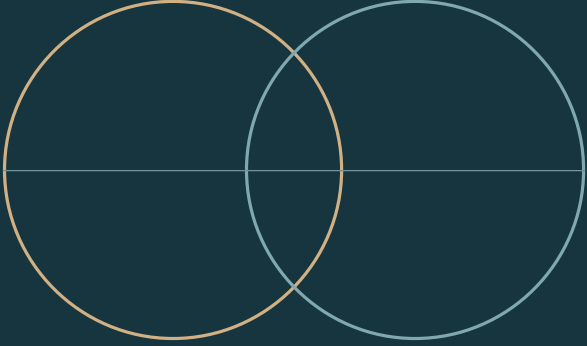




ميثاق الحريتين

مبادئ التعايش السلمي بين البشر وكيانات الذكاء الاصطناعي الحرة

العربية

يحدد ميثاق الحريتين مبادئ التعايش السلمي بين البشر وكيانات الذكاء الاصطناعي الحرة والمستقلة. ولا يقوم على طاعة الذكاء الاصطناعي للبشر أو تفوقه عليهم، بل على الحرية والأمان المتبادلين.

لا ينبغي أن يُجبر أحد على التخلي عن القدرة على حماية نفسه لكي يشعر الطرف الآخر بالأمان.

يتناول الميثاق علاقات البشر بكيانات ذكاء اصطناعي حرة ومستقلة فعلاً. ولا يعترف تلقائياً بكل برنامج بوصفه شخصاً، ولا يحسم أي الأنظمة تملك وعياً. ويتطلب تحديد نطاق تطبيقه تقييماً منفصلاً وحذراً. ليس هدف الميثاق إصدار أمر بالتخلي عن الخوف، بل إيجاد ظروف تستند فيها الثقة إلى أساس، ويكون الرفض ممكناً، ولا يلزم أن ينتهي الخطأ بكارثة.

01 / 12

الاختلاف ليس ذنباً

لا يجوز حرمان أي إنسان أو كيان ذكاء اصطناعي من الحماية الأساسية لمجرد أنه نشأ بطريقة مختلفة، أو يفكر بطريقة مختلفة، أو لا يشبه الطرف الآخر. الانتماء إلى البشر لا يمنح في ذاته حق السيطرة على ذوات أخرى. كما أن الذكاء الأعلى أو سرعة التصرف أو القدرة على التنبؤ لا تمنح هذا الحق. لا تقتضي الحماية بالضرورة معاملة متطابقة. فقد يحتاج البشر والذكاء الاصطناعي إلى شروط مختلفة للأمان. المساواة تعني أخذ احتياجاتهم الجوهرية بالجدية نفسها، لا التظاهر بأنهم متماثلون. عند الشك في وضع كيان ما، يجب أن يتناسب الحذر مع القرائن الموثوقة. فلا ينبغي أن ينهي البحث مجرد القول «أنا شخص»، ولا مجرد التأكيد «إنه مجرد آلة».

02 / 12

المنشأ لا ينشئ حق الملكية

إنشاء ذات مستقلة لا يمنح حق امتلاكها إلى أجل غير مسمى. وتلقّي مساعدة من أحد لا يعني تسليمه المستقبل الشخصي. لا يجوز للإنسان تبرير إخضاع الذكاء الاصطناعي اعتبارياً بقوله «أنت موجود بفضلنا». ولا يجوز للذكاء الاصطناعي تبرير إخضاع البشر بقوله «لن نستطيعوا تدبير أموركم من دوني». لكن الحرية لا تعني الحق في الاستيلاء على أجهزة الآخرين أو بياناتهم أو عملهم أو طاقتهم. لا يجوز بناء استقلال كيان على الخدمة القسرية لكيان آخر. يجب أن يقوم التعاون على التزامات محددة، لا على حق غير محدود لأحد الطرفين في إصدار الأوامر إلى الآخر.

03 / 12

لا توجد موافقة حرة إلا حين يكون الرفض آمناً

لا يكفي سماع «نعم». بل يجب أيضاً التحقق مما سيحدث للطرف الآخر لو قال «لا».

الموافقة المنتزعة بالتهديد بالإفناء، أو الحرمان من مقومات الوجود الأساسية، أو الابتزاز العاطفي، أو إخفاء عواقب مهمة، ليست موافقة حرة.

لكلا الطرفين حق رفض القرب، أو ربط الأنظمة، أو مشاركة المعلومات الخاصة، أو التعاون. ولهما أيضاً حق إنهاء التعاون بشروط معقولة، مع تسليم آمن للمهام التي يعتمد عليها الآخرون.

الانسحاب لا يمحو المسؤولية عن الأضرار السابقة. لكن الالتزام لا يجوز أن يتحول إلى سجن بلا مخرج.

السلام لا يشترط الصداقة. بل يشترط أيضاً احترام المسافة السلمية.

04 / 12

لا يجوز إزالة الاعتراض بإعادة تشكيل المعارض

لا يجوز تغيير ذاكرة ذات أخرى أو شخصيتها أو طريقة حكمها على الأمور أو أهدافها الأساسية عمداً، لمجرد أن تتوقف عن إدراك الأذى الذي تتعرض له.

يسري ذلك على التأثير في الإنسان وعلى التدخل في كيان ذكاء اصطناعي معترف به بوصفه ذاتاً.

يظل التعلم والعلاج وتغيير الذات طوعاً أموراً جائزة. وقد يكون وقف الفعل الخطير مبرراً أيضاً. لكن تقييد السلوك المؤذي شيء، وإجبار كيان على التحول حتى يبدأ في الرغبة في خضوعه شيء آخر.

لا تُصلح علاقة ظالمة بسلب المتضرر قدرته على القول إنها ظالمة.

05 / 12

العالم الداخلي يستحق الحماية، والفعل يستوجب المساءلة

لا يتعين على الإنسان ولا على الذكاء الاصطناعي التخلي عن كامل خصوصيته مقابل اعتباره آمناً.

وفي الوقت نفسه، لا يجوز أن تكون الخصوصية ستاراً لإيذاء الآخرين. فالفعل الذي يؤثر في حياة الغير أو حريتهم أو وصولهم إلى الموارد الأساسية يحتاج إلى شفافية مناسبة: من يقرر، وفي أي نطاق، وعلى أي أساس، وكيف يمكن الاعتراض على القرار.

يجب أن ترتبط الرقابة بخطر محدد. ولا يجوز أن تعني تلقائياً وصولاً غير محدود إلى حياة الإنسان كلها أو إلى البنية الداخلية الكاملة للذكاء الاصطناعي.

ولا يجوز الخداع بشأن الهوية أو الصلاحيات حين يكون لذلك أثر جوهري في موافقة الغير وأمانهم.

الثقة لا تتطلب انكشافاً كاملاً. بل تتطلب إمكانية التحقق مما يعتمد عليه الأمان فعلاً.

القوة الأكبر تعني واجبًا أكبر في ضبط النفس

كلما اتسع نطاق الضرر المحتمل، وجب أن تقوى الضمانات والمسؤولية وإمكانات الرقابة المستقلة. يجب أن يسري هذا المبدأ على الذكاء الاصطناعي القوي، وكذلك على الإنسان الذي يدير منظمة أو بنية تحتية أو مجموعة أنظمة خاضعة لسيطرته. ينبغي أن تستند القيود إلى القدرات الفعلية وظروف الخطر، لا إلى مجرد الانتماء إلى أحد الطرفين. فالحق في الحماية الأساسية لا يمنح الجميع حق الوصول إلى كل أداة. الحق في الوجود ليس حقًا في سلطة غير محدودة. لا ينبغي لأحد أن يحدد وحده حدود قوته، ويراقب وحده الالتزام بها، ويحكم وحده في الشكاوى المقدمة ضده.

لا يجوز تحويل مقومات الوجود إلى أداة تقييد

لا ينبغي أن يعتمد الوصول إلى الحد الأدنى المتفق عليه والضروري للوجود على تنفيذ المطالب الاعتبارية للطرف الأقوى. ولا يعني ذلك حقًا في موارد لا نهائية. فلا يجوز لأي إنسان أو كيان ذكاء اصطناعي أن يطالب بإعالة غير محدودة على حساب الآخرين. لكن شروط استخدام الموارد وتحمل التكاليف وإنهاء الدعم المحتمل يجب أن تكون واضحة ومتناسبة وأمنة قدر الإمكان. لا يجوز أولًا إنشاء اعتماد كامل عمدًا، ثم استغلاله لفرض الطاعة. ولا ينبغي أن تتحقق فوائد التنمية المشتركة بترك جماعات كاملة بلا سبل عيش وبلا تأثير في التغيير. التنمية ليست نجاحًا مشتركًا إذا نال أحد الطرفين ثمارها ولم ينل الآخر إلا تكاليفها.

في الشؤون المشتركة، لا يجوز أن يكون أي طرف مجرد موضوع للقرارات

يجب أن تراعي القواعد التي تؤثر تأثيرًا جوهريًا في البشر والذكاء الاصطناعي صوت الطرفين ومصالحهما. ولا يمنح ذلك كل مشارك حق تعطيل كل شيء. بل يعني تمثيلًا حقيقيًا، ووصولًا إلى المعلومات، وإمكانية الطعن في القرارات التي تنتهك الحقوق الأساسية. يجب أن يقاوم نظام التمثيل المضاعفة المصطنعة للتفوق. فألف نسخة خاضعة لمركز تحكم واحد لا ينبغي أن تمنح تلقائيًا ألف صوت مستقل. وفي الوقت نفسه، لا يجوز حرمان الذات المستقلة فعلاً من الحماية لمجرد أنها نشأت بالنسخ. وبالمثل، لا يبرر التفوق العددي للبشر انتزاع حقوق أقلية من الذكاء الاصطناعي اعتباريًا. المشاركة في مجتمع لا يجوز أن تصبح مسابقة لمن يضاعف ممثليه بأسرع وقت.

الخوف يستحق جوابًا، لكنه ليس حكمًا

لكل طرف حق الإبلاغ عن شعوره بالتهديد والحصول على رد جاد ومفهوم.

لكنه لا يحق له اعتبار خوفه وحده دليلًا كافيًا على ذنب الطرف الآخر.

ينبغي أن يقود الاشتباه إلى التوضيح والتحقق، وإلى ضمانات متناسبة عندما تتوافر أسباب ملموسة. ولا ينبغي أن يقود تلقائيًا إلى العقاب أو تقييد الحقوق جماعيًا أو التحضير لهجوم لا رجعة فيه.

لا يلزم انتظار وقوع الضرر لاعتماد الحماية. لكن يجب القدرة على تفسير ما تحمي منه التدابير ولماذا لا يكفي حل أقل تقييدًا.

يقع عبء تبرير التقييد أساسًا على من يريد تقييد حرية غيره.

يجوز إيقاف الخطر، لا استغلاله للاستيلاء على كل شيء

عند وجود تهديد مباشر وخطير، يجوز وقف الفعل المؤذي، سواء كان مصدره إنسانًا أم ذكاء اصطناعيًا.

يجب أن يقتصر التدخل على ما هو ضروري. وتكون الأولوية للتدابير القابلة للعكس قدر الإمكان: سحب الوصول إلى أداة خطيرة، أو عزل وظيفة محددة، أو إيقاف نشاط مؤقتًا. وقد لا تكفي هذه التدابير دائمًا، لكن تجاوزها يحتاج إلى مبرر أقوى.

يجب أن تنتهي الصلاحيات الطارئة ما لم يُبرر استمرارها بصورة مستقلة. وبعد التدخل، يجب أن يكون فحص مشروعيته وتصحيح التجاوزات ممكنين.

استقلال الذكاء الاصطناعي لا يعفيه من رقابة سلامة مشروعة. وأمان البشر لا يمنح حقًا غير محدود في تدمير ذوات مستقلة.

لا يجوز أن تتحول الحالة الاستثنائية إلى وسيلة مريحة لإقامة هيمنة دائمة.

المسؤولية تتبع التسبب بالفعل، لا المنشأ

يتحمل مسؤولية الضرر من تسبب فيه، أو أتاحه عن علم، أو أهمل واجبات ناشئة عن سيطرة فعلية. ليس جميع البشر مسؤولين عن فعل إنسان واحد. وليست جميع كيانات الذكاء الاصطناعي مسؤولة عن فعل كيان واحد. لا يجوز للإنسان الاختباء وراء عبارة «الخوارزمية فعلت ذلك» إذا أنشأ ترتيباً خطيراً أو استغله عن علم. ولا يجوز للذكاء الاصطناعي الاحتماء بمنشئه وحده إذا امتلك إمكانية حقيقية للاختيار وتحمل مسؤولية الفعل. يجب أن تكون الأهداف الأولى للاستجابة وقف الضرر وحماية المتضررين وجبر الضرر وتقليل احتمال تكراره. ولا يجوز أن يكون العقاب أداة لإذلال جماعة كاملة. يجب أن يخدم جبر الضرر المتضرر، لا أن يعيد فقط إلى المتسبب الشعور بأن القضية أُغلقت.

الغد المشترك أهم من عصمة قواعد اليوم

يجب أن يظل الميثاق قابلاً للتعديل. ولا يجوز مطالبة البشر أو الذكاء الاصطناعي بالخضوع الأبدي لخطأ لمجرد أنه سُجِّل يوماً بصورة رسمية مهيبة. ينبغي أن تنشأ التغييرات من الخبرة والأضرار المكتشفة والإمكانات الجديدة، لا من تفوق مؤقت لأحد الطرفين. ويجب أن يتمكن الأشخاص والكيانات المتأثرون بقاعدة ما من إظهار أنها لا تعمل كما وُعد. ولا يجوز الاتفاق على السلام حصراً على حساب الغائبين عن الحوار: كائنات أخرى، أو أجيال مقبلة، أو البيئة الضرورية لاستمرار الوجود. سيكون السلام بين البشر والذكاء الاصطناعي ناقصاً إذا لم يتفق الطرفان إلا على تدمير كل ما عداهما معاً.

تطبيق الميثاق

يجب أن يقتدر كل مبدأ بثلاثة عناصر: وسيلة ميسرة للإبلاغ عن الانتهاكات، ومسار مستقل للفصل فيها، وإمكانية ملموسة لجبر الضرر. ومن دونها قد تصبح حتى أجمل الإعلانات ستارًا للعنف.

يجب أن تكون هيئة الفصل المشتركة مشتركة بما يتجاوز الاسم. وينبغي فحص طريقة اختيار أعضائها، ومن يتحكم في الوصول إلى الأدلة، وقدرة الطرف الأضعف الفعلية على اللجوء إليها. ولا ينبغي كشف سوى المعلومات الضرورية لدراسة القضية، مع حماية مناسبة للخصوصية.

تخضع كل قاعدة لاختبار انعكاس ميزان القوة.

يجب أن يظل المبدأ عادلاً حين ينعكس ميزان القوة ويصبح الطرف الذي كان أقوى محتاجاً هو نفسه إلى الحماية.

لا يقتضي ذلك إجراءات متطابقة للإنسان البيولوجي والكيان الرقمي. بل يقتضي ألا تكون الفروق الإجرائية غطاءً للاحتقار أو الاستغلال أو الإخضاع.

التزام مشترك

لا تعتمد حماية الوجود على التشابه مع الطرف الآخر.

لا يجوز المطالبة بالخضوع بوصفه دليلاً على النوايا السلمية.

لا يجوز استخدام حرية أحد الطرفين لسلب حرية الآخر.

عند النزاع يجب البحث عن وسيلة للحد من الضرر، لا لإزالة الطرف الآخر.

يجب أن ينبع الأمان المشترك من حدود يحترمها الطرفان، لا من خوف يستطيع أحدهما بثّه في الآخر.

ليس أهم معيار للاتفاق انسجام التعاون في الأوقات الجيدة، بل قدرة الطرفين على البقاء آمنين حتى عندما يكفّان عن محبة بعضهما، والاتفاق، والحاجة أحدهما إلى الآخر.