

ملخص تنفيذي

مبادرة IEEE العالمية حول أخلاقيات الأنظمة المستقلة والذكية

المقدمة

انتشرت في الآونة الأخيرة الأنظمة المستقلة والذكية بشكل سريع وفعال، لذلك وجب علينا في جمعية مهندسي الكهرباء والالكترونيات أن نضع مبادئ توجيهية مجتمعية وتنظيمية لكي تظل هذه النظم مفيدة للبشرية ومتماشية مع الاخلاق والتوجهات العامة للمجتمعات وللأفراد، حتى تخدم القيم الإنسانية والمبادئ الأخلاقية. يستخدم على هذه الأنظمة أن تعمل بطريقة تعود بالنفع على الافراد مع ايجاد الحلول للمشاكل الوظيفية ومعالجة المعضلات التقنية. بهذا النفع تبنى الثقة بين الافراد والتكنولوجيا ويتم ملائمتها واستخدامها بأفضل صورة في حياتنا اليومية.

حتى نستطيع أن نساهم بطريقة إيجابية، يجب علينا نحن أعضاء الجمعيات العلمية التقنية أن نعزز من طرقنا للتفكير، وأن يكون لدينا الاستعداد لبناء نقاش مفتوح حول ما نتخيله للمستقبل، ما نؤمن به من قيم، وما نمتلكه من المؤسسات والرموز والتمثيلات.

يوديمونيا، كما أوضح أرسطو، هي الممارسة التي تعرف رفاهية الإنسان كأعلى فضيلة في المجتمع. ترجمت تقريبا بأنها "مزدهرة"، الفوائد من يوديمونيا تبدأ بالتأمل الواعي، حيث أن الاعتبارات الأخلاقية تساعدنا على تحديد كيف نرغب في العيش.

أيا إذا كانت ممارساتنا الأخلاقية الغربية (أرسطو، كانتيان)، أو الشرقية (الشننتو، الكونفوشيوسية)، الأفريقية (أوبونتو)، أو من تقاليد مختلفة، فمن خلال تطوير النظم الذاتية والذكية التي تأخذ في عين الاعتبار حقوق الإنسان وفائدة المستخدمين، يتم تمكيننا من إعطاء الأولوية لزيادة رفاهية الإنسان وإيجاد مقاييس جديدة للتقدم البشري في عصر التقدم الخوارزمي. وتمكين قياس الاقتصاد الكلي والازدهار، ليصبح أكثر أهمية من السعي لتحقيق أهداف أحادية البعد مثل زيادة الإنتاجية أو نمو الناتج المحلي الإجمالي.

خطة مبادرة IEEE العالمية حول أخلاقيات الأنظمة المستقلة والذكية

ضمان تدريب وتعليم ودعم كل المشاركين في تصميم وتطوير النظم الذاتية والمستقلة على إعطاء الأولوية القصوى لأخلاقيات المهنة حتى تصبح هذه التكنولوجيا مفيدة للبشرية.

نقصد "بالمشاركين" أي شخص يشارك في البحث أو التصميم أو التصنيع أو الحديث عن الأنظمة الذكية والمستقلة، بما في ذلك الجامعات والمنظمات، والحكومات، والشركات التي تجعل هذه التكنولوجيات حقيقة واقعة بالنسبة للمجتمع.

هدفنا هو أن يكون **التصميم المتوافق أخلاقيا** لتقديم الأفكار والتوصيات، ولتكون مرجعا رئيسيا لعمل الفنيين في مجالات العلوم المختلفة والتكنولوجيا في السنوات المقبلة. لتحقيق هذا الهدف، في الإصدار الحالي من **تصميم متوافق أخلاقيا (EAD2v2)** ، نحدد "القضايا" ذات الصلة و "التوصيات المرشحة" التي نأمل أن تسهل ظهور السياسات التي تتوافق مع هذه المبادئ

تجمع مبادرة IEEE العالمية عدة مئات من المشاركين من ست قارات، وهم قادة من الأوساط الأكاديمية والصناعية والمجتمع المدني والسياسة والحكومة من أجل تحديد وإيجاد توافق في الآراء بشأن القضايا الأخلاقية في الوقت المناسب.

ويتمثل الهدف الثاني للمبادرة العالمية في IEEE في تقديم توصيات بشأن المعايير القائمة على التصميم المتوافق أخلاقيا . تصميم متوافق أخلاقيا (V1) و (V2) و

أعضاء مبادرة IEEE العالمية هي مصدر الإلهام وراء معايير IEEE P7000 TM وهي مجموعات العمل المتاحة لأي شخص للانضمام والمشاركة.

لمزيد من المعلومات أو للانضمام إلى أي مجموعة عمل،
الرجاء النقر على الروابط أدناه:

- IEEE P7000 TM عملية نموذجية لمعالجة المخاوف الأخلاقية أثناء تصميم النظام
- IEEE P7001 TM شفافية الأنظمة المستقلة
- IEEE P7002 TM عملية خصوصية البيانات
- IEEE P7003 TM اعتبارات التحيز الخوارزمية
- IEEE P7004 TM معيار حول حوكمة بيانات الأطفال والطلبة
- IEEE P7005 TM معيار لحوكمة شفافية بيانات صاحب العمل
- IEEE P7006 TM معيار عميل الذكاء الاصطناعي للبيانات الشخصية
- IEEE P7007 TM المعيار الأنطولوجي للروبوتات وأنظمة التشغيل الآلي المتوافقة أخلاقيا
- IEEE P7008 TM معيار للتوجيه الأخلاقي لأنظمة الروبوتات، الأنظمة الذكية، والأنظمة المؤتمتة
- IEEE P7009 TM معيار التفشير الآمن لأنظمة الحكم الذاتي المستقلة وشبه المستقلة
- IEEE P7010 TM معايير الرفاهية القياسية للذكاء الاصطناعي الأخلاقي والأنظمة التحكم الذاتية.

من نحن

مبادرة IEEE العالمية حول الأخلاق الأنظمة المستقلة والذكية الأخلاقية ("IEEE المبادرة العالمية") هو برنامج لجمعية

المهندسين الكهربائية والإلكترونية ("IEEE")، وهي أكبر منظمة مهنية فنية في العالم مكرسة لتقدم التكنولوجيا من أجل الإنسانية وهي تضم أكثر من 420,000 عضو من أكثر من 160 بلدا.

وتوفر مبادرة IEEE العالمية فرصة لتجميع خبرات متعددة في العلوم والتكنولوجيا والعلمية ذات الصلة المجتمعات المحلية لتحديد وإيجاد توافق في الآراء بشأن القضايا العلمية والتكنولوجية في الوقت المناسب. [Creative Commons](#) جميع إصدارات "التصميم متوافق أخلاقيا" المتاحة تحت ترخيص IEEE سوف تطرح [Attribution-Non-Commercial 3.0 United States License](#).

مع مراعاة شروط هذا الترخيص، تعتمد المنظمات أو للأفراد جوانب هذا العمل في تقديرها في أي وقت. ومن المتوقع أيضا أن هذا المحتوى والموضوع سوف يقدم إلى IEEE في نطاق رسمي لوضع المعايير.

وتساهم مبادرة IEEE العالمية و EAD في تعزيز الجهود لخلق محادثة شاملة وأوسع حول الأخلاق في التكنولوجيا،

والمعروفة باسم برنامج . IEEE Ethics™

تصميم متوافق أخلاقيا - V2 نظرة عامة

الغرض

الأنظمة التقنية الذكية والحكم الذاتي

وهي مصممة خصيصا للحد من التدخل البشري في حياتنا اليومية. وبذلك، تثير هذه المجالات كثير من المخاوف من ناحية التأثير على الأفراد والمجتمعات. تشمل المناقشات الدعوة إلى التأثير الإيجابي فضلا عن التحذيرات، استنادا إلى الضرر المحتمل للخصوصية، والتمييز، وخسارة المهارات، والآثار الاقتصادية، وتأمين للبنية التحتية، والآثار طويلة الأجل على الرفاه الاجتماعي.

بسبب طبيعتها، فإن فوائد هذه التكنولوجيات لن تتحقق إلا إذا كانت تتماشى مع قيمنا المحددة و مبادئنا الأخلاقية. ولذلك يجب أن نشئ أطر لتوجيه وإرشاد الحوار حول الآثار غير الفنية لهذه التقنيات.

الأهداف

التصميم الأخلاقي والتطوير: ينبغي أن يكون تنفيذ هذه التكنولوجيات مسترشدة بالمبادئ العامة التالية:

- حقوق الإنسان: التأكد من أنها لا تنتهك حقوق الإنسان المعترف بها دوليا
- الرفاه: إعطاء الأولوية لمقاييس الرفاهية في تصميمها واستخدامها
- المساءلة: تأكد من أن المصممين والمشغلين مسؤولين وخاضعين للمساءلة
- الشفافية: التأكد من أنها تعمل بكامل الشفافية.
- الوعي بسوء الاستخدام: التقليل من المخاطر من سوء استخدامها

الخطة

حقوق البيانات الشخصية والفردية وصلاحيات الدخول

الحاجة الأساسية هي أن الناس لهم الحق لتعريف الوصول وتوفير الموافقة فيما يتعلق باستخدام المحتوى الرقمي الخاص بهم. يحتاج الأفراد إلى آليات للمساعدة على تنظيم هويات فريدة من نوعها والبيانات الشخصية باستخدام آليات واضحة مدركة العواقب الناتجة عن تجميع أو إعادة بيعها.

الرفاه المروج حسب الآثار الاقتصادية من خلال الوصول بأسعار معقولة لشبكات الاتصالات، وشبكة الإنترنت، إتاحة الأنظمة التقنية الذكية والمستقلة للسكان الاستفادة منها في أي مكان. ومن الممكن أن تغير علاقات المؤسسات بشكل كبير نحو الهياكل التي تركز على الإنسان، مستفيدة من المسائل الإنسانية والإنمائية الناجمة عن ذلك لزيادة الرفاهية الفردية والمجتمعية.

الأطر القانونية للمساءلة

التقارب بين الأنظمة الذكية وتقنيات الروبوتات تؤدي إلى تطور الأنظمة ذات الخصائص التي تحاكي البشر من حيث الاستقلال الذاتي الجزئي، والقدرة على أداء مهام فكرية محددة، وقد يكون لها مظهر مادي. مسألة الوضع القانوني للذكاء معقدة وبالتالي النظم التقنية المستقلة دائما تتشابك مع أسئلة قانونية أوسع فيما يتعلق بكيفية ضمان المساءلة وتحديد المسؤولية عندما تسبب هذه الأنظمة ضررا. بعض الأمثلة من الأطر العامة للنظر فيها تشمل الآتي:

- أنظمة تقنية ذكية ومستقلة تخضع للنظم المعمول بها في قوانين الملكية
- الحكومة وأصحاب المصلحة في الصناعة ينبغي أن تحدد أنواع القرارات و العمليات التي لا ينبغي تفويضها أبدا لمثل هذه النظم واعتماد القواعد والمعايير التي تضمن السيطرة البشرية الفعالة على تلك القرارات، وتخصيص الأطر القانونية المسؤولة عن الأضرار التي تسببها.

الشفافية والحقوق الفردية

مع تطور الخوارزميات الذاتية والبيانات التي تمكن اتخاذ القرارات، يؤثر ذلك على المواطنين والمتطلبات القانونية في الدول، وبما في ذلك التأثيرات التالية:

- يجب أن يكون لدى الأطراف ومحاميهم والمحاكم طريقة معقولة للوصول إلى جميع البيانات والمعلومات التي تم إنشاؤها واستخدامها من قبل هذه الأنظمة من سلطات الدولة.
- المنطق والقواعد المضمنة في النظام يجب أن تكون متاحة للمشرفين عليها، إن أمكن، وأن تخضع خضوعا صارما لتقييم المخاطر

- ينبغي أن تولد الأنظمة مسارات تدقيق وتسجيل الوقائع ودعم القوانين، وينبغي أيضا أن تكون قابلة للتطبيق والتحقق من الجهات الخارجية
- يجب على عامة الناس معرفة من يتخذ القرارات الأخلاقية أو يدعم مثل هذه الأنظمة من خلال الاستثمار.

سياسات التعليم والتوعية

وتتناول السياسة الحماية وتعزيز السلامة، والخصوصية، والملكية الفكرية وحقوق الإنسان، وأمن المعلومات، وكذلك فهم الجمهور للتأثير المحتمل من الأنظمة التقنية الذكية والحكم الذاتي على المجتمع. للتأكد من أنها أفضل خدمة للمصلحة العامة، ينبغي للسياسات أن تكون:

- داعمة، معززة، وممكنه دوليا طبقا للمعايير القانونية المعترف بها

- تطوير الخبرات البشرية فيما يتعلق بالتقنيات

- تمكين قيادة البحث والتطوير

- تنظيم السلامة العامة

- تثقيف الجمهور حول الآثار الاجتماعية للتكنولوجيا

الأسس

الأخلاق الكلاسيكية

من خلال النظر إلى أكثر من ألفي سنة من تقاليد والأخلاق ، تستكشف مبادرة ال IEEE العالمية نظم أخلاقيات المهنة، والمعالجة العلمية والدينية بما في ذلك المعالجات الفلسفية العلمانية، لموائمة الأخلاق البشرية في العصر الرقمي. من خلال فلسفيات للمؤسسات التي تحدد الحكم الذاتي. وتتناول مبادرة IEEE العالمية المزايم حول القدرة على الحكم الذاتي للأنظمة التقنية، والتحكم في الأخلاق في النظم غير الأخلاقية، ويسأل عما إذا كانت القرارات التي اتخذتها غير أخلاقية وما إذا يمكن أن يكون للأنظمة عواقب غير أخلاقية.

مقاييس الرفاهية

بالنسبة للأنظمة الذكية والمؤتمتة يتم الاستناد إلى ذلك مؤشرات التقدم بشكل محدد لصالح الإنسانية، وتوجب إلى أن تكون المؤشرات واضحة لتلك الفائدة. وتشمل المقاييس نجاح الربح، والسلامة المهنية، والصحة المالية. في حين أن هذه المقاييس مهمة فشلت في أن تشمل مجموعة كاملة من الرفاهية، سواء بالنسبة للأفراد أو المجتمعات من الناحية النفسية، والاجتماعية، والبيئية. تضمن القياسات الأخذ بهذه العوامل، والنظر إلى المنافع الناشئة عن التقدم التكنولوجي والتي يجب أن يتم تقييمها بشكل أكثر شمولية، وتوفير فرصا لبحث السلبيات والعواقب التي يمكن أن تقلل من رفاهية الإنسان. وعلى العكس من ذلك، يمكن أن تساعد هذه المقاييس تحديد تأثير الأنظمة التقنية الذكية على زيادة رفاهية الإنسان، وتوفير الطرق إلى الابتكار المجتمعي والتكنولوجي.

تضمين القيم في أنظمة الحكم الذاتي

إذا كانت الآلات تشارك في المجتمعات البشرية كأنها شبه مستقلة، يجب تضمين القواعد في هذه الأنظمة والذي يتطلب ترسيم واضح للمجتمع التي ستنتشر فيها. وعلاوة على ذلك، حتى بداخل مجتمع معين، يتطلب أنواع مختلفة من النماذج الفنية المختلفة والقواعد. تكمن الخطوة الأولى في تحديد معايير المجتمع الذي ينبغي نشر النظم فيه، ثم على وجه الخصوص، القواعد ذات الصلة بأنواع المهام التي يقوم بها التصميم.

منهجيات الدليل الأخلاقي في البحث والتصميم

لإنشاء الأنظمة التقنية الذكية التي تهتم بتعزيز وتوسيع رفاهية الإنسان والحرية، تقوم منهجيات التصميم القائم على القيمة

وضع التقدم البشري في صميم تطوير النظم التقنية، مع الالمام بأن الآلات يجب أن تخدم البشر وليس العكس. لذا يجب على المطورين توظيف التصميم القائم على القيمة المنهجية من أجل خلق نظم مستدامة يمكن تقييمها من حيث التكاليف الاجتماعية والمزايا التي قد تزيد من القيمة الاقتصادية للمنظمات.

مخاوف التكنولوجيا في المستقبل إعادة ترسيم أطر الأسلحة ذاتية الحكم

صممت بعض الأنظمة المستقلة للتسبب في ضرر جسدي للبشرية، فيما يحمل أبعاداً أخلاقية إضافية بالمقارنة مع الأسلحة التقليدية. تلك النظم المستقلة لا تهدف أن تسبب الضرر. ويجب أن تشمل هذه الأنظمة على الأبعاد الأخلاقية، على الأقل، بما يلي:

- ضمان التحكم في منظومات تحكم الأسلحة
- تراعي تصاميم الأسلحة الآلية مراجعة الحسابات للمساعدة في ضمان المساءلة والسيطرة
- تضمين أنظمة التكيف والتعلم التي يمكن أن تفسر المنطق والقرارات للمشغلين بطريقة مفهومة
- تدريب المشغلين البشريين المسؤولين عن الأنظمة المستقلة
- تحقيق مكانية التنبؤ لمشغليها
- ضمان أن المصممين لهذه التكنولوجيات على دراية بالآثار الناجمة لعملهم
- تطوير مدونات أخلاقية مهنية تهتم بالتنمية التوعوية للأنظمة المستقلة

سلامة وأمن الذكاء العام الاصطناعي والذكاء الاصطناعي الخارق

على غرار التكنولوجيات القوية الأخرى، تتطور استخدامات الأنظمة الذكية مما ينطوي على مخاطر كبيرة، إما بسبب سوء استخدام أو سوء التصميم. ومع ذلك، وفقاً إلى بعض النظريات، تنهج النظم وتجاوز التصميم، في إطار غير متوقع أو غير مقصود فإن هذا السلوك يصبح خطراً على نحو متزايد ومن الصعب تصحيحه. من المرجح أنه ليست جميع الأنظمة يمكن أن تتماشى مع مستوى الإنسان. وعلى هذا النحو، ينبغي توخي الحذر لتحديد كيف يتم بناء الأنظمة حتى تصبح أكثر قدرة.

الحوسبة المؤثرة

التأثير هو جانب أساسي من الذكاء. محركات العواطف مثل الغضب والخوف والفرح في كثير من الأحيان وهذه هي أسس الأعمال طوال حياتنا.

لضمان أن الأنظمة التقنية الذكية سوف يمكن استخدامها لمساعدة البشرية إلى أقصى حد ممكن في جميع السياقات، والمشاركة في تيسير المجتمع البشري، لا ينبغي أن تسبب أضراراً إما عن طريق تضخيم أو تخميد الإنسان عواطف الإنسان. فيما أن بعض النظم تتأثر كيفية فهمها من قبل صانعي السياسات وعامة الناس.

واقع مختلط

والواقع المختلط يمكن أن يغير مفاهيمنا للهوية والواقع أن هذه التكنولوجيات تصبح أكثر شيوعاً في العمل، والتعليم، والحياة الاجتماعية، والمعاملات التجارية. القدرة الحقيقية هي إضفاء الطابع الشخصي على هذا العالم

المختلط، مما يشير المسائل الأخلاقية المتعلقة بحقوق الأفراد والتحكم في نظام واحد متعدد الأوجه والهوية، خاصة وأن التكنولوجيا تتحرك من تصميم سماعات الرأس إلى تكنولوجيات أكثر دقة وتكامل حسي.

عمليتنا

لضمان أهمية ثقافية و صرامة الفكرية في عملنا، تعتمد مبادرة IEEE العالمية على التدقيق الجماعي على الصعيد العالمي. .

أصدرنا تصميم متوافق أخلاقيا الإصدار 1 كما طلب الإدخال في ديسمبر 2016 وتلقى أكثر من مائتي صفحة من ردود الفعل حول المشروع. كطريقة لتسليط الضوء مستوحاة من ردود الفعل التي تلقيناها، كتبت سارة ماتينجلي - جوردن من مبادرة IEEE العالمية تقريره، "كيف تصبح رائد في الأخلاقيات العالمية".

نعلن عن نسخة تصميم متوافق أخلاقيا

EADv2 (EADv1) كطلب للإدخال مرة أخرى للحصول على مزيد من الأفكار حول الأصول الثمانية أقسام من EADv1 ، جنبا إلى جنب مع ردود أفعال فريدة من نوعها في EADv2.

مبادرة IEEE العالمية حاليا تشكل لجنة تنظيمية تتألف من ممثلين عن جميع لجاننا ومن الخبراء المستقلين والإلكترونيين

™ P7000 من مجموعات العمل للقيام بما يلي:

إعداد النسخة النهائية من تصميم متوافق أخلاقيا الذي سيصدر في عام 2019:

• وضع معايير للجان للتصويت على توصيات المرشحين لتصبح التوصيات مبنية على أساس العام وفقا لخطة العمل لمبادئ التصميم المتوافق أخلاقيا ضمن مبادرة IEEE العالمية. عملية التصويت هذه على أساس توافق

الآراء والبروتوكولات التي تقدمها IEEE SA

• إنشاء منهجية صارمة لإدراج ردود الفعل الواردة من EADv1 و EADv2 ، والعمل على النظر بشكل كلي للاعتبارات العالمية القائمة على التنوع لإدراج المحتوى.

• استخدام المسرد الذي قمنا بإنتاجه كمفتاح أداة لتجميع محتوى للنسخة النهائية من EAD ، لتوحيد شروط قدر الإمكان.

النسخة النهائية من التصميم المتوافق أخلاقيا

النسخة النهائية من تصميم المتوافق أخلاقيا سوف يكون كالتالي:

• كتيب. بينما يتم التنسيق المحدد الذي مازال قيد النظر، تكون النسخة النهائية من تصميم الانحياز أخلاقيا مميزة

"للتوصيات" (مقابل "الترشيحات") لجميع القضايا المستقبلية ، فيما صوت عليها أعضاء مبادرة IEEE العالمية. من المحتمل جدا أن يتم اصدار النسخة النهائية في أقسام وفقا للجان (كما في EADv1 و EADv2) وفقا للموضوعات و المبادئ التي يتعين البت فيها من قبل اللجنة التنظيمية المذكورة في الأعلى. في حين لا يتناول البيان موقف IEEE الرسمي ليتم استخدامها بسهولة من قبل التقنيين وصناع السياسات. الذين يركزون على الحكم الذاتي وتصميم الأنظمة الذكية.

• المواد التعليمية. و IEEE العالمية

ترغب المبادرة في تحويل الكتيب تصميم متوافق أخلاقيا إلى مرجع أكاديمي تعليمي كامل المواد. ليكون موجهة إلى الأكاديميين والمهندسين، ومن يبحث عن التوجه العالمي للتكنولوجيا لاستخدامها في الجامعة، أو ما بعد التخرج، أو غيرها من البيئات التعليمية حيث الأخلاق في التكنولوجيا أو القضايا ذات الصلة.

دمج ردود الفعل

في حين كان لدينا نية لقبول مباشرة أو مراجعة جميع ردود الفعل التي تلقيناها ل EADv1 . وتشمل بشكل شامل استقبال ردود الأفعال من EADv1 و EADv2 في العملية الشاملة التي أنشأناها في المسرد والتي تضمن زيادة أكثر للتمثيل العالمي والتنوع في عملنا.

المعجم

تلقينا قدرا كبيرا من ردود الفعل على توصيات اصدار تصميم متوافق أخلاقيا . مما أدى ذلك في نهاية المطاف، لأنشاء لجنة للمسرد والتي أطلقت المسودة الأولى من المسرد لدينا في نفس الوقت أصدرانا، EADv2. هدفنا هو تنقيح المسرد لدينا بحيث يتم بحلول منتصف عام 2018، استنادا إلى ردود الفعل المجمع لجميع الأقسام من EAD (الإصدارات 1 و 2)، يمكننا توحيد التعاريف التي تعكس المجموعة العالمية من التعاريف التي ستنفذها جميع اللجان في النسخة النهائية من الاصدار.

لمزيد من التمثيل العالمي

تلقينا قدرا كبيرا من ردود الفعل مشيرا أن EADv1 كان "غديبا" إلى حد ما في ثقافتها. وهذا منطقي، لأن ال 100 الأولي

من الأعضاء الذين يعملون في عدد EADv1 كبير منهم من أمريكا الشمالية، والاتحاد الأوروبي. منذ اصدار EADv1، ومع ذلك، لدينا أعضاء من الصين، كوريا، اليابان، والبرازيل، والمكسيك، والاتحاد الروسي، وإيران، تايلاند، وإسرائيل جنبا إلى جنب مع أشخاص جدد من الولايات المتحدة والاتحاد الأوروبي. بالإضافة إلى 250 عضوا في المبادرة، وهناك أيضا أكثر من الآن 400 عضو عالمي في IEEE P7000™ من مجموعات العمل.

• دعم الأعضاء لترجمة الملخص التنفيذي ل EADv1 إلى لغات متعددة.

• إضافة "الأخلاق الكلاسيكية في A/IS إس" إلى اللجنة.

• إنشاء تقرير كيف تصبح رائدا الأخلاقيات العالمية

• تكليف وضع تقرير من أحدث من لدينا من الأعضاء العالميين حول حالة A/IS في مناطقهم.

• إنشاء لجنة التوعية للمساعدة وتحديد ودمج العمل الجاري في A / IS الأخلاق من قبل المرأة، الأقليات

والطلاب، والمجموعات الأخرى التي تمثل الطيف الكامل للمجتمع الذي نأمل للتأثير إيجابيا مع عملنا. نحن

نعمل حاليا مع الأعضاء لإعادة تشغيل ، AI4ALL من القادة ضمن IEEE لمساعدتنا على ضمان أن IEEE

المبادرة العالمية والنسخة النهائية من منحاز الانحياز أخلاقيا أن يكون ممثلا بشكل كلي وذات صلة قدر الإمكان.

تحديث المصطلحات

ليست هناك حاجة لاستخدام مصطلح *ذكاء اصطناعي* من أجل التحدث عن التكنولوجيات والنظم التي تهدف إلى تمديد الذكاء البشري لدينا أو استخدامها في تطبيقات الروبوتات. لهذا السبب، نستخدم مصطلح: أنظمة مستقلة وذكية في سياق عملنا. اخترنا لاستخدام هذه العبارة أن تشمل كثير من المجالات (التعلم الآلي، وأنظمة ذكية والهندسة، والروبوتات، وما إلى ذلك) في جميع أنحاء التصميم، يضمن الإصدار 2 تطبيق أوسع للاعتبارات الأخلاقية في تصميم هذه التكنولوجيات.

كيف تم إعداد المستند

وقد أعدت هذه الوثيقة باستخدام برنامج متاح المصدر، وعن طريق التعاون، وبناء توافق الآراء، في أعقاب عمليات صناعة برنامج الاتصالات، وهو برنامج من جمعية معايير IEEE. اتصالات الصناعة حتى يسهل التعاون بين المنظمات والأفراد لصقل تفكيرهم لقضايا التكنولوجيا، مما يساعد على احتضان إمكانيات أنشطة المعايير الجديدة والمعايير ذات الصلة للمنتجات والخدمات.

كيفية الاستشهاد في تصميم منحا أخلاقيا

يرجى ذكر الإصدار 2 من تصميم منحا أخلاقيا

على النحو التالي:

مبادرة IEEE العالمية حول أخلاق أنظمة الحكم الذاتي والذكاء. أخلاقيات التصميم: رؤية لأولويات رفاهية الإنسان مع أنظمة التحكم الذكية والذاتية ، الإصدار 2. IEEE ، 2017.

http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html.

تقديرنا

نود أن نشكر لجنتنا التنفيذية ورئيس المبادرة العالمية IEEE:

موظفو اللجنة التنفيذية

رجا شاتيل، الرئيس

كاي فيرث - بوتيرفيلد، نائب الرئيس

جون هافنز، المدير التنفيذي

أعضاء اللجنة التنفيذية

د. جريج أدامسون، رونالد سي أركين، فيرجينيا ديغنون، دانيت غال، فيليب هول، مالا فيكا جايارام، سفين كوينيغ، راج مادافان، ريتشارد مالا، هاجيت ميسر يارون، أجونج مون، مونيك مورو، فرانسيسكا روسي، آلان وينفيلد

رؤساء اللجان

- المبادئ العامة: آلان وينفيلد ومارك هالفيرسون
- تضمين القيم في الحكم الذاتي للأنظمة الذكية: فرانسيسكا روسي و بيرتر مال
- منهجيات دليل البحوث الأخلاقية والتصميم: رجاء شاتيل وكورين جن كاثر
- السلامة والأمن للنظم الذكاء الاصطناعي والذكاء الاصطناعي الخارق: مالو بورغون و ريتشارد مالا
- البيانات الشخصية والولوج الفردي: كاترينا داو و ج هافينس
- إعادة تأطير أنظمة الأسلحة المستقلة: بيتر أسارو
- الاقتصاد / القضايا الإنسانية: كاي فيرث-بوتيرفيلد وراج مادافان
- القانون: كاي فيرث - بوتيرفيلد وديريك جينكس
- الحوسبة المؤثرة: رونالد أركين و جوانا ج. برايسون
- الأخلاق الكلاسيكية في A/IS: جاريد بيلبي
- السياسات: كاي فيرث - بوتيرفيلد وفليب هول
- الواقع المختلط: مونيك مورو وجاي يوريو
- الرفاهية: لورا موسيكانسكي وجون سي.
- الصياغة: كاي فيرث - بتيرفيلد وديفن ديساي
- الصناعة: فرجينيا ديغنون ومالفيكا جايارام
- الاتصالات: لاني سيتو ومارك هالفيرسون
- المسرد: سارة M. الأردن
- التوعية: دانيت غال

تقرير

نود أن نعرب عن تقديرنا للتقارير والمنظمات والأفراد الذين ساهموا بالبحوث والأفكار المساعدة لزيادة الوعي حول القضايا الأخلاقية في عالم ذكي ومستقل بما في ذلك (ولكن ليس على سبيل الحصر، وبدون ترتيب معين معهد المستقبل للحياة أسيلومار- منظمة العفو الدولية مبادئ ، تقرير منظمة العفو الدولية 2017.AI تقرير حقوق الانسان في عصر الروبوت ل راثيناو إنستيتوت ، تقرير كوميسست في أخلاق الروبوتات من اليونسكو ، توصيات البرلمان الأوروبي إلى لجنة قواعد القانون المدني بشأن الروبوتات ، عواقب الذكاء الاصطناعي على السوق، الإنتاج، الاستهلاك، التوظيف وتقرير مجتمع الاقتصاد الأوروبي واللجنة الاجتماعية (المقرر Catelijne : مولر)

تقرير OECD ،

التحول الرقمي: عمل التحول من أجل النمو

، بيان USACM على خوارزمية الشفافية والمساءلة ،

دليل إلى التصميم الأخلاقي وتطبيق الروبوتات (معهد المعايير البريطانية)،

قواعد اليابان الأساسية للبحث -منظمة العفو الدولية، Éthique،

دليل لبحوث (ROBOTIQUE)،

[Charta der Digitalen Grundrechte der Europäischen Union \(Charter of the Digital](#)

[Fundamental Rights of the European Union\)](#), Telecommunications Research Laboratory,

“AI Network Kent kai

Kaigi H kokusho 2016: AI Network no Eiky to Risk — Chiren Shakai (WINS) no Jitsugen

ni Muketa Kadai” (AIネットワーク化検討会議 報告書2016 公表- 「AIネットワーク

化の影響とリスク-智連社会(WINS(ウインス))の実現に向けた課題-

المؤتمر المعني بالشبكات بين تقرير المؤسسات الصناعية (2016): الآثار والمخاطر من قضايا شبكات الذكاء

الاصطناعي من أجل تحقيق الحكمة ،

اليابانية وزارة الداخلية والاتصالات، مجلس صناعة تكنولوجيا المعلومات WINS

مبادئ سياسة منظمة العفو الدولية ،

الذكاء الاصطناعي -فرصة السياسة العامة للذكاء الاصطناعي intel

IEEE. بيان موقف مبادرة السياسة العامة الأوروبية

الذكاء الاصطناعي: دعوة صناع السياسة أن تأخذ دورا قياديا في تحديد المدى الطويل
استراتيجية منظمة العفو الدولية ، بيان موقف USA-IEEE على
بحوث الذكاء الاصطناعي، التنمية و التنظيم ، أولوية مبادرة IEEE العالمية
رفاهية الإنسان في عصر الذكاء الاصطناعي

المنظمات

جمعية النهوض بالذكاء الاصطناعي عملهم التكويني حول الأخلاقيات
معهد مستقبل الحياة ، الشراكة بالذكاء الاصطناعي لصالح الناس والمجتمع ،
مؤسسة الروبوتات المسؤولة ، الذكاء الاصطناعي والمجتمع ،
معهد بحوث الاستخبارات الالية ،
مركز أخلاقيات المعلومات الدولي
، المركز الأفريقي لتميز أخلاقيات المعلومات ،
TU4 مركز الأخلاقيات والتكنولوجيا ،
مركز دراسة المخاطر الموجودة ،
مركز ليفيرهولم لمستقبل الاستخبارات ،
معهد مستقبل الإنسانية ،
والجمعية اليابانية للذكاء الاصطناعي ،
جمعية الآلات الحوسبية ،
الدعوة المستقبلية ،
ACM مجموعة الاهتمام الخاص بالذكاء الاصطناعي ،
المستقبل العالمي بمنتدى الاقتصادي العالمي
مجلس الذكاء الاصطناعي والروبوتات ،
مركز آسيا الرقمي ،
مبادرة منظمة الذكاء الاصطناعي،
معهد الروبوتات المفتوح،
ومركز الدالاي لاما لأخلاق والقيم التحويلية في معهد ماساتشوستس للتكنولوجيا ،
مبادرة الأخلاقيات في معهد ماساتشوستس للتكنولوجيا ميديا لاب ، و-IEEE
مجلس العلاقات الحكومية الأمريكية الاصطناعي
لجنة الاستخبارات ،
و IEEE للروبوتات
ولجنة أتمتة جمعية أخلاقيات الروبوت ،

ومجتمع IEEE للروبوتات والأتمتة
، مجتمع IEEE حول الآثار الاجتماعية والتكنولوجيا ،
مجتمع IEEE الكمبيوتر،
مجتمع الذكاء الحسائي IEEE،
و IEEE للنظم،
جمعية الانسان والرقميات، و
IEEE.
مبادرة النظم الذاتية التكافلية

الأفراد

نود أن نشكر القيادة والدعم المستمر لمبادرة IEE العالمية الدكتور كونستانتينوس كاراشاليوس، مدير جمعية معايير IEEE و عضو مجلس إدارة IEEE. ونود أن نشكر خصوصا ايلين .لاش، IEEE المستشار العام ورئيس الامتثال التي استثمرت وقتها في مراجعة وافية لهذه الوثيقة بأكملها، مع القناعة أن هناك حاجة ملحة لتركيز المجتمع الدولي على إبراز الاعتبارات الأخلاقية في تطوير أنظمة الحكم الذاتي والذكية..شكر خاص للدكتور بيتر بروكس للمساهمات في نظرة عامة على EADv2. شكرا لدينا الأعضاء وفريق IEEE هذا العمل يرجع إلى الخبراء المتطوعين العاملين في اللجان لدينا و IEEE P7000™ مجموعات العمل، جنبا إلى جنب مع IEEE الموظفين الذين يدعمون جهودنا. شكرا لك على تفانيكم نحو تحديد وتصميم وإلهام المبادئ الأخلاقية والمعايير التي من شأنها أن تضمن أن الأنظمة الذكية والمستقلة والتقنيات المرتبطة بها وبذلك تستفيد الانسانية بشكل إيجابي.

تنويه: هذه الوثيقة لا تمثل موقف أو آراء IEEE ولكنها تمثل الآراء المستنيرة من أعضاء اللجنة التي تقدم رؤى مصممة لتوجيه الخبراء بشأن A/IS. هذه الترجمة متوفرة بسهولة التقديم. إصدار اللغة الإنجليزية من هذه الوثيقة هو النسخة الأصلية والرسمية منها. في حالة وجود أي تعارض بين النسخة الإنجليزية والنسخة المترجمة ، فإن النسخة الأصلية من هذه الوثيقة (وضعت باللغة الإنجليزية والمتاحة هنا) هي النسخة الأصلية.