité linguistique : l'intelligence artificielle offre de rejoindre les utilisateurs à travers le monde dans leur langue maternelle. Mais sous cette image positive se cache un problème insidieux. L'intelligence artificielle ne connait et ne comprend que le dialecte dominant de ces langues, et impose son utilisation aux locuteurs. Cette colonisation linguistique par l'IA menace d'effacer la diversité linguistique mondiale et de supprimer des communautés culturelles minoritaires. Dans cette présentation nous étudierons ce problème, ses origines, ses symptômes, et ses solutions.

Colette Brin – Explicabilité et transparence en journalisme: quelques défis de mise en pratique de valeurs d'IA responsable

Colette Brin est professeure titulaire au Département d'information et de communication de l'Université Laval et directrice du Centre d'études sur les médias. Spécialiste des transformations du journalisme et des pratiques d'information, elle est coresponsable de l'axe Arts, médias et diversité culturelle de l'Observatoire international sur les impacts sociétaux de l'IA et du numérique (OBVIA). Ses travaux récents portent notamment sur l'intégration responsable de l'intelligence artificielle dans les médias, les valeurs journalistiques et les besoins des publics. Elle a publié Médias publics, IA responsable et besoins des utilisateurs (2025) et, avec Véronique Guèvremont, Enjeux de l'intelligence artificielle pour les milieux de la culture et des médias d'information au Québec (2026). Elle a également codirigé avec Véronique Guèvremont l'ouvrage Intelligence artificielle, culture et médias (Presses de l'Université Laval, 2024).

Résumé : La transparence et l'explicabilité figurent parmi les principes les plus fréquemment associés à une utilisation responsable de l'intelligence artificielle. Leur mise en œuvre dans les pratiques journalistiques soulève toutefois des difficultés particulières. Comment rendre visible et compréhensible l'intervention de systèmes d'IA dans la production de l'information sans surcharger les publics d'indications techniques? Que signifie expliquer l'usage de l'IA lorsque les journalistes et les organisations médiatiques disposent eux-mêmes d'une connaissance limitée du fonctionnement des outils qu'ils mobilisent? Quels usages doivent faire l'objet d'une divulgation, de la transcription automatisée à la génération ou à la transformation de contenus?

Ces tensions seront examinées en considérant la transparence comme une pratique relationnelle de reddition de comptes entre médias/journalistes et publics. On s'intéressera notamment aux limites des dispositifs de divulgation, au risque d'une transparence essentiellement symbolique, ainsi qu'aux arbitrages entre explicabilité, efficacité professionnelle et confiance. À partir de travaux récents sur l'IA et le journalisme, la présentation proposera de distinguer plusieurs moyens et formes de transparence et d'explicabilité afin d'identifier les conditions dans lesquelles ces valeurs peuvent devenir des pratiques professionnelles significatives plutôt que de simples principes normatifs.

JEUDI 22 octobre 2026

Emanuelle Boutin-Gilbert – discrimination et invisibilisation des communautés queers en ligne

Emanuelle Boutin Gilbert s'intéresse aux transformations sociales engendrées par les technologies numériques. Son parcours, à l'intersection des communications, de l'innovation et de la recherche, l'a menée à collaborer avec des organisations variées, allant d'incubateurs pour startups du Québec au Mila - Institut québécois d'intelligence artificielle.

Formée en communication, elle a d'abord étudié les stratégies communicationnelles des groupes militants sur les médias sociaux et le rôle des émotions dans les discours visuels en ligne. Ses intérêts de recherche actuels portent sur les discriminations algorithmiques, la reproduction des dynamiques de pouvoirs sur les plateformes numériques et les stratégies de résistances déployées sur les médias sociaux par les groupes marginalisés. Elle privilégie une approche queer et féministe afin d'analyser les rapports de pouvoir, les représentations et les inégalités qui traversent les environnements numériques contemporains.

Résumé : Les plateformes numériques jouent un rôle central dans l'(in)visibilisation des communautés marginalisées. Derrière ces interfaces se trouvent des systèmes algorithmiques qui évaluent, classent et modèrent les contenus selon des règles souvent opaques qui influencent les comportements, les stratégies de publication et les possibilités d'exister en ligne. Cet atelier propose de mettre en lumière les conséquences de ces mécanismes sur le quotidien des créateurices queers, les formes d'autocensure qu'ils suscitent ainsi que les stratégies de résistance qui émergent. En s'appuyant sur des exemples concrets, l'atelier propose une réflexion critique sur les liens entre la reproduction des dynamiques de pouvoir, les normes de genre et l'intelligence artificielle.

Pierre Larouche – L'IA, l'ouverture et la flexibilité

Expert en droit de la concurrence et en gouvernance économique ainsi qu'en responsabilité civile, dans les traditions de droit civil et de common law, **Pierre Larouche** est professeur titulaire en droit et innovation à la Faculté de droit de l'Université de Montréal.

Diplômé des facultés de droit des universités McGill, de Bonn et de Maastricht, le Pr Larouche était depuis 2002 professeur titulaire de droit de la concurrence à l'Université de Tilburg, aux Pays-Bas, où il a notamment cofondé le Tilburg Law and Economics Center (TILEC), devenu l'un des plus grands centres de recherche en gouvernance économique au monde. Le Pr Larouche a aussi conçu et lancé un programme de 1er cycle innovateur, le Bachelor Global Law, à l'Université de Tilburg et lancé l'option doctorale « Innovation, science, technologie et droit » à l'Université de Montréal. Lors de son mandat de vice-doyen, il a aussi mené à bien la réforme du programme de droit de 1er cycle. Il a également enseigné au Collège d'Europe, à

Bruges, tout en étant professeur invité dans de nombreuses universités en Amérique (Northwestern, Pennsylvanie), en Europe (Sciences Po, Bonn) et en Asie (Singapour).

Le Pr Larouche a publié plus d'une soixantaine de monographies, d'articles et de contributions scientifiques, et ses travaux, cités par la Cour de justice de l'Union européenne et la Cour suprême du Royaume-Uni, ont influencé les politiques de la Commission européenne en matière de communications électroniques, de concurrence et de normalisation.

En plus d'avoir pratiqué le droit en cabinet privé, il a été auxiliaire juridique de l'honorable Charles D. Gonthier à la Cour suprême du Canada.

Résumé : Depuis des décennies, la recherche en sciences sociales, dans un cadre libéral et démocratique, cherche à incorporer les considérations liées à l'information dans le cadre analytique. S'est ensuivie une meilleure compréhension des effets dynamiques, comme en fait foi l'apparition, dans le discours académique et politique, des imperfections et incertitudes informationnelles, à travers des valeurs ou principes comme l'ouverture (au sens du non-déterminisme) et la flexibilité. L'IA, en phase avec l'amélioration de ses performances et sa diffusion dans l'économie et la société, affectera nécessairement ce discours. Cette présentation explore les implications de l'IA sur l'ouverture et la flexibilité dans le contexte plus spécifique du fonctionnement des marchés. Les nombreuses théories économiques du marché peuvent être classifiées selon qu'elles lui attribuent une fin d'efficience or de découverte. Selon la perspective théorique, l'impact de l'IA diffère. Si le marché mène à l'efficience, l'IA pourrait en améliorer le fonctionnement, mais non sans contrepartie. Si par contre, le marché est un outil de découverte, l'IA pourrait en saper les fondements. Dans les deux cas, une intervention juridique pourrait s'avérer nécessaire pour éviter le scénario catastrophe où la société cesse de récolter les bénéfices du marché. Cela ramène à l'avant-scène une autre controverse théorique sur les marchés, qui oppose naturalisme et constructivisme. La présentation se conclut en revenant sur le terrain plus général du discours académique et politique sur l'ouverture et la flexibilité.

Siddarth De Souza - A Data Justice perspective of Global Data Law

Siddharth Peter de Souza est professeur adjoint en IA et société au Centre for Interdisciplinary Methodologies de l'Université de Warwick. Ses travaux explorent la manière dont les données sont gouvernées à l'échelle mondiale dans des contextes contestés et pluralistes.

Il est le fondateur de Justice Adda, une entreprise sociale spécialisée en droit et en design en Inde, où il travaille sur des questions liées à l'accès à la justice et à l'autonomisation juridique.

Siddharth a travaillé comme consultant en politiques numériques auprès du Programme des Nations Unies pour le développement et de l'Organisation internationale de droit du développement. Il a auparavant été chercheur postdoctoral au sein du Global Data Justice Project au Tilburg Institute for Law, Technology and Society.

Résumé : Le droit mondial des données (Global Data Law) a émergé pour couvrir une variété de domaines, allant du droit de la protection des données aux législations relatives au partage

des données, aux réglementations sur l'accès aux données, aux cadres de gouvernance des données, aux réglementations relatives au commerce numérique et à la réglementation de l'intelligence artificielle. Avec ces développements, les questions relatives à la manière de penser les données, leurs infrastructures, la façon dont elles sont contrôlées, utilisées et valorisées deviennent essentielles en raison du pouvoir qu'ont les données d'influencer le développement socio-économique.

Cette présentation examinera différentes propositions relatives au droit mondial des données formulées par des organisations internationales, des entreprises privées et des États, notamment les États-Unis et la Chine, ainsi que par des entités régionales telles que l'Union européenne. Parmi ces propositions, elle abordera le défi croissant consistant à mettre en lumière les frictions qui constituent les modèles de droit des données à mesure que ceux-ci acquièrent un caractère « mondial ».

La présentation abordera ensuite les débats autour de la justice des données, qui mettent en évidence la manière dont la mise en données (datafication) a des implications pour les questions de représentation, d'autonomie, de capacités et d'inégalités structurelles qui existent au sein d'une économie numérique. Elle proposera enfin une réflexion sur ce à quoi pourrait ressembler le droit mondial des données s'il était étudié à partir d'une perspective de justice des données.

Olivier Servais – Quand l'algorithme tranche : qui fixe l'ordre des valeurs à l'ère de l'IA ?

Je suis historien et anthropologue, professeur titulaire à l'Université de Louvain, où j'ai également été doyen de la Faculté des sciences économiques, sociales et politiques. J'enseigne l'anthropologie des systèmes symboliques et leurs liens avec nos environnements, qu'ils soient considérés comme « naturels » ou « artificiels ». J'enseigne également à l'Institut des Arts de Diffusion (IAD) et au Campus Saint-Louis à Bruxelles.

Ma carrière m'a emmené dans une variété d'endroits : le Canada, les États-Unis, les Philippines, Maurice, la France, la Belgique... et le monde numérique. Aujourd'hui, mes recherches explorent les imaginaires des mondes virtuels, la sociabilité en ligne, les rituels et les formes émergentes de collectifs numériques. Je m'intéresse aux joueurs hardcore, aux hackers et aux programmeurs, aux interactions entre les humains et l'intelligence artificielle, mais aussi plus largement aux façons dont la technologie numérique reconfigure les façons dont nous formons des sociétés, défions et créons ensemble. Récemment, je me suis intéressé aux mouvements de déconnexion et de résistance à la technologie numérique et, enfin, aux non-humains non vivants en tant qu'acteurs sociaux (glaciers, pergélisol, eau, robots, IA, programmes).

Mes publications récentes comprennent Dans la peau des joueurs (Karthala, 2020), Jeux vidéo, nouvel opium du peuple ? (Karthala, 2020), Religieux et numérique (Social Compass, 2020), Ethnographies du catholicisme contemporain (Karthala, 2021), Déconfiner l'Europe (Karthala, 2024) et Catastrophes (Karthala/B, 2024).

Résumé : Les systèmes d'IA générative ne sont jamais neutres : chaque modèle encode, sans toujours l'avouer, une hiérarchie implicite entre les valeurs qu'il privilégie. Ce que l'on appelle

« alignement » recouvre en réalité des arbitrages silencieux entre efficacité et équité, innovation et précaution, liberté et protection. Or ces hiérarchies ne sont pas le fruit d'un choix collectif éclairé, mais le produit de logiques économiques, de biais culturels et de rapports de pouvoir qui se sont cristallisés dans les données et les objectifs d'optimisation. Cet intervention propose de rendre visible cette architecture de valeurs pour la soumettre au débat démocratique. Elle interroge en particulier la tension entre les valeurs inscrites dans la technologie elle-même et celles que les cadres de gouvernance prétendent imposer par-dessus. Trois questions structurantes guideront la réflexion : qui décide de l'ordre des valeurs, sur quelle légitimité, et comment rendre ces choix contestables ? L'enjeu n'est pas de trouver la hiérarchie « juste », mais d'instituer les conditions d'un pluralisme de valeurs capable de résister à l'uniformisation algorithmique. A partir d'exemple de terrain, l'objectif est de dégager des critères concrets pour évaluer la cohérence entre les valeurs proclamées par les régulateurs et celles effectivement opérationnalisées par les systèmes déployés et les usagers.

VENDREDI 23 octobre 2026

Georges Azzaria - L'effacement de la singularité juridique des individus par les algorithmes

Georges Azzaria est professeur à la Faculté de droit de l'Université Laval à Québec. Ses premières recherches ont porté sur les rapports entre l'art et le droit d'auteur, ainsi que sur le statut socio-économique des artistes. Depuis quelques années, il s'intéresse plus particulièrement à l'intelligence artificielle sous l'angle de la propriété intellectuelle et de la régulation juridique. De 2017 à 2025, il a été le directeur de l'École d'art de son université.

Résumé : La créativité humaine valorisée par la propriété intellectuelle et la consécration de la personnalité juridique ont accordé à l'individu, à partir du 18e siècle en Occident, une place centrale dans le droit. Or, la capacité de l'intelligence artificielle générative à remplacer l'humain opère aujourd'hui un renversement qui tend à effacer cette singularité et reléguer l'individu au rôle d'un relais informatique.

MONDAY october 19

Lyse Langlois - Governing AI: Norms, Knowledge and Power

Lyse Langlois has been Executive Director of the International Observatory on the Societal Impacts of AI and Digital Technology (Obvia) since 2018 and is a Full Professor in the Department of Industrial Relations at Université Laval. She directed the Institute of Applied Ethics (IDÉA) at the same university for eight years and is a regular researcher at the Interuniversity Research Centre on Globalization and Work (CRIMT), as well as at Université Laval's Institute Intelligence and Data (IID). In 2025–2026, she was a Visiting Professor at Korea University, with which she is pursuing several collaborative projects.

Her research has resulted in numerous scientific publications, books, articles, and reports on ethics and transformations of work related to technology. Her most recent contribution focuses on the geopolitics of science and science diplomacy, in a chapter published by Palgrave Macmillan in *Polycentric Science Diplomacy in an Age of Disruption*, edited by J. Lüdert and M. Wählisch.

Abstract: The governance of artificial intelligence is based on a growing set of ethical principles, technical standards, and regulatory frameworks. These are complemented by instruments designed to translate these standards into practice and assess their implementation, such as UNESCO's Readiness Assessment Methodology (RAM) and Ethical Impact Assessment (EIA), the NIST AI RMF, or the ISO/IEC 42001 standard.

Governing AI is not limited to determining what is acceptable. It also requires deciding what deserves to be known, measured, and made visible, and therefore what will remain outside the scope. This conference examines the relationships between norms, knowledge, and power. Three questions guide the reflection. Through what processes do certain values become governance norms, and what forms of knowledge give them legitimacy? What do these norms leave in the shadows, and finally, under what conditions can governance subject its own norms to criticism? I will argue that credible AI governance must accept being challenged. This requires independent and pluralistic evaluation mechanisms.

Sarah Brown - Relational Reinterpretations of AI

Dr. **Sarah M Brown** is Assistant Professor of Computer Science at the University of Rhode Island. Her research focuses on how to make AI systems better for all by studying how the technology, the people who make it, and its social impacts interact. Her work centers people at all scales: developing interventions that support data scientists developing real world systems and directly developing techniques that incorporate social values into AI systems. She holds BS, MS, and PhD degrees in Electrical Engineering from Northeastern University. Dr. Brown is a recipient of the NSF Graduate Research Fellowship and CAREER awards

Abstract : Many values typically attributed to AI are based in the language and some can be shaped with post-training alignment strategies. In other cases, certain values are traced back to technical constraints. However, often these values are not fundamental to the mathematical or computational object itself, but derive to a dominant philosophy of science. The value is imposed in the interpretation of the object and then stick with it through implementation of real technologies. These values are often presented as inescapable in the pursuit of building

technology, in part, because philosophy of science is not central in education of computational fields. In reality, nearly all ideas become contestable when examined deeply. In this talk, I will highlight some examples of phenomenon. Then I will pose an alternative philosophical foundation: relational realism. I will motivate its utility and appropriateness for AI to introduce the key tenets and conclude with examples of ways to reinterpret technical constraints relating to AI from this perspective that enable using AI in ways that promote values often under-represented in AI.

Marc-Antoine Dilhac - The Evolution of AI Governance Frameworks: From the Statement of Principles to Their Implementation

Marc-Antoine Dilhac is Professor of Ethics and Political Philosophy at the Université de Montréal, where his work focuses on the ethics and governance of artificial intelligence. In 2017, he initiated and led the co-construction process for the Montréal Declaration for a Responsible Development of Artificial Intelligence. He is Director of Governance and International Collaboration at OBVIA and a member of Mila, where he held a Canada–CIFAR AI Chair in AI Ethics. He serves as a UNESCO expert in the AI Ethics Experts Without Borders network and has led implementations of UNESCO's Readiness Assessment Methodology (RAM) in Uzbekistan and Québec. He previously served on the Government of Canada's AI Advisory Council and was the expert responsible for the Council of Europe's consultation process on the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.

Abstract: From the earliest ethical declarations on artificial intelligence, their concrete implementation, through both corporate procedures and public policies, was identified as a key priority. Ethical values and principles needed to be "operationalized" to effectively guide the design, deployment, and regulation of AI systems. However, meeting this imperative has proven more difficult than initially anticipated. Nearly ten years after the launch of the Montreal Declaration for a Responsible Development of Artificial Intelligence initiative, this conference offers an assessment of how AI governance has evolved. How have governance frameworks changed? What have we learned from their implementation? What remains to be done to ensure that we retain control over the development of AI? To address these questions, the conference will examine influential non-binding frameworks, such as UNESCO's Recommendation on the Ethics of Artificial Intelligence and the Hiroshima AI Process International Code of Conduct (G7/OECD), which is accompanied by a corporate self-reporting mechanism, as well as binding frameworks, such as the Council of Europe's Framework Convention on Artificial Intelligence and the European Union's AI Act. It will argue that values and principles are only effectively implemented when governance frameworks are equipped with methods of verification that make it possible to demonstrate, rather than merely declare, compliance with ethical standards.

Alison Gillwald - Redressing the uneven distribution of AI harms and opportunities through prioritising second and third generation rights and value of 'ubuntu'

Alison Gillwald (PhD) is the founding director of Research ICT Africa (RIA) and is an Adjunct Professor at the University of Cape Town's Nelson Mandela School of Public Governance, where she has run a Pan African digital policy doctoral program. With RIA a knowledge partner to South Africa's G20 Presidency, she served as a technical advisor to the Digital Economy Work Group and the High-Level Task Force on Artificial Intelligence, Data Governance, and Innovation for Sustainable Development. She also co-chaired the T20 Task Force on Digital Transformation. She is a member of the French G20 Presidencies' T7 working group on AI and Middle Powers and serves on the International Telecommunications Union's Academic Advisory Body on Enhanced Technologies. She has been active in the Global Partnership on Artificial Intelligence (GPAI), co-leading the Algorithmic Transparency in the Public Sector and Data Justice projects.

Alison led RIA's technical assistance to the African Union Commission, facilitating the International Task Team and drafting the AU Data Policy Framework and Implementation Plan. Currently, she is the principal investigator on the African Development Bank project on Digital Public Infrastructures for Africa, which examines derisking bank investments through a political economy lens. Her latest publications include the framing paper for the Club de Madrid Dialogue on Governance of Global Public Goods in a Fragmented World: Digital Public Goods and Infrastructure and Gillwald, A., and Fuller, J. (2026). Digital Transformation and the Limits of the G20: Lessons from the Emerging Economy G20 Presidencies.

TUESDAY October 20

Alexandra Bensamoun - Transparency: A Value or an Instrument? The AI Act and Cultural Challenges

Alexandra Bensamoun is Professor of Law at Université Paris-Saclay, specializing in intellectual property and digital regulation. She is Director of the Master 2/LLM programme Fundamental Intellectual Property and Digital Technologies, a joint degree programme with Université Laval (Quebec) and the Autonomous University of Madrid (Spain). An expert advisor to international (including UNESCO) and national authorities, and a qualified member of the French High Council for Literary and Artistic Property (CSPLA, French Ministry of Culture), she served on France's Interministerial Commission on Artificial Intelligence, which submitted a report on France's AI strategy to the President of the Republic, AI: Our Ambition for France (March 2024). Her reports for the French Ministry of Culture notably address the remuneration of cultural content used by AI systems (which led to Bill No. 220, unanimously adopted by the French Senate in April 2026), as well as the legal status of AI-generated outputs. She is the author of numerous individual and collaborative research publications, including a collective treatise on Artificial Intelligence Law (Droit de l'intelligence artificielle, LGDJ/Lextenso, 3rd ed., 2026).

Abstract: Transparency is often presented as one of the core values underpinning European AI regulation. In the cultural sectors, it takes on particular significance, as it is essential both to the effective protection of creators' rights and to public trust in the content being circulated. Drawing on a cross-cutting contemporary issue—that of the signature and authenticity of works—this presentation will examine the two stages of transparency under Regulation (EU) 2024/1689.Upstream, it will explore the interaction with the text and data mining exception and its opt-out mechanism, the sufficiently detailed summary of training content and the template published by the European Commission, as well as the proposed presumption of use under French law, which highlights the limitations of information disclosure. Downstream, it will address the marking of AI-generated content and the disclosure of information to the public, including the adapted regime provided for artistic works. The presentation will also assess whether transparency, conceived primarily as an obligation to provide information, can become a genuine tool for preserving cultural values.

Maurice Jones - ArtIA: Towards a New Digital Commons for Creative AI

Dr. Maurice Jones is a researcher, artist, and cultural strategist based in Tiohtià/Montréal. He recently defended his interdisciplinary PhD dissertation in Humanities at Concordia University. His research areas focus on festivals and digital culture infrastructures, cultural AI and the commons, transnational AI governance, as well as digital research-creation.

Abstract: Based on "ArtIA", an 18-month consultation conducted with more than 200 artists, cultural institutions, technologists, and creative sector workers across Canada, this presentation explores the possibilities of commons-based approaches to creative AI. Dr. Maurice Jones examines how collective ownership, shared infrastructures, critical AI literacies, and creative reappropriation can offer alternatives to dominant industry-driven models of AI development and adoption, and raises the question of what it would mean to build creative AI ecosystems rooted in cultural and community needs rather than in the extractive logics of platforms.

Antoine Henry - Operationalizing Values in Cultural Recommendation Systems: The Creation of a Fairness Score

Antoine Henry is an Associate Professor in Information and Communication Sciences and a member of GERIICO's Axis 4, whose theme is the circulation of information and the organization of knowledge.

His research focuses mainly on issues related to digital technologies (AI, digital commons, algorithmic ethics), organizational transformation, and collective intelligence.

He is also:

- a member of the Board of Directors of the French chapter of the scholarly association ISKO;
- a member of the research group of the Centre for Internet and Society (CNRS UPR 2000), within which he co-leads the working group on AI, art, and creativity.

Abstract: Algorithmic fairness represents a major challenge in the development of recommendation systems, particularly in the music sector. This presentation proposes an

innovative approach to current algorithmic fairness frameworks by introducing a new model for assessing fairness. We develop a method for synthesizing nine complex dimensions in order to influence the feedback loop of algorithms. Through this proposed Fairness Score, we aim to operationalize fairness and provide users with meaningful control over their experience and cultural practice.

Christophe Abrassard – Digital Sobriety, a New Paradigm for AI in an Uncertain World

Christophe Abrassart is an Associate Professor at the School of Design of the Faculty of Environmental Design at the Université de Montréal. A specialist in eco-design and strategic foresight, he is Co-Director of the Lab Ville Prospective at the Faculty of Environmental Design and Co-Lead of OBVIA's Research Axis on Digital Sobriety and the Socio-Ecological Transition. For the past 15 years, he has designed and facilitated co-design and participatory foresight workshops on the ecological and energy transition in Quebec with various organizations (e.g., the City of Montreal, Hydro-Québec). In particular, he coordinated a foresight exercise at OBVIA on digital sobriety in Quebec looking ahead to 2050.

Abstract: The current world is marked by a profound antinomy. On the one hand, environmental scientists are sounding the alarm about the climate emergency and the imperative of returning as quickly as possible to a "safe operating space for humanity" within planetary boundaries (Rockström, 2009). On the other hand, digital multinationals are accelerating their investments in ever-larger data centres that consume increasing amounts of materials, water and energy. To such an extent that the ecological and digital transitions, once presented as complementary, have become contradictory. For example, between the electrification of transportation, the adaptation of buildings to periods of extreme heat, and the electrification of data centres, not everything will be achievable without major tensions within the energy system over the coming decades (Shift Project, 2025).

Faced with this contradiction, a first approach is that of eco-design, aimed at optimizing the energy performance of data centres. However, this sector is characterized by continuous rebound effects. Another approach is that of systemic sobriety, which consists in critically questioning the uses of AI and exercising discernment between priority projects and others (e.g., between AI for healthcare and AI for entertainment or experiential marketing), or in reflecting on the appropriate place of digital technologies in society (e.g., in schools or organizations). To explore the potential of this approach and imagine the sociotechnical innovations to which it could lead, this presentation will draw on four foresight scenarios on digital sobriety in Quebec in 2050, resulting from a project coordinated by the speaker at OBVIA.

Jocelyn Maclure - Why the Value-Alignment Approach is Misguided

Jocelyn Maclure is Professor of Philosophy and the Jarislowsky Chair in Human Nature and Technology at McGill University. His research and teaching interests include ethics, political philosophy, the philosophy of AI, and social epistemology. His recent work has appeared in journals such as Minds & Machines, AI & Ethics, Digital Society, and AI & Society. He has

recently been a visiting professor at the Universities of Bonn and Grenoble, and served as Chair of Quebec's Commission de l'éthique en science et en technologie from 2017 to 2024.

The "value-alignment problem" (VAP) is arguably the most influential approach among philosophers and AI researchers in the fields of AI ethics and governance. VAP is centered on the question of how to build advanced AI systems whose goals and outputs are aligned with human values. While there is obviously nothing wrong with grounding the development of frontier AI in shared human values, this approach leaves out the most difficult aspects of practical ethics and policy making. After reviewing the main problems facing VAP (the pluralism, underdetermination and motivation problems), I will sketch out an alternative approach, which I call "Normative Integrity".

WENESDAY October 21

Jean-Gabriel Ganascia - Agentic AI and/or Human Oversight?

An engineer and philosopher by training, **Jean-Gabriel Ganascia** turned his attention to computer science and artificial intelligence at an early stage in his career. In 1983, he completed a doctoral thesis on expert systems at Paris-Saclay University, followed in 1987 by a thèse d'État on machine learning at the same institution. He is currently Professor Emeritus of Computer Science at Sorbonne University and a EurAI Fellow. He is also an honorary member of the Institut Universitaire de France and a member of LIP6 (Sorbonne University's Laboratory of Computer Science). His current research activities focus on artificial intelligence, computational ethics, computer ethics, the detection of fake news and digital humanities. He has previously contributed to French academic research in cognitive science by creating and leading the scientific interest group 'Sciences de la cognition'. He also initiated the OBVIL Labex (Observatoire de la vie littéraire – Observatory of Literary Life), which coordinated the efforts of his team at LIP6 and the literary teams at Sorbonne University in the field of digital humanities. Alongside his research activities, he chairs the Ethics Committee of France Travail (the French employment agency), the Ethics Committee of Docaposte and the Steering Committee of the CHEC (Cycle des Hautes Études de la Culture). He is also a member of COMEST (the UNESCO World Commission on the Ethics of Scientific Knowledge and Technology). He was President of the CNRS Ethics Committee (COMETS) from 2016 to 2021, and President of the Association Française pour l'Avancement des Sciences (AFAS) from 2022 to 2026. Throughout his career, he has published over 600 papers in conference proceedings, scientific journals, and books. He is also the author of a dozen books for a general audience, and since 2017 he has been writing several regular columns on ethics, one in the journal La Recherche and the other in the journal Sciences et avenir.

Abstract: Following the boom in generative AI, we have been entering the age of agentic AI since the end of 2025. The idea is an old one: an agent captures information—text, image, or sound—interprets it, makes decisions on its own on the basis of its inputs, and acts (hence the name "agent") in order to achieve the goals that have been set for it. How can we not be concerned about this? Agentic AI risks placing us under its control by bypassing us. In the ethics of care, the notion of "human guarantee" has been introduced to ensure that no decision is made

without first having been approved by a healthcare professional. Why not always proceed in this way? Because we do not have the time, and because humans are not a guarantee! Errare humanum est: we sometimes make mistakes more often than automated systems do. Retrospective studies of accidents involving airplanes or nuclear power plants show this. How can we resolve, in advance, potential conflicts of authority between humans and machines and decide which one should prevail? Sometimes, on the basis of rational arguments, we choose to leave control to the artificial agent. But this delegation of responsibility is only valid for a limited period of time and must remain revocable at any moment. At least, we hope so… These are all the questions that will be addressed here in order to consider what, in each situation, remains acceptable and what must be condemned.

Maroussia Lévesque - Human Rights: the North Star of AI Governance in practice

Maroussia Lévesque joined Queen's Law as an Assistant Professor in Law & AI in 2026. Her research focuses on AI governance, examining regulatory design, power imbalances and market concentration in the AI industry, and the respective roles of public and private actors in the governance ecosystem. Her current projects analyse how the corporate structures of frontier AI labs shape their behaviour in practice, along with the environmental implications and judicial uses of AI.

She is an affiliate at the Berkman Klein Center for Internet & Society, a Senior Fellow at the Centre for International Governance Innovation (CIGI), and a collaborator on the Indigenous-led Abundant Intelligences project.

Prior to joining Queen's, Professor Lévesque completed her SJD at Harvard Law School, where she lectured in Comparative AI Governance (with Professor Martha Minow) and was a Fellow at the Weatherhead Center for International Affairs. She previously led AI and human rights work at Global Affairs Canada, clerked for the Chief Justice at the Quebec Court of Appeal, collaborated with Amnesty International's Algorithmic Accountability Lab, and assisted in setting up the Global Partnership on AI. She is a member of the Quebec Bar.

Professor Lévesque's scholarship has appeared in the NYU Journal of Legislation & Public Policy, the University of Illinois Journal of Law, Technology & Policy, and NeurIPS proceedings, among others. She has written for The Guardian, the Toronto Star, La Presse, The Walrusand Tech Policy Press.

Abstract: Human rights are an appealing candidate to guide AI governance, as they provide an established vocabulary and international legitimacy, supported by existing institutional machinery. The UN Guiding Principles on Business and Human Rights extend that vocabulary to private actors, setting out a corporate responsibility to respect human rights alongside the state's binding duty to protect them. Asking which values AI governance should promote is easier than asking what makes a value operative.

This presentation will unpack the potential and limitations of human rights to guide AI governance. Early on in 2018 as machine learning picked up speed, the Toronto Declaration translated the Guiding Principles' due diligence requirement into concrete steps for machine

learning: providing actionable guidance to identify risks, prevent and mitigate them, track responses, and report publicly. More recently, Meta's Oversight Board described itself as an independent non-judicial grievance mechanism under the Guiding Principles, invoking human rights when reviewing automated content moderation. These examples show that human rights can be implemented to move the needle on AI governance.

That said, human rights come with limitations. Non-binding commitments can confer unearned legitimacy dividends to voluntary undertakings, with companies shrouding their actions in the language of pliable norms without the substance of enforcement. Voluntary frameworks may also buy companies time in terms of maintaining the status quo as to the absence of binding regulation, normalizing their problematic practices while societal expectations adjust. Finally, soft commitments give way under competitive pressure, as recent retreats from human rights pledges suggest.

Turning to current developments, investor protection rules requiring thorough risk disclosure may advance due diligence further as frontier AI labs go public. The proposed binding treaty on business and human rights would also require states to regulate rather than ask companies to restrain themselves. Ultimately, articulating values matters less than building the scaffolding of incentives and obligations to implement them.

Richard Khoury - Generative AI and the War Against Linguistic Diversity

Richard Khoury obtained his Bachelor's and Master's degrees in Electrical Engineering and Computer Engineering from Université Laval (Quebec City, QC) in 2002 and 2004 respectively, and his PhD in Electrical Engineering and Computer Engineering from the University of Waterloo (Waterloo, ON) in 2007. From 2008 to 2016, he was an Assistant Professor, then an Associate Professor, in the Department of Software Engineering at Lakehead University. In 2016, he joined Université Laval as an Associate Professor, before being appointed Full Professor in 2026. From 2021 to 2023, he was also President of the Canadian Artificial Intelligence Association. His research interests are in data mining and natural language processing, and also include knowledge management, machine learning, and artificial intelligence.

Abstract: Generative AI tools are now available in more than a hundred different languages. At first glance, this seems to be a victory for linguistic diversity: artificial intelligence makes it possible to reach users around the world in their native language. But beneath this positive image lies an insidious problem. Artificial intelligence only knows and understands the dominant dialect of these languages and imposes its use on speakers. This linguistic colonization by AI threatens to erase global linguistic diversity and suppress minority cultural communities. In this presentation, we will examine this problem, its origins, its symptoms, and its solutions.

Colette Brin – Explainability and Transparency in Journalism: Some Challenges in Putting Responsible AI Values into Practice

Colette Brin is a Full Professor in the Department of Information and Communication at Université Laval and Director of the Centre for Media Studies. A specialist in the transformations of journalism and information practices, she is co-lead of the Arts, Media and Cultural Diversity research axis of the International Observatory on the Societal Impacts of AI and Digital Technology (OBVIA). Her recent research focuses in particular on the responsible integration of artificial intelligence in the media, journalistic values, and audience needs. She has published Médias publics, IA responsable et besoins des utilisateurs (2025) and, with Véronique Guèvremont, Enjeux de l'intelligence artificielle pour les milieux de la culture et des médias d'information au Québec (2026). She also co-edited, with Véronique Guèvremont, the collective volume Intelligence artificielle, culture et médias (Presses de l'Université Laval, 2024).

Abstract: Transparency and explainability are among the principles most frequently associated with the responsible use of artificial intelligence. Their implementation in journalistic practices, however, raises particular difficulties. How can the involvement of AI systems in news production be made visible and understandable without overloading audiences with technical information? What does it mean to explain the use of AI when journalists and media organizations themselves have limited knowledge of how the tools they use actually work? Which uses should be subject to disclosure, from automated transcription to the generation or transformation of content?

These tensions will be examined by considering transparency as a relational practice of accountability between media organizations/journalists and audiences. Particular attention will be paid to the limitations of disclosure mechanisms, the risk of essentially symbolic transparency, and the trade-offs between explainability, professional efficiency, and trust. Drawing on recent research on AI and journalism, the presentation will propose distinguishing between several means and forms of transparency and explainability in order to identify the conditions under which these values can become meaningful professional practices rather than mere normative principles.

THURSDAY October 22

<u>Emanuelle Boutin-Gilbert – Under the Algorithmic Gaze: Discrimination and Queer In/Visibility Online</u>

Emanuelle Boutin Gilbert is interested in the social transformations driven by digital technologies. Her professional and academic background, at the intersection of communications, innovation, and research, has led her to collaborate with a wide range of organizations, from Quebec startup incubators to Mila, the Quebec Artificial Intelligence Institute.

Originally trained in communications, she first studied the communication strategies of activist groups on social media and the role of emotions in online visual discourse. Her current research interests focus on algorithmic discrimination, the reproduction of power dynamics on digital platforms, and the resistance strategies employed by marginalized groups on social media. She

adopts a queer and feminist perspective to analyze the power relations, representations, and inequalities that shape contemporary digital environments.

Abstract: Digital platforms play a central role in the (in)visibility of marginalized communities. Behind these interfaces lie algorithmic systems that evaluate, rank, and moderate content according to often opaque rules that shape user behavior, publishing strategies, and the very possibilities of existing online.

This workshop explores the impact of these mechanisms on the everyday experiences of queer content creators, the forms of self-censorship they encourage, and the strategies of resistance that emerge. Drawing on concrete examples, the workshop invites participants to critically examine the connections between the reproduction of power dynamics, gender norms, and artificial intelligence.

Pierre Larouche - AI, open-endedness and flexibility

As an expert in competition law, economic governance, and civil liability within the traditions of civil and common law, **Pierre Larouche** is a full professor in law and innovation at the Faculty of Law at the Université de Montréal. A graduate from the law faculties of McGill University, the University of Bonn, and Maastricht University, Mr. Larouche has been a full professor in competition law at Tilburg University, in the Netherlands, since 2002, where he notably co-founded the Tilburg Law and Economics Center (TILEC), a world-leading research centre in economic governance. Over the course of his academic career in Europe, Mr. Larouche also developed and established an innovative undergraduate program, the Bachelor in Global Law, at Tilburg University. Furthermore, he taught at the College of Europe in Bruges, and was a visiting professor at several universities in America (Northwestern, Pennsylvania), Europe (Sciences Po, Bonn), and Asia (Singapore).

Prof. Larouche has published some sixty monographs, articles, and scientific contributions, and his work, which has been cited by the Court of Justice of the European Union and the Supreme Court of the United Kingdom, informed the electronic communications and competition policies adopted by the European Commission.

In addition to his private law practice, he served as law clerk to the Honourable Charles D. Gonthier of the Supreme Court of Canada. Prof. Larouche also participates in the activities of the Public Law Research Centre (CRDP) and the Centre of the Law of Business and International Trade (CDACI). Additionally, he contributed greatly to the “Innovation, Science, Technology, and Law” doctorate option launched by the Faculty in the fall of 2017.

Abstract: A key plank of the liberal democratic social science research agenda of the last decades has been to incorporate insights from information research. This has typically led to a deeper understanding and appreciation for dynamic effects, as reflected in the embrace, in academic and policy discourse, of informational imperfections and uncertainty through values/principles such as open-endedness (dynamic version of openness) and flexibility. As its performance improves and it spreads throughout the economy and society, AI is likely to affect

this area of research and policy. This presentation explores the implications of AI on open-endedness and flexibility in the more specific context of markets. Competing economic theories of the market can be roughly sorted out according to the purpose they ascribe to markets (efficiency or discovery). Depending on the theoretical perspective, the anticipated impact of AI on markets will differ. If markets deliver efficiency, then AI could improve on their performance, but at a price. If however markets are a discovery tool, AI could undermine the very foundation of markets. In both cases, some legal intervention might therefore be necessary to avoid worse-case scenarios where socially-beneficial properties of markets would be lost. This sheds a new light on a different theoretical controversy about markets, namely the extent to which they are a natural or a constructed phenomenon. The presentation ends by a tentative extrapolation from the case of markets to the broader academic and policy discussion on open-endedness and flexibility.

Siddarth De Souza - A Data Justice perspective of Global Data Law

Siddharth Peter de Souza is an Assistant Professor in AI & Society at the Centre for Interdisciplinary Methodologies, University of Warwick. His work explores how data is governed globally in contested and plural settings. He is the founder of Justice Adda, a law and design social venture in India where he works on matters related to access to justice, and legal empowerment. Siddharth has worked as a digital policy consultant with United Nations Development Programme and the International Development Law Organization and was previously a post-doctoral researcher at the Global Data Justice Project at Tilburg Institute for Law, Technology and Society.

Abstract: Global Data Law has emerged to cover a variety of domains from data protection law, laws around data sharing, data access regulations, data governance frameworks, regulations around digital trade, and AI regulation. With these developments, questions around how to think about data, its infrastructure, how it is controlled, used and valued, become critical because of the power data has to influence socio-economic development.

This talk will discuss various proposals for Global Data Law from international organizations, private enterprises, and nation states including the US, China, and regional entities like the European Union. Amidst these proposals, it will discuss an increasing challenge to surface the frictions that constitute Data Law models, as they take on a 'global' character.

The talk will then discuss debates around data justice, which highlight how datafication has implications for questions of representation, autonomy, capabilities, and structural inequalities that exist within a digital economy. It will offer a proposal for what Global Data Law could look like if studied from a Data Justice perspective.

Olivier Servais – When the Algorithm Decides: Who Sets the Order of Values in the Age of AI?

I am a historian and anthropologist, full professor at the University of Louvain, where I also served as Dean of the Faculty of Economic, Social and Political Sciences. I teach anthropology

of symbolic systems and their links with our environments, whether they are considered "natural" or "artificial". I also teach at the Institut des Arts de Diffusion (IAD) and Campus Saint-Louis in Brussels.

My career has taken me to a variety of places: Canada, the United States, the Philippines, Mauritius, France, Belgium... and the digital world. Today, my research explores the imaginaries of virtual worlds, online sociability, rituals and emerging forms of digital collectives. I am interested in hardcore gamers, hackers and programmers, interactions between humans and artificial intelligence, but also more broadly in the ways in which digital technology is reconfiguring the ways in which we form societies, challenge and create together. Recently, I have been interested in movements of disconnection and resistance to digital technology and, finally, in non-living non-humans as social actors (glaciers, permafrost, water, robots, AI, programmes).

My recent publications include Dans la peau des gamers (Karthala, 2020), Jeux vidéo, nouvel opium du peuple ? (Karthala, 2020), Religious and the digital (Social Compass, 2020), Ethnographies du catholicisme contemporain (Karthala, 2021), Déconfiner l'Europe (Karthala, 2024) and Disasters (Karthala/B, 2024).

Abstract: Generative AI systems are never neutral: every model encodes, without always explicitly acknowledging it, an implicit hierarchy among the values it prioritizes. What is referred to as "alignment" actually encompasses silent trade-offs between efficiency and fairness, innovation and precaution, freedom and protection. Yet these hierarchies are not the result of informed collective choices, but rather the product of economic logics, cultural biases, and power relations that have become crystallized in data and optimization objectives. This presentation proposes to make this architecture of values visible in order to subject it to democratic debate. In particular, it examines the tension between the values embedded in the technology itself and those that governance frameworks claim to impose upon it. Three key questions will guide the discussion: who decides the ordering of values, on what grounds of legitimacy, and how can these choices be made open to challenge? The aim is not to identify the "right" hierarchy, but to establish the conditions for a pluralism of values capable of resisting algorithmic homogenization. Drawing on examples from the field, the objective is to identify concrete criteria for assessing the consistency between the values proclaimed by regulators and those effectively operationalized by deployed systems and their users.

FRIDAY ocotber 23

Georges Azzaria - The Erasure of Individuals' Legal Singularity by Algorithms

Georges Azzaria is a Professor at the Faculty of Law of Université Laval in Quebec City. His early research focused on the relationship between art and copyright, as well as on the socio-economic status of artists. In recent years, he has taken a particular interest in artificial intelligence from the perspective of intellectual property and legal regulation. From 2017 to 2025, he was Director of his university's School of Art.

Abstract: Human creativity, valued by intellectual property, and the recognition of legal personality have, since the 18th century in the West, given the individual a central place in law. However, the capacity of generative artificial intelligence to replace humans is now bringing about a reversal that tends to erase this singularity and relegate the individual to the role of a computer intermediary.