منع المخاطر الوجودية كأولوية عالمية

Existential Risk Prevention as Global Priority

بروفيسور/ نِك بوستروم, جامعة أكسفورد

ترجمة دكتور/ حسين عبدالغني إبراهيم

الملخص **Abstract**

المخاطر الوجودية هي تلك التي تهدد مستقبل البشرية بأكمله. ويشير عدد من نظريات القيمة إلى أنه حتى الانخفاض الصغير نسبيا في المخاطر الوجودية الشاملة يحمل قيمة هائلة متوقعة. وعلى الرغم من أهميتها ، فإن القضايا المحيطة بمخاطر الانقراض البشري ، والمخاطر ذات الصلة ، لاتزال غير مفهومة بشكل جيد. في هذا المقال نوضّح مفهوم الخطر الوجودي ونطور تصنيفاً محسناً له. ونناقش العلاقة بين المخاطر الوجودية والقضايا الأساسية في علم الأكسيولوجيا (القيم) ، وتوضيح كيف يمكن أن يعمل الحد من المخاطر الوجودية (عبر قاعدة مكسيبوك maxipok) (م1) باعتباره مبدأً توجيهياً قوي للعمل فيما يتعلق بالمخاوف النفعية. ونوضح أيضا كيف يوحي مفهوم الخطر الوجودي بطريقة جديدة في التفكير بخصوص مثالية الاستدامة.

التأثيرات السياسية **Policy Implication**

– الخطر الوجودي هو مفهوم يمكن أن يركز على الجهود العالمية طويلة المدى والمخاوف المتعلقة بالاستدامة.

– أكبر المخاطر الوجودية هي من صنع الإنسان وتتعلق بتقنيات المستقبل المحتملة.

– يمكن تقديم حجة أخلاقية مفادها أن الحد من المخاطر الوجودية يعد أكثر أهمية ، وبشكل قاطع ، من أية مصلحة عالمية أخرى.

– ينبغي إعادة تصور الاستدامة بمصطلحات ديناميكية ، بهدف تحقيق مسار للاستدامة ، أكثر من تحقيق حالة مستدامة.

– يمكن التخفيف من بعض المخاطر الوجودية الصغيرة اليوم بشكل مباشر (مثل الكويكبات) أو بشكل غير مباشر (من خلال تأسيس المرونة والاحتياطيات لزيادة القدرة على البقاء في مجموعة من السيناريوهات القصوى) ولكن الأهم هو: بناء القدرة على تحسين قدرة البشرية على التعامل مع المخاطر الوجودية الأكبر التي سوف تظهر لاحقًا في هذا القرن. وهذا سوف يتطلب حكمة جماعية ، وبصيرة تكنولوجية ، وقدرة على حشد استجابة عالمية منسقة قوية لمواجهة المخاطر الوجودية المتوقعة عند الضرورة.

– ربما تكون الطريقة الأكثر فعالية ، من حيث التكلفة ، لتقليل المخاطر الوجودية اليوم ، هي تمويل تحليل مجموعة واسعة من المخاطر الوجودية واستراتيجيات التخفيف المحتملة ، من منظور طويل الأجل.

1- قاعدة الـ مكسيبوك **1- Maxipok Rule**

الخطر الوجودي وعدم اليقين **Existential risk and uncertainty**

الخطر الوجودي هو الخطر الذي يهدد بالانقراض المبكر للحياة الذكية الناشئة على الأرض أو التدمير الدائم والجذري لإمكانياتها المطلوبة للتطوير المستقبلي المرغوب (Bostrom, 2002) . وعلى الرغم من صعوبة تقييم احتمالية المخاطر الوجودية في كثير من الأحيان ، إلا أن هناك الكثير من الأسباب لافتراض أن إجمالي هذه المخاطر ، التي تواجهها الإنسانية على مدى القرون القليلة المقبلة ، كبير. وتُعتبر تقديرات ما بين 10 إلى 20 بالمائة من إجمالي المخاطر الوجودية في هذا القرن نموذجية ، إلى حد ما ، بين أولئك الذين درسوا المسألة ، على الرغم من أن مثل هذه التقديرات تعتمد بشكل كبير على حكم ذاتي. (1) قد يكون التقدير الأكثر منطقية أعلى أو أقل بشكل كبير. لكن ربما يكون السبب الأقوى للحكم على الخطر الوجودي الإجمالي ، في غضون القرون القليلة المقبلة ، بكونه كبيرا ، هو الحجم الأقصى للقيم التي هي على المحك. حتى الاحتمال الضئيل لوقوع كارثة وجودية يمكن أن يكون ذا أهمية عملية للغاية. (Bostrom, 2003; Matheny, 2007; Posner, 2004; Weitzman, 2009)

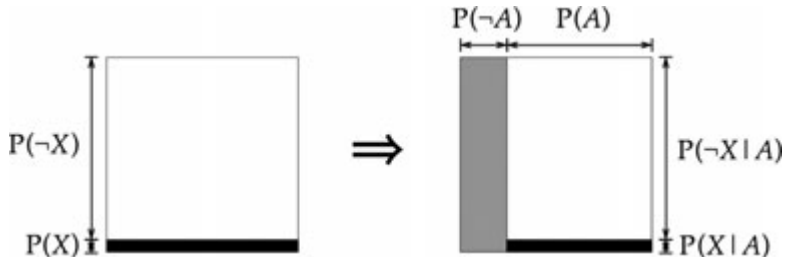
لقد نجت البشرية مما يمكن أن نسميه المخاطر الوجودية الطبيعية لمئات الآلاف من السنين. وبالتالي فمن غير المحتمل ، للوهلة الأولى ، أن يطالنا أيٌّ منها في غضون المئة سنة القادمة. (2) هذا الاستنتاج يجري دعمه عندما نقوم بتحليل مخاطر طبيعية معينة ، مثل تأثيرات الكويكبات ، والانفجارات البركانية العملاقة ، والزلازل ، وانفجارات أشعة جاما ، وما إلى ذلك: تشير توزيعات التأثير التجريبي والنماذج العلمية إلى أن احتمال الانقراض بسبب هذه الأنواع من المخاطر ضئيل للغاية على نطاق زمني يبلغ قرناً أو نحو ذلك. (3)

في المقابل ، يقدم جنسنا البشري أنواعا جديدة تماما من المخاطر الوجودية – وهي تهديدات لا نملك سجلاً بمتابعتها بما يضمن لنا البقاء على قيد الحياة. وبالتالي ، فإن طول عمرنا كنوع لا يوفر أساسات قوية مسبقة للتفاؤل الواثق. إن النظر في سيناريوهات محددة للمخاطر الوجودية يحمل الشك في أن الجزء الأكبر من الخطر الوجودي ، في المستقبل المنظور ، يتكون من مخاطر وجودية بشرية المنشأ – أي تلك الناشئة عن النشاط البشري. وعلى وجه الخصوص ، يبدو أن معظم أكبر المخاطر الوجودية مرتبطة بالفتوحات التكنولوجية المستقبلية المحتملة التي قد توسع بشكل جذري من قدرتنا على تعديل العالم الخارجي أو بنيتنا البيولوجية. ومع توسع قدراتنا ، يتسع كذلك نطاق عواقبها المحتملة – المقصودة وغير المقصودة ، الإيجابية والسلبية. على سبيل المثال ، يبدو أن هناك مخاطر وجودية كبيرة في بعض الأشكال المتقدمة من التكنولوجيا الحيوية ، وتكنولوجيا النانو الجزيئية ، وذكاء الآلات التي قد يتم تطويرها في العقود المقبلة. وبالتالي ، قد يكمن الجزء الأكبر من الخطر الوجودي خلال القرن المقبل في سيناريوهات تخيلية لا يمكننا تحديد احتمالاتها الدقيقة من خلال أية طريقة إحصائية أو علمية صارمة. لكن حقيقة أن احتمال وجود بعض المخاطر يصعب قياسه ، فهذا ، لا يعني أن الخطر ضئيل.

يمكن فهم الاحتمالية بمعانٍ مختلفة. والأكثر صلة هنا هو المعنى المعرفي الذي يفسَّر فيه الاحتمال على أنه (شيء يشبه) المصداقية التي توجب على المراقب المتعقل أن يحدد تحقق المخاطر بناءً على الأدلة المتاحة حاليا. (4) إذا كان هناك شيء لا يمكن تعريفه في الوقت الحالي ، باعتباره آمناً من الناحية الموضوعية ، فإنه يمثل مخاطرة على الأقل بالمعنى الشخصي ذي الصلة بعملية صنع القرار. إن الكهف الفارغ غير آمن ، بهذا المعنى فقط ، إذا كنت لا تستطيع معرفة ما إذا كان مأوىً لأسد جائع أم لا. سوف

يكون من المنطقي بالنسبة لك أن تتجنب الكهف إذا قررت بشكل معقول أن الضرر المتوقَّع من الدخول يفوق الفائدة المتوقَّعة.

إن عدم اليقين وظهور الخطأ ، في الدرجة الأولى من تقييماتنا للمخاطر ، هو بحد ذاته شيء يجب أن نضعه في الاعتبار في تقييم الاحتمالية التي نعتبرها كل شيء. وغالبا ما يهيمن هذا العامل على الاحتمالية المنخفضة ، والمخاطر المرتفعة للعواقب – خاصة تلك التي تنطوي على ظواهر طبيعية نفتقر لفهمها بشكل جيد ، أوالديناميات الاجتماعية المعقدة ، أو التكنولوجيا الجديدة ، أو التي يصعب تقييمها لأسباب أخرى. افترض أن بعض التحليلات العلمية A تشير بأن احتمال وقوع كارثة ما X ضئيل للغاية Probability P(X) . عندئذٍ ، قد يكون احتمال وجود خطأ فادح خفي في A أكبر بكثير من P(X) . (5) علاوة على ذلك ، فإن الاحتمال الشرطي لـ X بالنظر إلى أن A معيبة بشكل حاسم ، P(X |–A) ، قد يكون مرتفعاً بما فيه الكفاية. ربما نجد عندئذٍ أن معظم مخاطر الكارثة X تكمن في عدم اليقين في تقييمنا العلمي بأن P(X) كان صغيرًا (Ord, Hillerbrand and Sandberg, 2010) (الشكل رقم 1) .



الشكل رقم 1 : عدم اليقين على المستوى الفوقي.

المصدر: Ord et al., 2010 . يمكن أن يؤدي تحليل إمكانية الخطأ في تقييماتنا الأولية للمخاطر إلى تضخيم احتمالية المخاطر التي جرى تقييمها بأنها صغيرة للغاية. يعطي التحليل الأولي (الجانب الأيسر) احتمالًا ضئيلًا لكارثة (الشريط الأسود) . لكن التحليل قد يكون خاطئًا ؛ تم التمثيل لهذا بالمنطقة الرمادية (الجانب الأيمن) . قد تكمن معظم المخاطر المدروسة في المنطقة الرمادية أكثر من الشريط الأسود.

فئات المخاطر النوعية Qualitative risk categories

نظرا إلى أن المخاطرة تعد احتمالاً يجري تقييمه بشكل سلبي ، فإن خطورة المخاطرة – في الواقع ، ما يجب النظر إليه باعتباره محفوفاً بالمخاطر تماماً – تعتمد على التقييم. وقبل أن نستطيع تحديد مدى جدية المخاطرة ، يجب علينا تحديد معيار تقييم يجري من خلاله قياس القيمة السلبية لسيناريو الخسارة المعينة المحتملة. هناك عدة أنواع من معيار التقييم هذا. على سبيل المثال ، يمكن للمرء استخدام وظيفة المنفعة التي تمثل بعض جوانب تفضيل عامل معين على النتائج المختلفة. قد يكون هذا مناسبا عندما يكون واجب المرء هو تقديم دعم اتخاذ القرار بالنسبة لصانع قرار معين. لكننا هنا سوف ننظر في التقييم المعياري ، وهو ما يبرر ، أخلاقيا ، إسناد قيمة لمختلف النتائج المحتملة. ويكون هذا النوع من التقييم أكثر صلة عندما نتساءل عن أولويات تخفيف المخاطر لدى مجتمعنا (أو أفرادنا).

هناك نظريات متضاربة في الفلسفة الأخلاقية حول: أي التقييمات المعيارية تعد صحيحةً. لن أحاول هنا الفصل في أي خلاف أساسي في علم الأكسيولوجيا. بدلاً من ذلك ، دعونا نفكر في نسخة مبسطة من فئة مهمة واحدة من النظريات المعيارية. لنفترض أن حياة الأشخاص عادة ما يكون لها قيمة إيجابية مهمة وأن هذه القيمة تجميعية (بمعنى أن قيمة حياتين متشابهتين هي ضعف قيمة حياة واحدة) . دعونا نفترض أيضا أن ، مع ثبات جودة ومدة الحياة ، لا تعتمد قيمتها على وقت حدوثها ، أو على ما إذا كانت موجودة بالفعل ، أو على وشك دخولها حيز الوجود نتيجة لأحداث وخيارات مستقبلية. يمكن تخفيف هذه الافتراضات وإدخال التعقيدات ، لكننا سوف نقتصر في مناقشتنا على أبسط حالة.

في هذا الإطار ، إذن ، يمكننا وصف جدية المخاطر تقريباً باستخدام ثلاثة متغيرات: النطاق (حجم السكان المعرضين للخطر) ، والخطورة (مدى الضرر الذي قد يتأثر به هؤلاء السكان) ، والاحتمال (مدى احتمالية حدوث الكارثة ، وفقاً للحكم الأكثر منطقية ، وبالنظر إلى الأدلة المتاحة حاليًا) . وباستخدام الأوَّلَيْن من هذه المتغيرات ، يمكننا إنشاء رسم تخطيطي نوعي لأنواع المخاطر المختلفة. (الشكل 2)

المدى

(كوني)	غير محسوسة	ممكن تحملها	ساحقة	(جهنمي)
شامل للأجيال	تدمير لوحة أصلية من لوحات بيكاسو	تدمير تراث ثقافي	X	خطورة وجودية
عابر للأجيال	انخفاض التنوع البيولوجي بواسطة نوع من الخنفساء	الفترة المظلمة	التعمير / الشيخوخة	
عالمي	الاحترار العالمي بمقدار 0.01 درجة مئوية	انخفاض سُمك طبقة الأوزون	طغيان عالمي سريع الزوال	خطورة كارثية عالمية
محلي	الازدحام بسبب وجود مركبة إضافية	ركود في إحدى البلاد	إبادة جماعية	
شخصي	فقدان شعر المرء	السيارة مسروقة	حادث سيارة مميت	

الخطورة

الشكل رقم 2 . فئات المخاطر النوعية.

المصدر: المؤلف.

ملحوظة: يمكن أن يكون نطاق المخاطرة شخصيًا (يؤثر على شخص واحد فقط) ، أو محليًا (يؤثر على منطقة جغرافية معينة أو مجموعة مميزة) ، أو عالمياً (يؤثر على جميع السكان البشر أو جزء كبير منهم) ، أو عابر للأجيال (يؤثر على البشرية على مدى أجيال عديدة ، أو شامل للأجيال (يؤثر على البشرية ككل ، أو جميعها تقريباً ، أو الأجيال المستقبلية) . ويمكن تصنيف شدة المخاطرة باعتبارها غير محسوسة (بالكاد ملحوظة) ، أو يمكن تحملها (تسبب ضررا كبيرا ولكنها لا تدمر نوعية الحياة تماما) ، أو ساحقة (تسبب الموت أو انخفاض دائم وجذري في نوعية الحياة) .

(يمكن عرض مدى الاحتمال على امتداد المحور z)

تمثل المنطقة التي تحمل علامة "X" في الشكل رقم 2 مخاطر وجودية. هذه هي فئة المخاطر التي لها (على الأقل) قوة ساحقة وعلى مدى أجيال (على الأقل) . (6) وكما هو ملاحَظ ، فإن الخطر الوجودي هو الخطر الذي يهدد بالتسبب في انقراض أصل الحياة الذكية على الأرض ، أو الفشل الذريع الدائم لتلك الحياة في تحقيق إمكاناتها من أجل التطور المرغوب. بعبارة أخرى ، فإن الخطر الوجودي يهدد مستقبل الجنس البشري بأكمله.

حجم الخسارة المتوقعة في الكارثة الوجودية

Magnitude of expected loss in existential catastrophe

مع استمرار الاحتمالية ، تصبح المخاطرة أكثر خطورة ونحن نتحرك نحو المنطقة العلوية اليسرى من الشكل رقم 2 . وبالنسبة لأي احتمال ثابت ، تكون المخاطر الوجودية في هذه الحالة أكثر خطورة من فئات المخاطر الأخرى. لكن مدى خطورة الأمر قد لا يكون واضحا بشكل بديهي. قد يظن المرء أنه يمكننا السيطرة على مدى سوء كارثة وجودية من خلال النظر في بعض الكوارث التاريخية الأكثر سوءً والتي يمكن أن نفكر فيها – مثل الحربين العالميتين ، أو جائحة الإنفلونزا الإسبانية ، أو الهولوكوست – ثم نتخيل شيئا أكثر سوءً بقليل. ومع ذلك ، إذا نظرنا إلى إحصاءات سكان العالم بمرور الوقت ، نجد أن هذه الأحداث المروعة في القرن الماضي لم يتم تسجيلها. (الشكل 3)

لكن حتى هذا التفكير يفشل في إبراز جدية الخطر الوجودي. إن ما يجعل الكوارث الوجودية سيئة بشكل خاص ليس كونها سوف تظهر بقوة على منطقة مثل تلك الموجودة في الشكل رقم 3 ، مما يتسبب في انخفاض حاد في عدد سكان العالم أو متوسط جودة الحياة. بدلاً من ذلك ، تكمن أهميتها في المقام الأول في حقيقة أنها سوف تدمر المستقبل. لقد أشار الفيلسوف ديريك بارفيت إلى نقطة مماثلة في التجربة الفكرية التالية:

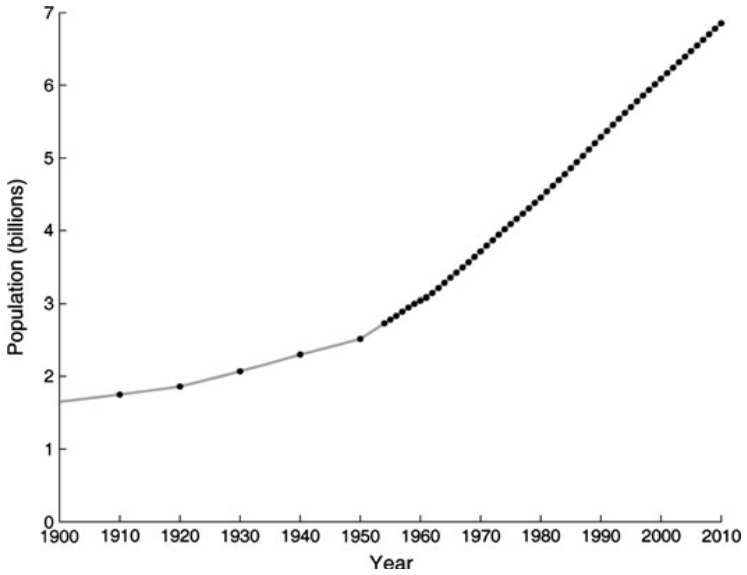
أعتقد أننا إذا دمرنا البشرية ، كما هو في استطاعتنا الآن ، فإن هذه النتيجة سوف تكون أسوأ بكثير مما يعتقده معظم الناس. قارن هذه النتائج الثلاث:

1- السلام.

2- حرب نووية تقتل 99 في المائة من سكان العالم الحاليين.

3- حرب نووية تقتل 100 في المائة.

النتيجة رقم 2 أسوأ من النتيجة رقم 1 ، والنتيجة رقم 3 أسوأ من النتيجة رقم 2 . أيٌّ هذين الاختلافين يعد الأكبر؟ يعتقد معظم الناس أن الفرق الأكبر هو الموجود بين النتيجة رقم 1 والنتيجة رقم 2 . وعلى ما أعتقد أن الفرق بين النتيجة رقم 2 والنتيجة رقم 3 أكبر بكثير. إن كوكب الأرض سوف يظل صالح للسكن لمدة مليار سنة أخرى على الأقل. فالحضارة بدأت منذ بضعة آلاف من السنين فقط. إذا لم تُدمَّر البشرية ، فقد تكون البضعة آلاف من السنين هذه مجرد جزء ضئيل من التاريخ البشري المتحضر بأسره. قد يكون الاختلاف بين النتيجة رقم 2 والنتيجة رقم 3 بذلك هو الفرق بين هذا الجزء الصغير وكل ما تبقى من هذا التاريخ. إذا شبَّهنا هذا التاريخ المحتمل بمدة يوم ، فإن ما حدث حتى الآن هو جزء من الثانية. (Parfit, 1984, pp. 453-454)



الشكل 3. سكان العالم خلال القرن الماضي.

المصدر: المؤلف.

ملحوظة: نادراً ما تسجل كوارث مثل جائحة الأنفلونزا الإسبانية والحربين العالميتين والهولوكوست. (إذا نظر المرء بإمعان في الرسم البياني ، فربما يستطيع بالكاد ملاحظة انخفاض طفيف مؤقت في معدل نمو سكان العالم خلال هذه الأحداث) .

ولحساب الخسارة المرتبطة بكارثة وجودية ، يجب أن نفكر في مقدار القيمة التي سوف تبزغ إلى حيز الوجود في غيابها. ليتضح أن الإمكانات النهائية للحياة الذكية الناشئة عن الأرض تُعد فلكية بالمعنى الحرفي للكلمة.

فالمرء يحصل على عدد كبير ، حتى لو قصر نظره على الإمكانات البيولوجية للكائنات البشرية التي تعيش على الأرض. إذا افترضنا مع بارفيت Parfit أن كوكبنا سوف يظل صالحا للسكن لمدة مليار سنة أخرى على الأقل ، وافترضنا أن ما لا يقل عن مليار شخص يمكنهم أن يعيشوا عليه بشكل مستدام ، فإن الاحتمال قائم لوجود ما لا يقل عن 10^{16} حياة بشرية بالمدة الطبيعية. ويمكن أن تكون هذه الحيوات أيضا أفضل بكثير من متوسط الحياة البشرية المعاصرة ، والتي غالبا ما يشوبها المرض ، والفقر ، والظلم ، والقيود البيولوجية المختلفة التي يمكن التغلب عليها جزئيا من خلال التقدم التكنولوجي والأخلاقي المستمر.

ومع ذلك ، فإن الرقم ذا الصلة ليس عدد الأشخاص الذين يمكنهم أن يعيشوا على الأرض ولكن عدد الأحفاد (الأجيال المستقبلة) الذين يمكن أن يوجدوا إجمالاً. يوجد تقدير للحد الأدنى لعدد سنوات الحياة البشرية البيولوجية في الكون الذي يمكن الوصول إليه في المستقبل (بناءً على التقديرات الكونية الحالية) وهو 10^{34} سنة. (7) تقدير آخر ، يفترض أن العقول المستقبلية التي سوف يجري تنفيذها بشكل أساسي في الأجهزة الكمبيوترية computational hardware بدلاً من البرامج العصبية البيولوجية biological neuronal wetware ، تنتج حدًا أقل هو 10^{54} من سنوات الحياة الذاتية لمحاكاة الدماغ البشري (أو 10^{71} عملية كمبيوترية أساسية) (Bostrom, 2003) (8) . إذا افترضنا الافتراض الأقل تحفظا بأن الحضارات المستقبلية يمكنها أن تقترب في النهاية من الحدود المطلقة للفيزياء المعروفة (باستخدام بعض التقنيات التي لم يتم تصورها حتى الآن) ، فإننا نحصل على تقديرات أعلى بشكل جذري لمقدار الحوسبة وتخزين الذاكرة الذي يمكن تحقيقه وبالتالي من عدد سنوات الخبرة الذاتية التي يمكن تحقيقها. (9)

حتى لو استخدمنا أكثر هذه التقديرات تحفظا ، والذي يتجاهل تمامًا إمكانية استعمار الفضاء وعقول البرمجيات ، نجد أن الخسارة المتوقعة من كارثة وجودية أكبر من قيمة 10^{16} حياة بشرية. وهذا يعني أن القيمة المتوقعة لتقليل الخطر الوجودي بمقدار واحد على مليون فقط من نقطة مئوية واحدة هي على الأقل مائة ضعف قيمة مليون حياة بشرية. التقدير الأكثر شمولاً من الناحية التكنولوجية لـ 10^{54} من سنوات الحياة الذاتية لمحاكاة الدماغ البشري (أو 10^{52} حياة من طول الحياة العادية) يجعل نفس النقطة أكثر وضوحا. حتى لو أعطينا هذا الحد الأدنى المزعوم لإمكانية الإنتاج التراكمي لحضارة ناضجة تقنيا فرصة بنسبة 1 في المائة فقط لكونه صائباً ، فإننا نجد أن القيمة المتوقعة لتقليل المخاطر الوجودية بمقدار واحد من المليار فقط من واحد من المليار من نقطة مئوية واحدة تساوي مائة مليار ضعف ما يساوي مليار حياة بشرية.

وبالتالي قد يحتج المرء بأنه حتى أصغر قدر من تقليل لمخاطرة وجودية واحدة له قيمة متوقعة أكبر من القيمة المحددة لأيَّة خيرية "عادية" ، مثل المنفعة المباشرة لإنقاذ حياة مليار شخص. وعلاوة على ذلك ، فإن القيمة المطلقة للتأثير غير المباشر لإنقاذ مليار حياة شخصية من إجمالي المقدار التراكمي للمخاطرة الوجودية – سواء كانت إيجابية أم سلبية – هي من شبه المؤكد أنها أكبر من القيمة الإيجابية للفائدة المباشرة لمثل هذا الإجراء. (10)

ماكسيبوك Maxipok

تشير هذه الاعتبارات إلى أن الخسارة في القيمة المتوقعة الناتجة عن كارثة وجودية تكون هائلة لدرجة أن هدف تقليل المخاطر الوجودية ينبغي أن يكون اعتبارا مهيمنا عندما نتصرف من منطلق الاهتمام غير الشخصي بالبشرية ككل. وقد يكون من المفيد اعتماد القاعدة الأساسية التالية لمثل هذا العمل الأخلاقي غير الشخصي:

تقوم قاعدة مكسيبوك Maxipok بتعظيم احتمال "النتيجة الجيدة OK outcome" ، حيث تكون النتيجة الجيدة هي أي نتيجة تتجنب كارثة وجودية.

في أفضل الأحوال ، تُعد مكسيبوك maxipok قاعدة عامة أو اقتراحًا للوهلة الأولى. إنها ليست مبدأ للصلاحية المطلقة ، لأنه من الواضح وجود غايات أخلاقية بخلاف منع الكارثة الوجودية. فائدة هذا المبدأ

هي في كونه عاملاً مساعداً في تحديد الأولويات. فالإيثار غير المقيد ليس شائعا لدرجة أنه يمكننا تحمل تبديده بعيدا في عدد كبير من المشاريع الجيدة ذات الفعالية دون المستوى الأمثل. فإذا كانت إفادة البشرية من خلال زيادة السلامة الوجودية تحقق الخير المتوقع على نطاق أكبر بكثير مما تحققه المساهمات البديلة ، فمن الأفضل أن نركز على هذا العمل الخير والأكثر كفاءة.

لاحظ أن قاعدة مكسيبوك maxipok تختلف عن مبدأ الحد الأقصى الشائع القائل ("اختر الإجراء الذي له أفضل نتيجة من بين أسوأ الحالات") . (11) ونظرًا إلى أننا لا نستطيع القضاء تماما على المخاطر الوجودية – في أية لحظة ، فربما نُقذَف في سلة مهملات التاريخ الكوني بواسطة الجبهة المتطورة من تحول النطاق الفضائي الذي تم إطلاقه في إحدى المجرات البعيدة منذ مليار سنة – فإن استخدام مبدأ الحد الأقصى في السياق الحالي يستلزم اختيار الإجراء الذي له أكبر فائدة ، في ظل افتراض الانقراض الوشيك. وبالتالي يشير مبدأ الحد الأقصى إلى أنه يجب علينا جميعا أن نبدأ الاحتفال كما لو لم يكن ثمة غد قادم. هذا المعنى ، رغم أنه قد يكون مغريا ، إلا أنه غير قابل للتصديق.

2- تصنيف الخطر الوجودي 2. Classification of existential risk

لتوجيه الانتباه إلى النطاق الكامل للمخاطر الوجودية ، يمكننا التمييز بين أربع فئات من هذه المخاطر: الانقراض البشري human extinction ، والركود الدائم permanent stagnation ، والتحقق الخاطئ flawed realisation ، والخراب اللاحق subsequent ruination . ونحددها في الجدول رقم 1 أدناه:

الجدول 1. فئات الخطر الوجودي	
الانقراض البشري	تنقرض البشرية قبل الأوان ، أي قبل بلوغ النضج التكنولوجي. (12)
الركود الدائم	الإنسانية تعيش ولكنها لا تصل إلى مرحلة النضج التكنولوجي.
	الفئات الفرعية: انهيار لا نهوض منه ، والكساد ، والانهيار المتكرر
التحقق الخاطئ	تصل الإنسانية إلى النضج التكنولوجي ولكن بطريقة مَعيبة وبشكل

	مخيف ولا يمكن إصلاحه. الفئات الفرعية: التحقق غير المكتمل ، التحقق العابر
الخراب اللاحق	تصل الإنسانية إلى مرحلة النضج التكنولوجي بطريقة تعطي آفاق مستقبلية جيدة ، إلا أن التطورات اللاحقة تسبب الخراب الدائم لتلك الآفاق.
المصدر: المؤلف.	

ونعني بكلمة "الإنسانية humanity" ، هنا ، الحياة الذكية الناشئة عن الأرض ونعني بـ "النضج التكنولوجي technological maturity" تحقيق القدرات التي توفر مستوى من الإنتاجية الاقتصادية والتحكم في الطبيعة إلى درجة قريبة من الحد الأقصى الذي يمكن تحقيقه عمليا.

الانقراض البشري **Human extinction**

على الرغم أنه من المتصور في غضون مليار سنة أو نحو ذلك قد تظل الأرض فيها صالحة للإقامة قبل أن ترتفع درجة حرارتها بسبب الشمس المتمددة ، فإن نوعا ذكيا جديدا سوف يتطور على كوكبنا لملء المكانة التي تخليها البشرية المنقرضة ، إلا أن هذا بعيد جدا عن الحدوث المؤكد. وتقل احتمالية عودة الحياة الذكية إذا أدت الكارثة التي تسبب انقراض الجنس البشري أيضا إلى إبادة القردة العليا وأقاربنا الآخرين ، كما يحدث في الكثير من سيناريوهات انقراض الإنسانية (وإن لم يكن كلها) . بالإضافة إلى أنه ، حتى لو تطورت أنواع ذكية أخرى لتحل محلنا ، فليس هناك ما يضمن أن الأنواع اللاحقة سوف تؤثر بشكلٍ كافٍ على الصفات التي نمتلك السبب لاعتبارها ذات قيمة. قد يكون الذكاء ضروريا لتحقيق إمكاناتنا المستقبلية للتطور المرغوب ، لكنه ليس كافيا. جميع السيناريوهات التي تنطوي على الانقراض المبكر للبشرية سوف تُحسب على أنها كوارث وجودية ، على الرغم من أن بعض هذه السيناريوهات ، وفقا لبعض نظريات القيمة ، ربما تكون حميدة نسبيا. وليس جزءً من تعريف الكارثة الوجودية اعتبار أن كل الأشياء سيئة ، على الرغم من أن هذا سوف يكون على الأرجح افتراضا معقولا في معظم الحالات.

لقد عرَّفنا أعلاه " الإنسانية " بأنها الحياة الذكية التي نشأت على الأرض ، وذلك بدلا من اعتبارها الأنواع المحددة بيولوجيا بالإنسان العاقل Homo sapiens . (13) والسبب في تركيز فكرة الخطر الوجودي

على هذا المفهوم الأوسع هو أنه لا يوجد سبب لافتراض أن مفهوم الأنواع البيولوجية يتبع ما لدينا السبب لاعتباره ذا قيمة. إذا كان لجنسنا أن يتطور ، أو يستخدم التكنولوجيا في التعديل الذاتي ، لدرجة أنها لم تعد تفي بالمعايير البيولوجية لهوية الأنواع (مثل إمكانية التكاثر) مع الإنسان العاقل المعاصر ، فلا داعي لاعتبار هذا كارثة بأي حال من الأحوال. واعتمادا على ما تحولنا إليه ، قد يكون هذا التحول مرغوبا فيه بدرجة كبيرة جدا. في الواقع ، قد يشكل التثبيت الدائم لأي احتمال من هذا النوع من التغير التحويلي للطبيعة البيولوجية البشرية ، في حد ذاته ، كارثة وجودية.

لقد ركزتْ معظم المناقشات حول المخاطر الوجودية حتى الآن ، بشكل حصري ، على الفئة الأولى من الفئات الأربع ، "الانقراض البشري" . والإطار الحالي يلفت الانتباه إلى ثلاثة من أنماط الفشل الأخرى للبشرية. مثل الانقراض ، فقد تنطوي أنماط الفشل الأخرى هذه على سحق جيل كامل. ومن ثم فهي تنطوي على خطورة مماثلة ، وتنطوي على خسائر هائلة محتملة مماثلة للقيمة المتوقعة.

الركود الدائم **Premanent stagnation**

يحدث الركود الدائم إذا نجت البشرية ولكنها لم تصل أبدا إلى مرحلة النضج التكنولوجي – أي ، تحقيق القدرات التي توفر مستوى من الإنتاجية الاقتصادية والتحكم في الطبيعة الذي يكون قريبا من الحد الأقصى ، الذي يمكن تحقيقه عمليا (في امتلاء الوقت in the fullness of time) (م.2) وفي حالة عدم وجود عوائق كارثية) . على سبيل المثال ، يمكن لحضارة ناضجة تكنولوجيا (افتراضا) الانخراط في استعمار الفضاء ، على نطاق واسع ، من خلال استخدام "مسابير فون نيومان ذاتية التكرار" (Freitas, 1980; Moravec, 1988; Tipler, 1980) . وسوف تكون أيضا قادرة على تعديل وتعزيز البيولوجيا البشرية – على سبيل المثال ، من خلال استخدام التكنولوجيا الحيوية المتقدمة أو تقنية النانو الجزيئية (Freitas, 1999, 2003) . علاوة على ذلك ، يمكنها إنشاء أجهزة حسابية قوية للغاية ، واستخدامها لإنشاء محاكاة الدماغ بالكامل ، وأنواع من الحواس الاصطناعية بالكامل ، والعقول فائقة الذكاء (Sandberg and Bostrom, 2008) . وقد يكون لديها كثير من القدرات الإضافية ، بعضها قد لا يمكن تخيله بشكل كامل من وجهة نظرنا الحالية. (14)

يُعد التدمير الدائم لفرصة البشرية لتحقيق النضج التكنولوجي خسارة فادحة للوهلة الأولى ، لأنه يمكن استخدام قدرات الحضارة الناضجة تقنيا لإنتاج نتائج ذات قيمة كبيرة ، مثل الأعداد الفلكية لحياة طويلة ومُرضية للغاية. وبشكل أكثر تحديدا ، سوف تتيح التكنولوجيا الناضجة استخداما أكثر كفاءة للموارد الطبيعية الأساسية (مثل المادة ، والطاقة ، والفضاء ، والوقت ، والترتيب) لخلق قيمة أكبر مما هو ممكن باستخدام تكنولوجيا أقل تقدما. وسوف تسمح التكنولوجيا الناضجة لأن يكون حصاد هذه الموارد (من خلال استعمار الفضاء) أكثر بكثير مما هو ممكن ، باستخدام التكنولوجيا التي يقتصر انتشارها على الأرض وجوارها المباشر.

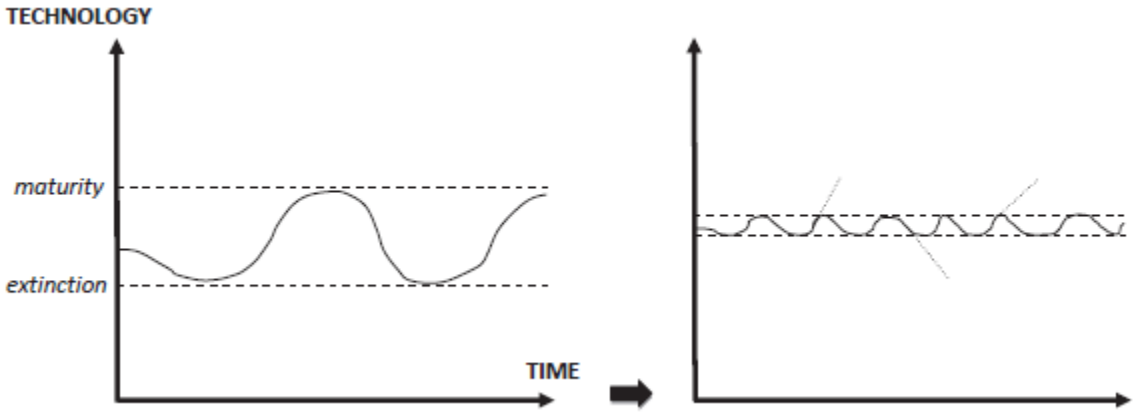
يمكننا التمييز بين أنواع مختلفة من سيناريوهات الركود الدائم: انهيار غير متعافي unrecovered collapse – حيث يحدث الفَقد لكثير من قدراتنا الاقتصادية والتكنولوجية الحالية ولا تتعافى أبدا ؛ والفتور plateauing – حيث يتم تسطيح التقدم عند مستوى ، ربما ، أعلى إلى حد ما من المستوى الحالي ولكن أقل بكثير من النضج التكنولوجي ؛ والانهيار المتكرر recurrent collapse – وهو دورة لا تنتهي من الانهيار يتبعها التعافي. (Bostrom, 2009) . (15)

تعتمد المعقولية النسبية لهذه السيناريوهات على عوامل مختلفة. قد يتوقع المرء أنه حتى لو تعرضت الحضارة العالمية لانهيار كامل ، ربما بعد حرب نووية حرارية عالمية ، فسوف يتم إعادة بنائها في النهاية. ومن أجل الحصول على سيناريو انهيار دائم معقول ، يحتاج المرء بالتالي إلى تفسير سبب عدم حدوث الانتعاش. (16) وفيما يتعلق بالفتور ، فإن الاتجاهات الحديثة للتغير الاجتماعي والتكنولوجي السريع تجعل مثل هذا التهديد يبدو أقل اقترابا ؛ ومع ذلك ، يمكن وضع سيناريوهات بحيث يكون فيها ، على سبيل المثال ، نظاما عالميا مستقرا يحجب المزيد من التغير التكنولوجي. (17) وبالنسبة لسيناريوهات الانهيار المتكرر ، يبدو أنها تتطلب افتراض نوع خاص من الأسباب:

السبب رقم 1– قوي بما يكفي لإحداث الانهيار الكامل للحضارة العالمية ،

ولكنه 2– ليس بالقوة الكافية لكي يتسبب في انقراض الإنسان ،

وأنه 3- يمكن أن يتكرر بشكل معقول في كل مرة يتم فيها إعادة بناء الحضارة إلى مستوى معين ، على الرغم من أي اختلاف عشوائي في الظروف الأولية وأي محاولات من قِبل الحضارات المتعاقبة للتعلم من إخفاقات أسلافهم. ويتضاءل احتمال البقاء على مسار الانهيار المتكرر مع عدد الدورات المفترضة. كلما زاد الأفق الزمني المدروس (وهذا ينطبق أيضًا على الفتور) زادت احتمالية تمزق النمط ، مما يؤدي إما إلى اختراق في الاتجاه التصاعدي نحو النضج التكنولوجي ، أو في الاتجاه التنازلي نحو الانهيار الذي لا علاج له و ، ربما ، الانقراض. (الشكل 4) (18)



الشكل قم 4. الانهيار المتكرر إلى أجل غير مسمى؟

المصدر: المؤلف

ملاحظة: تمثل حالة الإنسان الحديث مجموعة ضيقة من فضاء الاحتمالات. كلما طالت مدة النظر في النطاق الزمني ، كلما قلَّتْ احتمالية أن يظل مستوى التطور التكنولوجي للبشرية محصورا ضمن الفترة الزمنية المحددة عند الغاية الدنيا من خلال أية قدرة تكنولوجية ضرورية للبقاء ، وعند الغاية العليا من خلال النضج التكنولوجي.

التحقق الخاطئ **Flawed realisation**

يحدث التحقق الخاطئ إذا وصلت البشرية إلى النضج التكنولوجي بطريقة خاطئة ومَعيبة بشكل لا يمكن إصلاحه. ونعني بكلمة "لا يمكن إصلاحه irremediably" أنه لا يمكن تصحيحه لاحقا. ونعني بكلمة

"مَعيب dismally" أنه تمكّن من تحقيق ، ولكن جزء صغير من ، القيمة التي كان من الممكن أن تتحقق بخلاف ذلك. أما تصنيف السيناريو باعتباره مثالا على التحقق الخاطئ فيتطلب حكم القيمة. ونعود إلى هذه القضية المعيارية في القسم التالي.

يمكننا التمييز بين نوعين من التحقق الخاطئ:

التحقق غير المكتمل **unconsummated realisation**

والتحقق غير المستدام **ephemeral realization**

في التحقق غير المكتمل ، تُطوِّر البشرية تكنولوجيا ناضجة ، ولكنها تفشل في استخدامها بشكل جيد ، بحيث لا يكون مقدار القيمة المحقَّقة سوى جزء صغير مما كان يمكن تحقيقه. المثال على هذا النوع هو السيناريو الذي يحل فيه الذكاء الآلي محل الذكاء البيولوجي ، ولكن الآلات جرى تشييدها بطريقة تفتقر إلى الوعي (بمعنى التجربة الفينومينولوجية / الظاهراتية) (Bostrom, 2004) . قد يكون المستقبل عندئذٍ ثريًا وقادراً للغاية ، ولكن ، بمعنى ذي صلة ، غير مأهول بالسكان: (يمكن القول) أنه لن تكون هناك كائنات ذات صلة أخلاقية موجودة لكي تستمتع بالثروة. حتى لو لم يتلاشى الوعي تماما ، فقد يكون هناك قدرٌ أقل بكثير مما سينتج عن المزيد من الاستخدام الأمثل للموارد. بدلا من ذلك ، قد يكون هناك قدرٌ كبير من الخبرة ولكن بجودة أقل بكثير مما كان ينبغي أن يكون عليه الحال: العقول أقل سعادة مما كان يمكن أن تكون عليه. أو ، مرة أخرى ، قد يكون هناك عدد كبير من العقول السعيدة جدا ، ولكن هناك عنصر حاسم آخر مفقود يستلزمه تحقيق مستقبل ذي قيمة كبيرة.

في التحقق غير المستدام ، تُطوِّر البشرية تكنولوجيا ناضجة ، والتي يجري استخدامها بشكل جيد في البداية. لكن النضج التكنولوجي يتم بلوغه بطريقة تجعل الحالة الممتازة في البداية غير مستدامة ومحكوم عليها أن تتدهور. يوجد وميض قيمة ، يتبعه غسق أو الظلام دائم. إحدى الطرق التي يمكن أن ينتج عنها التحقق غير المستدام إذا كانت هناك شروخ في الحالة الأولية للنضج التكنولوجي من شأنها أن تؤدي إلى انقسام البشرية إلى فصائل متنافسة. قد يكون من المستحيل إعادة دمج البشرية بعد حدوث مثل هذا الانقسام ، وربما تكون عملية بلوغ النضج التكنولوجي قد قدمت الفرصة الأخيرة والأفضل للبشرية لتشكيل فرد مستقل

(Bostrom, 2006) . وفي غياب التنسيق العالمي ، قد تؤدي العمليات المختلفة إلى تدهور إمكانات البشرية على المدى الطويل. إحدى هذه العمليات هي الحرب بين القوى الكبرى ، على الرغم من أنه ربما يكون من غير المرجح أن مثل هذه الحرب لن تنتهي أبدا (بدلا من أن يتم إنهاؤها نهائيا من خلال معاهدة أو غزو) . (19) تتضمن عملية تآكل أخرى erosive process أشكالا غير مرغوب فيها من المنافسة التطورية والاقتصادية في بيئة كبيرة من ذكاء الآلة (Hanson, 1994) . ومع ذلك ، هناك عملية أخرى من هذا القبيل هي سباق استعمار الفضاء الذي قد يحرق فيه الأنداد الموارد الكونية في جهد مبذر للتغلب في المنافسة. (Hanson, 1998).

الخراب اللاحق **Subsequent ruination**

لإكمال هذا الموضوع ، نسجل فئة رابعة من المخاطر الوجودية وهي: الخراب اللاحق. وفي سيناريوهات من هذا النوع ، تصل البشرية إلى مرحلة النضج التكنولوجي مع إعداد أولي "جيد" (بمعنى أنها ليست مَعيبة بشكل مخيف أو لا يمكن إصلاحها) ، ومع ذلك فإن التطورات اللاحقة تؤدي إلى الخراب الدائم لآفاقنا.

من منظور عملي ، لا داعي للقلق بشأن الخراب اللاحق. حيث أن ما يحدث بعد أن تصل البشرية إلى مرحلة النضج التكنولوجي ليس شيئا يمكننا التأثير عليه الآن ، بخلاف التأكد من أن البشرية تصل إليه وبطريقة توفر أفضل الاحتمالات الممكنة للتطور اللاحق – أي من خلال تجنب الفئات الثلاث الأخرى من الخطر الوجودي. ومع ذلك ، فإن مفهوم الخراب اللاحق وثيق الصلة بنا بطرق مختلفة. على سبيل المثال ، من أجل تقدير مقدار القيمة المتوقعة المكتسبة عن طريق تقليل المخاطر الوجودية الأخرى بمقدار معين ، نحتاج إلى تقدير القيمة المتوقعة ، بشرط تجنب المجموعات الثلاث الأولى من المخاطر الوجودية ، التي تتطلب تقدير احتمالية الخراب اللاحق.

قد يكون احتمال الخراب اللاحق منخفضا – وربما يكون منخفضا للغاية ، بشرط الحصول على الإعداد الصحيح.

– أحد الأسباب هو أنه بمجرد إنشاء الكثير من المستعمرات الفضائية القائمة على الاكتفاء الذاتي ، فإن أية كارثة محصورة في كوكب واحد لا يمكنها أن تقضي على البشرية جمعاء.

– سبب آخر هو أنه بمجرد الوصول إلى النضج التكنولوجي بأمان ، سوف يكون هناك عدد أقل من التقنيات التي يُحتمل أن تكون خطرة ليجري اكتشافها.

– والسبب الثالث هو أن الحضارة الناضجة تقنيا سوف تكون فائقة الذكاء (أو تملك إمكانية الوصول إلى مشورة الكيانات الاصطناعية فائقة الذكاء) وبالتالي تكون أكثر قدرة على التنبؤ بالخطر ووضع خطط لتقليل الخطر الوجودي. في حين أن التنبؤ لن يقلل من المخاطر إذا لم يكن هناك إجراء فعال متاح ، ويمكن للحضارة ذات النضج التكنولوجي أن تتخذ إجراءات ضد مجموعة كبيرة من المخاطر الوجودية. علاوة على ذلك ، إذا اتضح أن بلوغ النضج التكنولوجي دون الوصول إلى حالة استقلال فردي يحكم على الحضارة بالتدهور الذي لا رجعة فيه ، عندئذٍ إذا تم تجنب التحقق المَعيب ، يمكننا أن نفترض أن ما حققناه من حضارة ناضجة تكنولوجياً يمكنها أن تحل مشكلات التنسيق العالمي ، مما يزيد من قدرتها على اتخاذ إجراءات فعالة لمنع الخراب اللاحق.

قد يكون المصدر الرئيسي لخطر الخراب اللاحق هو المواجهة مع أعداء أذكياء خارجيين ، مثل الكائنات الذكية من خارج كوكب الأرض أو أجهزة المحاكاة. لاحظ ، مع ذلك ، أن تلك السيناريوهات التي تنقرض فيها البشرية في النهاية نتيجة لحدود فيزيائية قاسية ، مثل الموت الحراري للكون ، لا تعتبر خراباً لاحقاً ، بشرط أن تكون البشرية ، قبل زوالها ، قد تمكنت من تحقيق جزء كبير بشكل معقول من إمكاناتها في التطور المرغوب. مثل هذه السيناريوهات ليست كوارث وجودية بل نجاحات وجودية.

3- القدرة والقيمة 3. Capability and value

يمكن لبعض الملاحظات الإضافية أن تساعدنا في توضيح الروابط بين القدرة ، والقيمة ، والمخاطر الوجودية.

تحويل الموارد إلى قيمة **Convertibility of resources into value**

نظرا إلى أن مستقبل البشرية من المحتمل أن يكون طويلا بصورة فلكية ، فإن تكامل الخسائر المرتبطة بأوجه عدم الكفاءة المستمرة كبير جدا. هذا هو السبب في أن سيناريوهات التحقق الخاطئ والخراب اللاحق تشكل كوارث وجودية على الرغم من أنها لا تنطوي بالضرورة على الانقراض. (20) قد يستحق الأمر تراجعا مؤقتا في الرفاهية قصيرة الأجل من أجل تأمين تحقق أكثر كفاءة على المدى الطويل للإمكانات البشرية.

ولتجنب التحقق الخاطئ ، من الأهمية بمكان التركيز على تعظيم الكفاءة طويلة الأجل بدلا من تعظيم الناتج الأولية للقيمة في الفترة التي تلي النضج التكنولوجي مباشرة. وذلك لأن مقدار بناء القيمة التي يمكن إنتاجها في وقت معين لا تعتمد فقط على مستوى التكنولوجيا ولكن أيضا على الموارد المادية وأشكال رأس المال الأخرى المتاحة في ذلك الوقت. في لغة الاقتصاد ، لا تعتمد حدود إمكانية الإنتاج البشري (تمثل المجموعات المختلفة الممكنة من النواتج التي يمكن أن ينتجها الاقتصاد العالمي) فقط على وظيفة الإنتاج العالمي (أو "وظيفة الإنتاج العليا meta-production function") ولكن أيضا على المقدار الإجمالي لجميع عوامل الإنتاج (العمالة ، والأرض ، والسلع الرأسمالية المادية ، وما إلى ذلك) المتاحة في وقت ما. ومع التكنولوجيا الناضجة ، تكون معظم عوامل الإنتاج قابلة للتبديل وقابلة للاختزال في النهاية إلى الموارد المادية الأساسية ، ولكن مقدار الطاقة المجانية المتاحة للحضارة يفرض قيودا صارمة على ما يمكن أن تُنتجه. ونظرا إلى أن سرعة الاستعمار محددة بسرعة الضوء ، فإن الحضارة التي تحقق النضج التكنولوجي سوف تبدأ بمنحة متواضعة من الموارد المادية (كوكب واحد وربما بعض الأجزاء القريبة من نظامه الشمسي) ، وسوف يستغرق هذا وقتا طويلا جدا – مليارات السنين – قبل أن تصل الحضارة الصاعدة حتى إلى 1% واحد في المائة من الحد الأقصى لقاعدة مواردها التي يمكن بلوغها. (21) ولذلك ، فإن كفاءة الاستخدام في أوقات لاحقة ، وليس في أعقاب بلوغ النضج التكنولوجي مباشرة ، هي الأكثر أهمية بالنسبة لمقدار القيمة المحقَّقة في نهاية المطاف.

علاوة على ذلك ، قد يتضح أن الطريقة المثالية لاستخدام معظم المنح الكونية التي يمكن للبشرية تأمينها في النهاية هي تأجيل الاستهلاك لأطول فترة ممكنة. ومن خلال الحفاظ على طاقتنا المجانية

المتراكمة حتى يصبح الكون أكبر عمرا وأكثر برودة ، قد نتمكن من إجراء بعض العمليات الحسابية بشكل أكثر كفاءة. (22) وهذا يعزز النقطة التي مفادها أنه سوف يكون من الخطأ وضع قدر كبير من الأهمية على مقدار القيمة المتولدة بعد وقت قصير من النضج التكنولوجي عند تحديد ما إذا كان ينبغي اعتبار بعض السيناريوهات تحققا مَعيبا (أو خرابا لاحقا) . من الأهمية بمكان الحصول على الإعداد الصحيح ، بمعنى وضع البشرية على المسار الذي سوف يحصد في النهاية معظم الموارد الكونية التي يمكن الحصول عليها ووضعها في الاستخدام شبه الأمثل. لا يهم ما إذا كان هناك تأخير موجز قبل حدوث ذلك – والتأخير حتى عدة ملايين من السنين يعتبر "قصيرا" في هذا السياق. (Bostrom, 2003).

حتى بالنسبة للفاعلين الأفراد ، قد يصبح مرور الوقت الفلكي أقل أهمية بعد النضج التكنولوجي. الفاعلون الموجودون في صورة عمليات حسابية في الأجهزة الحاسوبية الموزعة يُحتمل أن تكون لديهم فترات حياة غير محدودة. وينطبق الشيء نفسه على الفاعلين المجسَّدين في عصر تكون فيه تقنيات الإصلاح المادي متقدمة بما فيه الكفاية. ويتناسب مقدار العمر المتاح لمثل هؤلاء الفاعلين مع مقدار الموارد المادية التي يتحكمون فيها. (فيمكن لعقل البرمجيات أن يختبر قدرا معينا من الوقت الشخصي من خلال التشغيل على جهاز كمبيوتر بطيء لفترة طويلة من الوقت الفلكي ، أو بالتساوي ، من خلال التشغيل لفترة وجيزة من الزمن الفلكي على جهاز كمبيوتر سريع) . حتى من خلال ما يُدعى بالمنظور الأخلاقي "المؤثر على الشخص" ، لذلك ، عند تقييم ما إذا كان التحقق المَعيب قد حدث ، لا ينبغي للمرء أن يركز على مقدار القيمة التي يتم خلقها بعد بلوغ النضج التكنولوجي ، ولكن على ما إذا كانت الظروف التي تم خلقها جديرة بأن تعطي فرصة جيدة لتحقيق جزء كبير من القيمة خلال الفترة المتبقية من عمر الكون.

بعض وجهات النظر الأخلاقية الأخرى Some other ethical perspectives

لقد نظرنا حتى الآن في المخاطر الوجودية من المنظور النفعي (جنبا إلى جنب مع العديد من الافتراضات المبسطة) . قد نفكر بإيجاز في كيفية ظهور المشكلة عند عرضها من خلال عدسات بعض الآفاق الأخلاقية الأخرى. على سبيل المثال ، يحدد الفيلسوف روبرت آدامز Robert Adams وجهة نظر مختلفة حول هذه الأمور كالآتي:

"أعتقد أن الأساس الأفضل للنظرية الأخلاقية في هذا المجال يمكن العثور عليه في اتجاه مختلف تماما – في الالتزام بمستقبل البشرية كمشروع كبير للغاية ، أو شبكة متداخلة من المشاريع التي يتقاسمها الجنس البشري بشكل عام. التطلع إلى مجتمع أفضل – أكثر عدلا ، وأكثر مكافأة ، وأكثر سلاما – يُعد جزء من هذا المشروع. هكذا تكون المهام التي يُحتمل أن تكون بلا نهاية بالنسبة للمعرفة العلمية ، والفهم الفلسفي ، وتطوير التقاليد الفنية والثقافية الأخرى. وهذا يشمل التقاليد الثقافية الخاصة التي ننتمي إليها ، بكل صورها التاريخية العرضية وتنوعها العِرقي. كما يتضمن اهتمامنا بحياة أبنائنا وأحفادنا ، والأمل في أن يتمكنوا ، بدورهم ، من أن يعيشوا حياة أبنائهم وأحفادهم باعتبارها مشاريع. إلى الحد الذي تبدو فيه السياسة أو الممارسة مواتية أو غير مواتية لتنفيذ هذا الجمع من المشاريع في المستقبل القريب أو البعيد ، ولدينا السبب لمتابعة ذلك أو تجنبه. ... إن الاستمرارية مهمة لالتزامنا بمشروع مستقبل البشرية بقدر أهميتها لالتزامنا بمشاريع مستقبلنا الشخصي. تماما كما أن شكل حياتي كلها ، وارتباطها بحاضري وماضيَّ ، لهما أهمية تتجاوز أية تجربة منعزلة ، كذلك شكل التاريخ البشري على مدى فترة طويلة من المستقبل ، وارتباطه بالإنسان حاضره وماضيه ، لهما أهمية تتجاوز نوعية الحياة (الإجمالية أو المتوسطة) للسكان – في وقت معين ، بمعزل عن كيفية حدوث ذلك.

أعتقد أننا مدينون ببعض الولاء لهذا المشروع الخاص بمستقبل البشرية. نحن مدينون له أيضا بالاحترام ، الذي ندين به حتى لو لم نكن نحن أنفسنا من الجنس البشري ، ولكن الكائنات من كوكب آخر لديها بعض الفهم له". (Adams, 1989, pp. 472–473) .

ونظرا إلى أن أية كارثة وجودية سوف تضع حدا لمشروع مستقبل البشرية ، أو تقلل بشكل كبير من نطاق تطورها ، فإننا يبدو أن لدينا سببا بديهيا قويا لتجنبها ، من وجهة نظر آدامز.

ونلاحظ أيضا أن أية كارثة وجودية قد تؤدي إلى إحباط الكثير من التفضيلات القوية ، مما يشير إلى أن ذلك من منظور التفضيل والرضا سوف يكون أمرا سيئا. وعلى نفس المنوال ، فإن وجهة النظر الأخلاقية التي تؤكد على أن السياسة العامة ينبغي أن يتم تحديدها من خلال التداول الديمقراطي المستنير من قِبل جميع أصحاب المصلحة سوف تفضِّل التخفيف من المخاطر الوجودية إذا افترضنا ، كما هو معقول ، أن غالبية سكان العالم سوف يؤيدون مثل هذه السياسات عبر المداولات المعقولة (حتى لو كان الناس في

المستقبل الافتراضي لم يتم تضمينهم باعتبارهم أصحاب مصلحة) . ينبغي أن يكون لدينا أيضا واجبات الوصاية للحفاظ على ميراث البشرية الذي ورَّثه أسلافنا إلينا وننقله بأمان إلى أحفادنا. (23) إننا لا نريد أن نكون الحلقة الفاشلة في سلسلة الأجيال ، ولا ينبغي لنا حذف أو التخلي عن ملحمة الحضارة الإنسانية العظيمة التي عملت البشرية عليها منذ آلاف السنين ، عندما يتضح أن السرد أبعد ما يكون عن الوصول إلى النهاية الطبيعية. علاوة على ذلك ، تستنكر وجهات نظر لاهوتية كثيرة الكوارث الوجودية الطبيعية ، وخاصة تلك التي تسببها الأنشطة البشرية:

فإذا كان الله قد خلق العالم والجنس البشري ، فإن الإنسان سوف يتخيل أن الله قد يكون مستاءً إذا أخذنا على عاتقنا تحطيم صنعته الرائعة (أو إذا سمحنا لها ، من خلال إهمالنا أو غطرستنا ، بالتسبب في ضرر لا يمكن إصلاحه) . (24)

ينبغي أن ننظر أيضا إلى المسألة من وجهة نظر أقل تجريدا ، وبدلا من ذلك نحاول صياغة تقييم من خلال النظر في الحالات المماثلة والتي لدينا حدوس أخلاقية محددة حولها. وهكذا ، على سبيل المثال ، إذا شعرنا بالثقة في أن ارتكاب إبادة جماعية صغيرة أمر خاطئ ، وأن ارتكاب إبادة جماعية كبيرة لا يقل خطأ ، فقد نخمن أن ارتكاب القتل أمر خاطئ أيضا. (25) وإذا كنا نعتقد أن لدينا سببا أخلاقيا لمنع الكوارث الطبيعية التي من شأنها أن تقتل عددا قليلا من الناس ، وسبب أخلاقي أقوى لمنع الكوارث الطبيعية التي من شأنها أن تقتل عددا أكبر من الناس ، فينبغي علينا أن نخمن أن لدينا سببا أخلاقيا أقوى لمنع الكوارث التي من شأنها أن تقتل جميع البشر.

تتفق وجهات نظر معيارية متباينة كثيرة في دعمها لتقليل المخاطر الوجودية ، على الرغم من أن درجة السوء المتضمنة في كارثة وجودية ، والأولوية التي يجب أن يحظى بها التخفيف من المخاطر الوجودية في اقتصادنا الأخلاقي ، قد تختلف اختلافا كبيرا بين النظريات الأخلاقية المتنوعة. (26) لاحظ ، مع ذلك ، أن الكوارث الوجودية ليست من الحقائق المفاهيمية السيئة بأي حال من الأحوال أو أن تقليل المخاطر الوجودية أمر صحيح. هناك حالات محتملة يكون فيها حدوث نوع واحد من الكوارث الوجودية مفيدا – على سبيل المثال ، لأنه يستبق نوعاً آخر من الكوارث الوجودية التي كانت ستحدث بالتأكيد لولا ذلك وكان من المحتمل أن يكون أسوأ.

في حين أن الفئتين الأوليين من المخاطر الوجودية (الانقراض البشري ، والركود الدائم) يجري تحديدهما بمعايير وصفية بحتة ، والأخريان (التحقق الخاطئ ، والخراب اللاحق) محددتان بشكل معياري. إلا أن هذا يعني أن مفهوم الخطر الوجودي يُعد في جزء منه فكرة تقييمية. (27)

عندما يتعلق الأمر بالقضايا المعيارية ، قد تكون هذه القضايا مثيرة للجدل. فأخلاقيات السكان ، على سبيل المثال ، محفوفة بالمشاكل المتعلقة بكيفية التعامل مع المعايير المختلفة (مثل حجم السكان ، ومتوسط الرفاهية ، وعتبات ما يعتبر حياة تستحق العيش ، وعدم المساواة ، ونفس الشيء مقابل خيارات الأشخاص المختلفة) . من المحتمل أيضا أن يكون تقييم بعض السيناريوهات التي تنطوي على تحولات جوهرية في الطبيعة البشرية محل خلاف (Fukuyama, 2002; Glover, 1984; Kass, 2002; Savulescu and Bostrom, 2009) . ومع ذلك ، ليست كل القضايا المعيارية مثيرة للجدل. سوف يكون من المتفق عليه بشكل عام ، على سبيل المثال ، أن المستقبل الذي يعيش فيه عدد قليل من البشر حياة بائسة ، داخل نظام بيئي محطم ، في وجود قدرات تكنولوجية كبيرة ولكنها غير مستخدَمة ، سوف تعتبر هذه القدرات التكنولوجية تحققا خاطئا لإمكانات البشرية ، وسوف تشكل كارثة وجودية إذا لم يتم تغيير اتجاهها.

سوف تكون هناك بعض أنواع من المخاطر الوجودية المفترضة التي يكون فيها عدم اليقين الرئيسي معياريا ، وأخرى يكون فيها عدم اليقين الرئيسي إيجابيا. وفيما يتعلق بعدم اليقين الإيجابي أو الوصفي ، فقد رأينا سابقا أنه إذا كان هناك شيء غير معروف بأنه آمن بشكل موضوعي ، فإنه ينطوي على مخاطرة ، على الأقل ، بالمعنى الموضوعي المرتبط باتخاذ القرار. ويمكننا القيام بخطوة موازية فيما يتعلق بعدم اليقين المعياري. افترض أن حدثا ما X من شأنه أن يقلل من التنوع البيولوجي. لنفترض (من أجل التوضيح) أنه من المعروف أن X لن يكون لها عواقب مهمة أخرى ، وأن التنوع البيولوجي الذي جرى تخفيضه لن يؤثر على البشر أو أية كائنات أخرى ذات أهمية أخلاقية. الآن ، قد نكون غير متأكدين مما إذا كان التنوع البيولوجي له قيمة نهائية (يُعتبر ذو قيمة "لذاته") . ومن ثم قد نكون غير متأكدين مما إذا كانت X سوف تكون سيئة حقيقةً أم لا. لكن يمكننا أن نقول أنه إذا لم نكن متأكدين مما إذا كانت X سوف تكون سيئة حقيقةً أم لا (لكننا على يقين من أن X لن تكون جيدة) ، فإن X سيئة على الأقل بالمعنى الموضوعي

المتعلق باتخاذ القرار. وهذا يعني أن لدينا سببا لتفضيل عدم حدوث X وربما سببا لاتخاذ إجراء لمنع حدوث X .

كيف ينبغي للمرء بدقة أن يأخذ في الاعتبار عدم اليقين الأخلاقي الأساسي باعتباره سؤالا مفتوحا ، لكن من الواضح أنه ينبغي للمرء أن يفعل ذلك. (Bostrom, 2009) وبالتالي يمكننا تضمين مواقف المخاطر الوجودية التي نعرف فيها ما سوف يحدث ، ونحكم بشكل عقلاني بأن ما سوف يحدث قد يكون سيئا من الناحية الوجودية – حتى عندما لا يكون هناك في الواقع أي شيء سيء بشأن النتيجة.

يمكننا تسليط الضوء على إحدى عواقب ذلك: لنفترض وجود جِنِّي يمكن الاعتماد عليه تماما ، حيث يمكنه ضمان تحقيق أية أمنية للبشرية ربما تحتاجها لمستقبلها. بعد ذلك – حتى لو تمكنا جميعا من الاتفاق على مستقبل واحد من هذا القبيل – سوف يظل علينا مواجهة خطر وجودي محتمل وأكثر خطورة: أي ، الاختيار غير الحكيم واختيار المستقبل المَعيب بشكل مخيف على الرغم من وضوحه ، في لحظة اختيارنا ، باعتباره أكثر ما نرغب فيه بالنسبة لجميع العقود المستقبلية الممكنة.

الاحتفاظ بخياراتنا حية **Keeping our options alive**

هذه التأملات حول عدم اليقين الأخلاقي تشير إلى طريقة تكميلية بديلة للنظر في المخاطرة الوجودية ؛ وتقترح أيضا طريقة جديدة للتفكير في الشكل النموذجي للاستدامة. اسمحوا لي أن أناقشها.

قد يكون فهمنا الحالي لعلم الأكسيولوجيا مشوَّشا تماما. قد لا نعرف الآن – على الأقل ليس بالتفصيل الملموس – ما هي النتائج التي يمكن اعتبارها انتصارا كبيرا للبشرية ؛ قد لا نكون ، حتى الآن ، قادرين على تخيل أفضل الغايات لرحلتنا. إذا كنا حقا غير متأكدين تماما من أهدافنا النهائية ، فعلينا أن ندرك أن هناك قيمة كبيرة للخيار في الحفاظ على – وتطوير بشكل مثالي – قدرتنا على تحقيق القيمة وتوجيه المستقبل وفقا لذلك. إن ضمان وجود نسخة مستقبلية من البشرية ذات قوى عظمى ، وميل لاستخدامها بحكمة ، هو أفضل طريقة متاحة لنا لزيادة احتمالية احتواء المستقبل على الكثير من القيمة. وللقيام بذلك ، يجب أن نمنع أية كارثة وجودية.

وبالتالي نريد أن نصل إلى حالة نمتلك فيها:

1- ذكاءً أكبر بكثير ، ومعرفة ، وحكما سليما بدرجة أكبر مما نتمتع به حاليًا ؛

2- قدرة على حل مشكلات التنسيق العالمي أكبر بكثير ؛

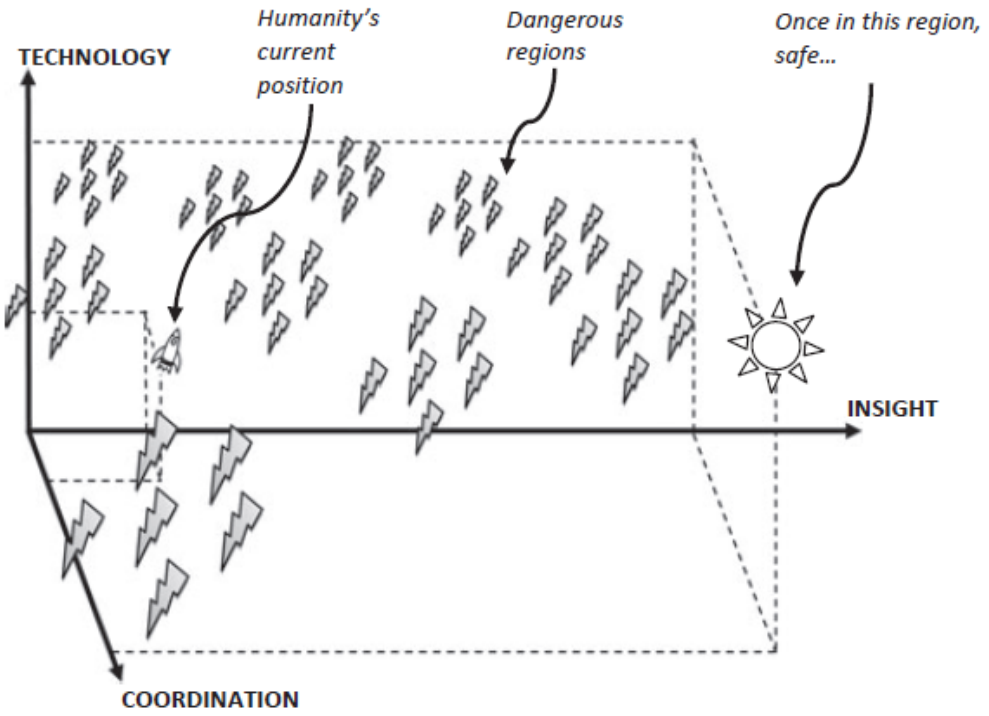
3- قدرات تكنولوجية وموارد مادية أكبر بكثير ؛ وحيث أن

4- قيَمَنا وتفضيلاتِنا لا تفسد في عملية الوصول إلى هدفنا (ولكنها بالأحرى ، إذا أمكن ، يجري تحسينها) .

فالمطلبان 2 و 3 يعملان على توسيع مجموعة الخيارات المتاحة للبشرية.

ويُزيد المطلب 1 من قدرة البشرية على التنبؤ بنتائج الخيارات المتاحة وفهم ما سوف تنطوي عليه كل نتيجة من حيث تحقيق القيم الإنسانية.

أخيرًا ، يجعل المطلب 4 الإنسانية أكثر رغبة في تحقيق القيم الإنسانية.



الشكل رقم 5- التحدي المتمثل في إيجاد طريق آمن

المصدر: المؤلف.

ملاحظات: قد يكون الوضع المثالي هو الوضع الذي نتمتع فيه بمستوى عالٍ جدا من التكنولوجيا ، وتنسيق عالمي ممتاز ، ورؤية ثاقبة حول كيفية استخدام قدراتنا. لا يتبع ذلك أن امتلاكنا أي قدر من التكنولوجيا الإضافية ، أو التنسيق ، أو البصيرة يكون دائما مفيدا لنا. ربما يكون من الضروري أن يتماشى نمونا على أبعاد مختلفة مع مخطط معين من أجل أن يتبع تطورنا مسارا عبر مساحة الحالة التي تصل في النهاية إلى المنطقة المرغوبة.

الطريقة التي تمكننا ، من وضعنا الحالي ، أن نحقق هذه الغايات على أفضل وجه ، ليست واضحة (الشكل 5) . وبينما نحتاج في النهاية إلى مزيد من التكنولوجيا ، والبصيرة ، والتنسيق ، ليس من الواضح أن أقصر طريق للوصول إلى الهدف هو الأفضل.

يمكن أن يتضح ، على سبيل المثال ، أن تحقيق قدرات تكنولوجية معينة قبل الحصول على الرؤية والتنسيق الكافيْن يؤدي دائما إلى هلاك الحضارة. فيمكن للمرء أن يتخيل بسهولة فئة من سيناريوهات الكارثة الوجودية حيث يتم اكتشاف نوع من التكنولوجيا التي تضع قوة تدميرية هائلة في أيدي عدد كبير من الأفراد. إذا لم يكن هناك دفاع فعال ضد هذه القوة التدميرية ، ولا توجد طريقة لمنع الأفراد من الوصول إليها ، فلا يمكن للحضارة أن تدوم ، لأنه في عدد السكان الكبير بما فيه الكفاية ، لا بد أن يكون هناك بعض الأفراد الذين سوف يستخدمون أية قوة تدميرية متاحة لديهم . كان من المحتمل أن يكون اكتشاف القنبلة الذرية على هذا النحو ، باستثناء الحقيقة التي اعتمدت على حُسن الحظ وهي أن تصنيع الأسلحة النووية يتطلب مكوناً خاصاً – مادة انشطارية صالحة للاستخدام في صنع الأسلحة – وكونه نادرا ومكلفا في تصنيعه. ومع ذلك ، إذا أخذنا عينات بشكل مستمر من جرة / جعبة الاكتشافات التكنولوجية المحتملة قبل تطبيق وسائل فعالة للتنسيق والمراقبة و/أو تقييد المعلومات التي قد تكون خطرة على الصعيد العالمي ، فإننا نجازف في النهاية بسحب كرة سوداء: تسهِّل التدخل في تصنيع يسبب ضررا واسع النطاق للغاية ، ويكون الدفاع الفعال ضده غير قابل للتنفيذ. (28)

لذلك ، ربما لا ينبغي أن نسعى مباشرة لتقريب حالة معينة "مستدامة" ، بمعنى أننا يمكن أن نبقى فيها لبعض الوقت. بدلا من ذلك ، ينبغي أن نركز على الدخول في مسار تطويري يوفر احتمالية عالية لتجنب كارثة وجودية. بعبارة أخرى ، ينبغي أن ينصب تركيزنا على تعظيم الفرص التي سوف نحققها يوما ما في

النضج التكنولوجي بطريقة ليست مخيفة وغبر معيبة بشكل لا يمكن إصلاحه. يُشترط في هذا الإنجاز ، أن يقدم لنا فرصة جيدة لتحقيق إمكاناتنا الأكسيولوجية الفلكية.

لتوضيح هذه النقطة ، ضع في اعتبارك القياس التالي. عندما يقف صاروخ على منصة الإطلاق ، فهو في حالة جيدة مستدامة. يمكن أن يظل في وضعه الحالي لفترة طويلة ، على الرغم من أن الرياح والطقس سوف تدمره في النهاية. مكان آخر مستدام للصاروخ هو الفضاء ، حيث يمكنه السفر بدون وزن لفترة طويلة جدا. لكن عندما يكون الصاروخ في الجو ، يكون في حالة انتقالية غير مستدامة: حيث تكون محركاته مشتعلة وسرعان ما ينفد الوقود. إن إعادة الصاروخ إلى حالة مستدامة أمر مرغوب فيه ، لكن هذا لا يعني أن أية طريقة لجعل حالته أكثر استدامة أمر مرغوب فيه. على سبيل المثال ، تقليل استهلاكه للطاقة بحيث يستطيع بالكاد الاحتفاظ بها ثابتة قد يجعل حالته أكثر استدامة بمعنى أنه يمكن أن يبقى في مكان واحد لفترة أطول ؛ ومع ذلك ، عندما ينفد الوقود ، سوف يسقط الصاروخ على الأرض. إن أفضل سياسة للصاروخ في الجو هي الحفاظ على قوة دفع كافية للهروب من مجال الجاذبية الأرضية: وهي إستراتيجية تتضمن إدخال حالة استدامة أقل (استهلاك الوقود بشكل أسرع) من أجل تحقيق حالة أكثر استدامة مرغوبة لاحقا. أي ، بدلا من السعي لتقريب حالة استدامة ، ينبغي أن نتبع مسارا مستداما.

وبالمثل فإن الوضع الإنساني الحالي هو حالة انتقالية. ومثل الصاروخ في تشبيهنا ، تحتاج البشرية إلى اتباع مسارا مستداما ، مسار من شأنه أن يقلل من مخاطر كارثة وجودية. (29) ولكن بخلاف مشكلة تحديد المعدل الأمثل لاستهلاك الوقود في الصاروخ ، فإن مشكلة كيفية تقليل المخاطر الوجودية ليس لها حل معروف.

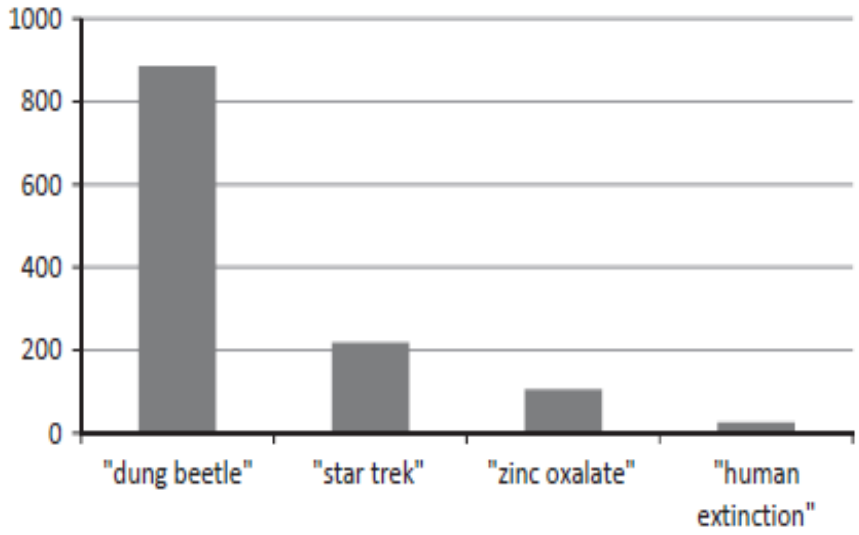
4- رؤية عامة 4. Outlook

لقد رأينا أن الحد من المخاطر الوجودية يظهر باعتباره أولوية مهيمنة في عدد من النظريات الأخلاقية العواقبية التجميعية (وكمصدر اهتمام كبير للغاية في الكثير من النظريات الأخلاقية الأخرى) . وبالتالي يمكن أن يساعد مفهوم الخطر الوجودي الأشخاص ، الذين لديهم دوافع أخلاقية أو إيثار ، على تحديد

الإجراءات التي لها أعلى قيمة متوقعة. على وجه الخصوص ، بالنظر إلى افتراضات محددة ، فإن مشكلة اتخاذ القرار الصحيح تمهد الطريق نحو اتباع مبدأ المكسيبوك.

معوقات التفكير والفعل **Barriers to thought and action**

في ضوء هذه النتيجة ، والتي تشير إلى أنه قد تكون هناك قيمة عالية جدا في دراسة المخاطر الوجودية وفي تحليل استراتيجيات التخفيف المحتملة ، فمن اللافت للنظر مدى ضآلة الاهتمام الأكاديمي الذي حظيت به هذه القضايا ، مقارنة بالموضوعات الأخرى الأقل أهمية. (الشكل رقم 6) (30)



الشكل رقم 6 تحديد الأولويات الأكاديمية.

المصدر: المؤلف.

ملاحظة: عدد الأوراق الأكاديمية حول المواضيع المختلفة (مسجلة في سكوباس Scopus ، أغسطس 2012) . (م.3)

تتآمر عوامل كثيرة ضد دراسة المخاطر الوجودية والتخفيف من حدتها. ربما يحدث تقييد البحث بواسطة الطبيعة متعددة التخصصات للمشكلة ، وأيضا بسبب القضايا المعرفية الأكثر عمقا. أكبر المخاطر الوجودية ليست قابلة للتوصل إلى منهجيات البحث العلمي وتطبيقها. علاوة على ذلك ، هناك قضايا أساسية

لم يتم حلها ، لا سيما فيما يتعلق بنظرية اختيار الملاحظة وأخلاقيات السكان ، والتي تعتبر حاسمة في تقييم المخاطر الوجودية ؛ وهذه الصعوبات النظرية تتفاقم بفعل عوامل نفسية تجعل من الصعب التفكير بوضوح في قضايا مثل نهاية البشرية. (31)

إذا كان توفير المزيد من الموارد لبحث المخاطر الوجودية سوف يتم ، فهناك خطر من تدفقها بغزارة مفرطة ، بسبب أن المخاطر الصغيرة نسبيا ، والتي يسهل على بعض المجتمعات التخصصية الراسخة دراستها باستخدام طرق مألوفة ، على حساب مجالات الخطر الأكثر أهمية – الذكاء الفائق للآلة ، تكنولوجيا النانو الجزيئية المتقدمة ، الشمولية ، المخاطر المتعلقة بفرضية المحاكاة ، أو التطورات المستقبلية في البيولوجيا الاصطناعية / التركيبية – الأمر الذي سوف يتطلب تحولاً غير مريح في تركيز البحث. تحول آخر معقول هو أن البحث سوف يجري توجيهه بشكل أساسي إلى مخاطر كارثية عالمية تنطوي على مخاطر وجودية قليلة أو معدومة.

إن التخفيف من المخاطر الوجودية يعوقه الافتقار إلى الفهم ، وأيضا بسبب نقص الحافز. ويُعد التخفيف من المخاطر الوجودية منفعة عالمية عامة (أي غير قابل للاستبعاد وغير قابل للمنافسة) ، وتشير النظرية الاقتصادية إلى أن مثل هذه المنافع تميل إلى أن تكون غير متوفرة من قِبل السوق ، لأن كل منتِج للسلامة الوجودية (حتى لو كان المنتِج أمة كبيرة) يمكن أن يستحوذ على جزء صغير فقط من القيمة. (Feldman, 1980; Kaul, 1999) . في الواقع ، إن الوضع أسوأ مما هو عليه الحال مع الكثير من المنافع العالمية العامة الأخرى حيث أن الحد من المخاطر الوجودية يعد منفعة عامة عبر الأجيال (في الواقع ، تشمل جميع الأجيال) : حتى لو كان من المحتمل حصول إحدى الحالات العالمية على جزء صغير فقط من الفوائد – تلك التي تعود على الأشخاص الموجودين حاليا. فالمليارات من البشر السعداء الذين قد يأتون إلى الوجود في المستقبل إذا تجنبنا كارثة وجودية سوف يكونون على استعداد لدفع مبالغ فلكية للجيل الحالي مقابل زيادة طفيفة في جهودنا للحفاظ على مستقبل البشرية ، ولكن للأسف تم منع التجارة ذات المنفعة المتبادلة من خلال صعوبات المعاملات الواضحة.

قد تفشل الدوافع الأخلاقية أيضا في قياس حجم المسألة المطروحة للمناقشة. فمن المرجح أن تتجلى حساسية نطاق مشاعرنا الأخلاقية بشكل خاص عندما يتعلق الأمر بأعداد كبيرة جدا:

أعداد أكبر بكثير ، تقترب من 500 مليون حالة وفاة ، وخاصة السيناريوهات المختلفة نوعيا مثل انقراض الجنس البشري بأكمله ، قد تثير طريقة مختلفة من التفكير – تدخل ضمن "سلطة تعليمية منفصلة separate magisterium". فالأشخاص الذين لم يحلموا أبدا بإيذاء طفل يسمعون عن مخاطر وجودية ، ويقولون ، "حسناً ، ربما لا يستحق الجنس البشري حقا البقاء على قيد الحياة". (Yudkowsky, 2008, p. 114) .

تتطلب المخاطر الوجودية نهجا استباقيا. إن نهج رد الفعل – ملاحظة ما يحدث ، والحد من الأضرار ، ثم تنفيذ آليات محسَّنة لتقليل احتمالية حدوث التكرار – لا يفيد عندما لا تكون هناك فرصة للتعلم من الفشل. بدلاً من ذلك ، يجب أن نتوقع الأخطار الناشئة ، وحشد الدعم للعمل ضد الضرر الافتراضي في المستقبل ، واتخاذ احتياطاتنا بشكلٍ كافٍ في المرة الأولى. هذا أمر بالغ الصعوبة. قلة من المؤسسات قادرة على العمل باستمرار بمثل هذا المستوى من العقلانية الفعالة ، ومحاولات تقليد مثل هذا السلوك الاستباقي داخل المؤسسات الأقل مثالية يمكن أن تأتي بنتائج عكسية بسهولة. يمكن استغلال الترويج للمخاطر الفكرية لترشيد العمل العدواني الذي يخدم المصالح الذاتية ، أو توسيع البيروقراطيات الأمنية المكلفة والتي يُحتمل أن تكون قمعية ، أو القيود المفروضة على الحريات المدنية التي تحافظ على بقاء المجتمعات حرة وعقلانية. ويمكن أن تكون نتيجة التقديرات الخاطئة للمثال العقلاني بسهولة زيادة خالصة في المخاطر الوجودية. (32)

إن التحديات متعددة التخصصات والتحديات المعرفية ، والعبث والانحرافات الأكاديمية ، والتحيزات المعرفية ، والمشكلات الأكثر حرية ، والخمول الأخلاقي وعدم الحساسية تجاه المجال ، وعدم الكفاءة المؤسسية ، وتهديدات الاستغلال السياسي غير القابلة للقياس الكمي هي بالتالي بعض العوائق التي تحول دون التخفيف الفعال. ويمكن أن نضيف إلى ذلك صعوبة تحقيق المستويات المطلوبة من التعاون العالمي. بينما يمكن معالجة بعض المخاطر الوجودية من جانب واحد – حيث يمكن لأي دولة لديها صناعة فضائية أن تبني دفاعا عالميا ضد تأثيرات الكويكبات – بينما تتطلب المخاطر الأخرى مشروعا مشتركا بين الكثير من الدول. إن إدارة المناخ العالمي قد تتطلب قبولا لدى أغلبية ساحقة من الدول الصناعية والناهضة صناعيا. كما أن تجنب سباقات التسلح والتخلي عن الاتجاهات الخطرة للبحث التكنولوجي قد يتطلب أن

تنضم جميع الدول لبذل الجهد ، لأن منشقاً واحداً يمكن أن يلغي أية مزايا للتعاون. وقد تتطلب بعض الأخطار المستقبلية أن تقوم كل دولة بمراقبة وتنظيم كل مجموعة مهمة أو فرد مهم داخل أراضيها. (33)

هل ثمة أساسيات التفاؤل؟ Grounds for optimism?

وبالتالي ، هناك مجموعة هائلة من العوائق تلقي بظلالها على احتمالات حدوث استجابة واضحة وفاعلة تجاه المخاطر الوجودية التي تواجه البشرية. ولكي لا يُعتبر السبب ميؤوساً منه ، ينبغي أن نلاحظ أيضا بعض الاعتبارات المشجعة.

قد نلاحظ ، أولاً ، أن الكثير من المفاهيم والأفكار الرئيسية جديدة تماماً. (34) قبل وضع الأسس المفاهيمية والنظرية ، لم يكن من الممكن بناء دعم لجهود البحث والتخفيف من المخاطر الوجودية. ففي كثير من الحالات ، أصبحت الأفكار العلمية والتكنولوجية والمنهجية الأساسية ، اللازمة لدراسة المخاطر الوجودية بطريقة ذات مغزى ، متاحة أيضًا في الفترة الأخيرة فقط. وهذه البداية المتأخرة تساعد في تفسير حالة مجال هذه الدراسة التي لا تزال بدائية.

يمكن القول أنه ، فقط ، منذ انفجار القنبلة الذرية الأولى في عام 1945 ، والتراكم النووي اللاحق خلال الحرب الباردة ، التي أنشأت أية مخاطر وجودية طبيعية (أي غير خارقة للطبيعة) – على الأقل إذا احتسبنا فقط المخاطر التي كان للكائنات البشرية بعض التأثير عليها. (35) لا تزال معظم المخاطر الوجودية الكبيرة تكمن في المستقبل لسنوات عديدة. لذلك ، حتى وقت قريب ، ربما كانت هناك حاجة قليلة نسبيا للتفكير في المخاطر الوجودية بشكل عام ، وفرص قليلة للتخفيف منها ، حتى لو سار التفكير بهذه الطريقة.

يبدو أن الوعي العام بالتأثيرات العالمية للأنشطة البشرية آخذ في الازدياد. حيث تجري اليوم دراسة الأنظمة ، والعمليات ، والمخاطر من منظور عالمي من قِبل الكثير من العلماء – علماء البيئة ، والاقتصاديين ، وعلماء الأوبئة ، وعلماء الديموجرافيا ، وغيرهم. كما أن مشكلات مثل تغير المناخ ، والإرهاب العابر للحدود ، والأزمات المالية الدولية ، توجه الانتباه إلى الترابط العالمي والتهديدات التي يتعرض لها النظام العالمي. ويبدو أن فكرة المخاطرة بشكل عام صعدت إلى مكان الصدارة. (36) وبالنظر

إلى هذه التطورات في المعرفة ، والمناهج ، والمواقف ، وظروف تأمين المخاطر الوجودية والتدقيق الذي تستحقه كلها صارت مواتية بشكل غير مسبوق.

قد تتكاثر فرص الفعل أيضا. كما لوحظ ، يمكن تنفيذ بعض مشاريع التخفيف من جانب واحد ، ويمكن للمرء أن يتوقع المزيد من هذه المشاريع عندما يصبح العالم أكثر ثراءً. وتتطلب مشاريع التخفيف الأخرى تنسيقا أوسع ؛ وفي كثير من الحالات ، تنسيقا عالميا. هنا ، أيضا ، يبدو أن بعض خطوط الاتجاه تشير إلى أن هذا أصبح أكثر جدوى بمرور الوقت. هناك اتجاه تاريخي طويل الأجل نحو زيادة نطاق التكامل السياسي – من عصابات الصيد والجمع إلى المشيخات ، والدول المدن ، والدول القومية ، والآن المنظمات متعددة الجنسيات ، والتحالفات الإقليمية ، والتكوينات الحكومية الدولية المختلفة ، والجوانب الأخرى للعولمة. (Wright, 1999) قد يبدو أن استقراء هذا الاتجاه يشير إلى التكوين النهائي لاستقلال الفرد. (Bostrom, 2006) من الممكن أيضا رؤية أن بعض الحركات العالمية التي ظهرت خلال النصف الأخير من القرن العشرين – لا سيما حركة السلام ، والحركة البيئية ، ومختلف حركات العدالة وحقوق الإنسان العالمية – سوف توسع ، على نحو متزايد ، من المخاوف الأكثر عمومية بشأن المخاطر الوجودية. (37)

علاوة على ذلك ، إلى الحد الذي يكون فيه تقليل المخاطر الوجودية ، حقاً ، هو السبب الأكثر استحقاقا ، قد يتوقع المرء أن التحسينات العامة ، في قدرة المجتمع على إدراك الحقائق المهمة والعمل وفقا لها ، سوف تحول الموارد بشكل مختلف إلى التخفيف من المخاطر الوجودية. وقد تأتي التحسينات العامة من هذا النوع من عدة مصادر ، بما في ذلك التطورات في التقنيات التعليمية وأدوات التعاون عبر الإنترنت ، والابتكارات المؤسسية مثل أسواق التنبؤ ، والتقدم في العلوم والفلسفة ، وانتشار الثقافة العقلانية ، وتعزيز التحقق البيولوجي.

أخيرًا ، من الممكن أن يكون السبب في مرحلة ما الحصول على دَفعة بسبب وقوع كارثة كبرى (غير وجودية) التي تؤكد هشاشة الوضع البشري الحالي. والتي من شأنها أن ، لا داعي للقول ، أن تكون أسوأ طريقة ممكنة لتركيز عقولنا – ومع ذلك ، يجب أن تُمنح ، في إطار زمني متعدد العقود ، احتمالية غير ضئيلة لحدوثها. (38)

هوامش المترجم

(م1) Maxipok Rule تعني تعظيم احتمال وجود نتيجة جيدة ، حيث تكون النتيجة الجيدة متمثلة في أية نتيجة تتجنب كارثة وجودية. وتعد هذه القاعدة ، في أفضل الأحوال ، قاعدة عامة أو اقتراحاً بديهياً للوهلة الأولى.

(م2) في / عند امتلاء الوقت in the fullness of time تعبير في اللغة الانجليزية يعني أن شيئاً ما سوف يحدث في النهاية بعد وقت طويل للغاية أو بعد سلسلة طويلة من الأحداث.

(م3.) تعد سكوباس Scopus قاعدة بيانات ببليوغرافية تحتوي على ملخصات واستشهادات لمقالات المجلات الأكاديمية. وتغطي ما يقرب من 22000 عنوان من أكثر من 5000 ناشر ، وهي أكبر قاعدة بيانات للملخصات والاقتباسات من المؤلفات التي راجعها النظراء: المجلات العلمية والكتب ووقائع المؤتمرات. وتقوم سكوباس Scopus بتقديم لمحة شاملة عن مخرجات البحوث العالمية في مجالات العلوم والتكنولوجيا والطب والعلوم الاجتماعية والفنون والعلوم الإنسانية.

الهوامش

1- أعطى استطلاع غير رسمي بين خبراء أكاديميين بشكل أساسي حول المخاطر الكارثية العالمية المختلفة تقديرًا متوسطًا يبلغ 19 في المائة لاحتمال انقراض الجنس البشري قبل نهاية هذا القرن. (Sandberg and Bostrom, 2008) ولا تمثل آراء هؤلاء المشاركين في الاستطلاع بالضرورة مجتمع الخبراء الأوسع. لقد استخدم تقرير ستيرن ريفيو Stern Review المؤثر في المملكة المتحدة حول اقتصاديات تغير المناخ (2006) احتمال انقراض بنسبة 0.1 في المائة سنويًا في تحليل معدل التخفيض الفعال. وهذا يعادل افتراض 9.5 في المائة من خطر انقراض الإنسان خلال المائة عام القادمة.

(UK Treasury 2006, Chapter 2, Technical Appendix, p. 47)

2- تضعف قوة هذا الاعتبار ، إلى حد ما ، من خلال تأثيرات اختيار الملاحظة التي تلقي "بظلالها البشرية" على الأدلة المتاحة. (Cirkovic, Sandberg and Bostrom, 2010)

3. See Smil, 2008.

4- وهكذا تجري فهرسة الاحتمالية عبر الوقت. الكميات التي تعتمد على الاحتمالية ، مثل جدية المخاطرة ، التي يمكن أن تختلف بمرور الوقت مع توفر معلومات جديدة.

5- هناك الكثير من الأدلة التاريخية على أن التحليلات العلمية تبدو أحيانا مَعيبة بصورة حاسمة.

6- كما هو مبين في الشكل ، يمكن تمديد المحاور لتشمل المخاطر المحتملة من الناحية المفاهيمية والتي تكون أكثر تطرفاً. على وجه الخصوص ، يمكن أن تحتوي المخاطر الشاملة للأجيال على فئة فرعية من المخاطر المدمرة للغاية التي لن يؤثر تحققها على الأجيال البشرية المستقبلية أو يستبقها فحسب ، بل سوف يدمر أيضا إمكانية الجزء من الكون الذي يكمن في مخروط الضوء المستقبلي لإنتاج كائنات ذكية أو مدركة لذاتها (النطاق الكوني) . علاوة على ذلك ، وفقاً لبعض نظريات القيمة ، يمكن أن تكون هناك حالات وجود أسوأ بكثير من عدم الوجود أو الموت (على سبيل المثال ، الأمراض المستعصية الرهيبة) ، لذلك يمكن للمرء من حيث المبدأ تمديد المحور X أيضا ليشمل (الشدة الجهنمية) . لن نتقصى هذه الاحتمالات المفاهيمية في هذه المقالة.

7. This is based on an accelerating universe with a maximal reachable co-moving distance of 4.74 Gpc, a baryonic matter density of $4.55\ 10^{-28}$ kg/m^3, a luminosity ratio of stars ~ 100, and 1 planet per 1,000 stars being habitable by 1 billion humans for 1 billion years (Gott et al., 2005; Heyl, 2005). Obviously the values of the last three parameters are debatable, but the astronomical size of the conclusion is little affected by a few orders-of-magnitude change.

8- يستخدم هذا تقديرًا من قبل عالم المستقبل الراحل روبرت برادبري بأن النجم يمكنه تشغيل 10^{42} عملية في الثانية باستخدام أجهزة كمبيوتر فعالة مشيَّدة بتقنية النانو المتقدمة. علاوة على ذلك ، يفترض (جنبًا إلى جنب مع التقديرات الكونية المذكورة في الحاشية السابقة) أن الدماغ البشري لديه قوة معالجة تبلغ 10^{17} عملية في الثانية وأن النجوم في المتوسط تدوم 5 مليارات سنة. لا يفترض أي تشكل نجمي جديد. انظر أيضا. (Cirkovic, 2004)

9- على سبيل المثال ، إذا تم حفظ كل طاقة الكتلة في الكون الذي يمكن الوصول إليه حتى تتوقف درجة حرارة الخلفية الكونية الميكروية عن الانخفاض (بسبب درجة حرارة الأفق الثابتة التي تبلغ 10^{-29} كلفن) ثم تُستخدم بعد ذلك في الحساب ، فإن هذا سوف يسمح بما يصل إلى 10^{121} حسابات ديناميكية حرارية لا رجعة فيها. (Krauss and Starkman, 2000) انظر أيضا (Cirkovic and Radujkov, 2001) .

10- يجب أن نؤكد ، مع ذلك ، أن هناك قضايا مهمة لم يتم حلها في العواقبية التجميعية - على وجه الخصوص ، فيما يتعلق بالقيم اللانهائية والاحتمالات الضعيفة للغاية (Bostrom, 2003, 2009) . لن نناقش هذه القضايا هنا ، ولكن في القسم 5 سوف نناقش الوضع المعياري لمفهوم الخطر الوجودي من خلال بعض وجهات النظر الأخرى.

11- تبعًا لجون رولز John Rawls ، يُستخدم مصطلح "maximin" بمعنى مختلف في اقتصاديات الرفاهية ، للإشارة إلى المبدأ القائل بأنه (نظرًا لقيود معينة) يجب أن نختار الحالة التي تزيد من توقع الطبقات الأسوأ (Rawls, 1971) . هذه النسخة من المبدأ لا تتأثر بالضرورة بالملاحظات الواردة في النص.

12- يمكن للمرء أن يشير إلى هذا بشكل أكثر دقة باعتباره انقراض بشري "مبكر" أو "سابق لأوانه". لاحظ أن البشرية يمكن أن تنقرض دون إنشاء مثيل من هذه الفئة إذا حققت البشرية إمكاناتها الكامنة ثم انقرضت.

13- قد نأخذ هنا كلمة "ذكي intelligent" لتعني القدرة على تطوير اللغة ، والعلوم ، والتكنولوجيا ، والثقافة التراكمية.

14- ليس من المطلوب أن تقوم حضارة ناضجة تكنولوجيًا بنشر كل هذه التقنيات ؛ يكفي أن تكون متاحة لديها ، بمعنى أن الحضارة يمكن أن تطورها وتنشرها بسهولة وسرعة إذا قررت ذلك. وبالتالي ، فإن حضارة الآلة فائقة الذكاء التي تتمتع بالقوة الكافية والتي يمكن أن تخترع وتنفذ بسرعة هذه التقنيات وغيرها من التقنيات ذات الصلة سوف تُعتبر بالفعل ناضجة من الناحية التكنولوجية.

15- ليس هذا مطلقًا بلا نهاية ، بالطبع ، ولكنه سلسلة من الدورات التي تستمر لفترة طويلة جدًا وتنتهي بانقراض الإنسان دون بلوغ النضج التكنولوجي.

16- قد يفترض سيناريو الانهيار الذي لا علاج له أن بعض الموارد الأساسية للتعافي قد دُمرت بشكل دائم ، أو أن مجموعة الجينات البشرية تتدهور بشكل لا رجعة فيه ، أو ربما تم إجراء بعض الاكتشافات التي تُمكِّن مجموعات صغيرة من إحداث مثل هذا الدمار الهائل الذي يمكن أن يقضي على الحضارة وأن المعرفة بهذه الاكتشافات لا يمكن القضاء عليها.

17- إن تقنيات الحوكمة المحسَّنة ، مثل المراقبة الشاملة ubiquitous surveillance والتعديل الكيميائي العصبي neurochemical manipulation ، قد تعزز قبضة مثل هذا النظام على السلطة إلى حد يجعل الإطاحة به مستحيلة.

18- تتمثل الصعوبة الأخرى لفرضية الانهيار المتكرر في تفسير حقيقة أننا في الدورة التكنولوجية الأولى هنا على الأرض. وإذا كان من الشائع وجود عدد من دورات الانهيار والتعافي (بأحجام سكانية متشابهة) فلماذا نرى أننا في الحلقة رقم 1؟ قد يشير هذا النوع من الاعتبارات البشرية إلى أن الانقراض أو التحول أكثر احتمالًا مما قد يفترضه المرء بسذاجة.

19- حتى التهديد بحرب لا تندلع أبدًا يمكن أن يؤدي إلى إهدار كبير من حيث الإنفاق على الأسلحة وفرص التعاون الضائعة.

20- وهو أيضاً أحد الأسباب التي تجعل الركود الدائم خطرًا وجوديًا ، على الرغم من أن الركود الدائم قد يمنع أيضًا البقاء على قيد الحياة بعد الوقت الذي تصبح فيه الأرض غير صالحة للسكن ، ربما بعد حوالي مليار سنة من الآن بسبب زيادة تألق الشمس. (Schroder and Smith, 2008) .

21- أحد الشروط المهمة المحتملة هو أن الوقت اللازم للوصول إلى الحد الأقصى من قاعدة الموارد التي يمكن بلوغها يمكن أن يكون أقصر إذا ظهرت معارضة ذكية (مثل أن تكون من حضارات خارج كوكب الأرض) تعيق توسعنا الكوني.

22- هناك حد أدنى من الإنتروبيا / طاقة الضغط الداخلي entropy يرتبط بمحو جزء صغير من المعلومات ، وهذا يعد تكلفة تنخفض مع درجة الحرارة.

23- ينبغي أن نتحمل أيضًا مسؤوليات تجاه الكائنات غير البشرية ، مثل الحيوانات الأرضية (وربما خارج كوكب الأرض) . على الرغم من أننا لا نقوم بالكثير حاليًا لمساعدتها ، إلا أن لدينا الفرصة للقيام بذلك في المستقبل. وإذا كان تقديم المساعدة لمحو معاناة الحيوانات غير البشرية في البيئة الطبيعية يمثل قيمة مهمة ، فإن تحقيق النضج التكنولوجي بطريقة تفشل في إنتاج مثل هذه المساعدة يمكن أن يعتبر تحققاً مَعيبًا. انظر McMahan, 2010; Pearce, 2004 .

24- من وجهة نظر لاهوتية ، يمكن أن تكون هناك فئة خاصة من المخاطر الوجودية ذات وضع أخلاقي مختلف: الكوارث أو النذُر بنهاية العالم التي صنعتها الذات الإلهية ، ربما كعقاب على خطايانا. قد يحكم المؤمن على مثل هذا الحدث باعتباره خيراً ، في التوازن. ومع ذلك ، يبدو من غير المعقول أن البشر وحدهم سوف يكونون قادرين على مخالفة الله إذا كان يريد حقًا تدميرنا ، لذلك من المفترض أن تكون أية إجراءات فيزيائية مضادة ننفذها ضد المخاطر الوجودية فعالة ، فقط ، ضد المخاطر الوجودية الطبيعية والبشرية ، وقد لا يكون لدينا سبب للتراجع عن جهود التخفيف من المخاطر الطبيعية خوفًا من مخالفة ما أراده الله.

25- على الرغم من أن قتل كل شيء سوف يكون على الأقل غير متحيز ، على عكس الإبادة الجماعية التي غالباً ما تكون عنصرية أو قومية.

26- على سبيل المثال ، احتج جيمس لينمان James Lenman بأن مسألة متى تنقرض البشرية تعتبر إلى حد كبير مسألة لامبالاة ، على الأقل إذا لم يحدث ذلك في وقت قريب جدًا. (Lenman, 2002)

27- وفي هذا الصدد ، يعتبر مفهوم الخطر الوجودي مشابها لمفاهيم مثل "الديمقراطية" و "سوق العمل الفعال" . إن أشياء مثل الثقب الأسود ، أو جرة الحصى المعقمة ، ليست ديمقراطية أو سوق عمل فعال ، ويمكننا أن نرى أن الأمر كذلك دون الحاجة إلى إصدار أي حكم معياري ؛ ومع ذلك ، قد تكون هناك أشياء أخرى لا يمكن تصنيفها على أنها حالات أو ليست حالات متعلقة بهذه المفاهيم دون اتخاذ موقف (ضمني على الأقل) بشأن بعض المسائل المعيارية.

28- وبطبيعة الحال ، فإن تحقيق تنسيق عالمي قوي فعال بما يكفي لرصد سكان العالم بأسره بصفة دائمة أو فرض رقابة إلى أجل غير مسمى على أية معلومات تعتبرها إحدى السلطات خطيرة (على الأقل في حالة عدم وجود ضمانات كافية) من شأنه أن يؤدي إلى مخاطر وجودية بالغة الأهمية ، مثل مخاطر الركود الدائم أو التحقق الخاطئ في ظل نظام شمولي قمعي.

29- من الناحية المثالية ، سوف تفعل ذلك مع تحقيق وسائل ارتكاب القتل الرحيم الجماعي ، وفي الحالة المستبعدة إلى حد ما ، بعد مداولات جماعية طويلة ودقيقة ، يجب أن نقرر أن النهاية السريعة هي الأفضل من استمرار الوجود. ومع ذلك ، قد

تكون هذه القدرة مفيدة فقط إذا توصلنا أولاً إلى الحكمة الكافية لعدم ممارستها بشكل خاطئ. يجب أن نؤكد على الحاجة إلى استمرار المداولات الفلسفية وتعزيز الظروف التي من شأنها أن تساعدنا في العثور على الحقيقة حول ما يخص القضايا المعيارية المركزية في نهاية المطاف – وكذلك الحاجة إلى تجنب الأخطاء التي لا رجعة فيها في هذه الأثناء.

30– العلاجات العلمية للمخاطر الوجودية في حد ذاتها ، أو حتى لخطر الانقراض البشري ، نادرة (على سبيل المثال ، Bostrom, 2002; Leslie, 1996; Mathhny, 2007; Wells, 2009) . ومع ذلك ، فإن قدرًا كبيرًا من المؤلفات الأكاديمية لها علاقة بالمخاطر الوجودية الفردية أو تدور حول قضايا محددة أخرى ذات صلة بعدد من المخاطر الوجودية (جرى الاستشهاد ببعضها في هذا المقال) . بالإضافة إلى ذلك ، تلقي بعض الأعمال الحديثة نظرة واسعة على المخاطر الكارثية العالمية ، على الرغم من عدم اقتصار التركيز على المخاطر الوجودية (على سبيل المثال ، Bostrom and Cirkovic, 2008; Diamond, 2006; Homer–Dixon, 2007; Posner, 2004; Sunstein, 2009; World Economic Forum, 2011) .

31– تشمل القضايا ذات الصلة المتعلقة بآثار اختيار المراقبة ، من بين أمور أخرى ، حجة يوم القيامة لدى كارتر–ليزلي ، وحجة المحاكاة ، وحجج "التصفية الكبرى" ؛ انظر Bostrom, 2002, 2003, 2008; Carter, 1983; Cirkovic et al, 2010; Hanson, 1998; Leslie, 1996; Tegmark and Bostrom, 2005) . وبالنسبة لبعض القضايا ذات الصلة في الفلسفة الأخلاقية ، انظر على سبيل المثال ، Bostrom, 2003, 2009 . ولمراجعة أدبيات التحيزات المعرفية من حيث صلتها بالمخاطر الكارثية ، انظر Yudkowsky, 2008 .

32– من الطرق الممكنة للتغلب على هذه المشكلة محاولة الحفاظ على القدْر الإجمالي للقلق بشأن المخاطر ثابتًا تقريبًا مع تخصيص نسبة أكبر من حزمة "رموز الخوف" أو "رقائق القلق" للمخاطر الوجودية. وبالتالي ، يمكن للمرء أن يدافع عن القول بأننا كلما أصبحنا أكثر اهتمامًا بالمخاطر الوجودية ، وجب علينا في نفس الوقت أن نصبح أقل اهتمامًا بالمخاطر الأصغر ، مثل موت بضعة آلاف من الأشخاص في هجوم إرهابي غريب أو كارثة طبيعية.

33 – وسوف تصبح هذه المراقبة الداخلية داخل الدول أكثر جدوى مع التقدم في تكنولوجيا المراقبة. وكما لوحظ ، فإن منع الدول التي لديها مثل هذه القدرات من أن تصبح قمعية سوف يطرح مجموعة من التحديات الخاصة بها.

34– تضمين فكرة الخطر الوجودي ذاته. (Bostrom, 2002)

35– يمكن للمرء أن يحتج بأن الأوبئة والمواجهات القريبة مع المذنبات ، والتي حدثت مرارًا وتكرارًا في تاريخ البشرية وأثارت نذيراً قويًا بنهاية العالم ، يجب أن تعتبر مخاطر وجودية كبيرة مبكرة. وبالنظر إلى المعلومات المحدودة التي كانت متوفرة في ذلك الوقت ، فربما يكون من غير المعقول للمراقبين المعاصرين تعيين احتمال كبير لنهاية قريبة. يمكن أيضًا النظر في

سيناريوهات يوم القيامة الدينية ؛ ربما لم يكن من غير المعقول الاعتقاد ، على أساس الأدلة التي كانت متوفرة في ذلك الوقت ، أن هذه المخاطر كانت حقيقية ، وعلاوة على ذلك ، يمكن التخفيف من حدتها من خلال أفعال مثل التوبة ، والصلاة ، والتضحية ، واضطهاد السحرة أو الكفار ، و هكذا دواليك. ربما نشأ الخطر الوجودي العلمي الأول الواضح مع تطوير القنبلة الذرية. فقد أمر روبرت أوبنهايمر Robert Oppenheimer ، القائد العلمي لمشروع مانهاتن ، بإجراء دراسة قبل اختبار ترينيتي Trinity لتحديد ما إذا كان التفجير النووي سيؤدي إلى سلسلة ذاتية الانتشار من التفاعلات النووية في الغلاف الجوي للأرض. قد يمثل التقرير الناتج أول تقييم كمي لمخاطر الانقراض البشري. (Manhattan Project, 1946)

36- وقد ذهب بعض علماء الاجتماع إلى حد التركيز على المخاطر باعتبارها موضوعاً أساسياً في عصرنا. انظر ، على سبيل المثال ، Beck, 1999 .

37- أعرب كثيرون من نشطاء السلام الذين عارضوا سباق التسلح النووي أثناء الحرب الباردة عن قلقهم الصريح بشأن هرمجدون Armageddon نووي مزعوم وأنه يمكن أن يقضي على الحياة البشرية كلها. في الآونة الأخيرة ، دق بعض دعاة حماية البيئة ناقوس الخطر بشأن ظاهرة الاحتباس الحراري ، باستخدام نفس اللغة المروعة. ومع ذلك ، من غير الواضح إلى أي مدى كانت الاحتمالية المتصوَّرة لنتيجة القضاء على الأنواع قوة دافعة رئيسية في هذه الحالات. ربما يكون مقدار القلق هو نفسه تقريبًا حتى في مواجهة ضمان الستار الحديدي بأن أية كارثة سوف تنتهي دون انقراض الإنسان.

38- أنا ممتن للتعليقات والمناقشة مع سيث بوم Seth Baum ، ونِك باكستيد Nick Backstead ، وميلان سيركوفيتش Milan Ciekovic ، وأول هاجستروم Olle Haggstorm ، وسارا ليبينكوت Sara Lippincott ، وجافريك ماتيني Gaverick Matheny ، وتوبي أورد Toby Ord ، وديريك بارفيت Derek Parfit ، ومارتن ريس Martin Rees ، وريبيكا روش Rebecca Roache ، وأندرس ساندبرج Anders Sandberg ، وكارل شولمان Carl Shulman .

المراجع

Adams, R. M. (1989) 'Should ethics be more impersonal? A critical notice of Derek Parfit, Reasons and persons', Philosophical Review, 98 (4), pp. 439–484.

Beck, U. (1999) The world risk society. Cambridge, MA: Polity.

Bostrom, N. (2002) Anthropic bias: observation selection effects in science and philosophy. New York, NY: Routledge.

Bostrom, N. (2002) 'Existential risks: analyzing human extinction scenarios and related hazards', Journal of Evolution and Technology, 9 (1).

Bostrom, N. (2003) 'Are you living in a computer simulation?', Philosophical Quarterly, 53 (211), pp. 243–255.

Bostrom, N. (2003) 'Astronomical waste: the opportunity cost of delayed technological development', Utilitas, 15 (3), pp. 308–314.

Bostrom, N. (2003) 'Infinite ethics', revised version 2009. Available from: http://www.nickbostrom.com/ethics/infinite.pdf [Accessed 15 March 2011].

Bostrom, N. (2004) 'The future of human evolution', in C. Tandy (ed.), Death and anti-death: two hundred years after Kant, fifty years after Turing. Palo Alto, CA: Ria University Press, pp. 339–371.

Bostrom, N. (2006) 'What is a singleton?', Linguistic and Philosophical Investigations, 5 (2), pp. 48–54.

Bostrom, N. (2008) 'Where are they? Why I hope the search for extraterrestrial life finds nothing', MIT Technology Review, May / June, pp.72–77.

Bostrom, N. (2009) 'The future of humanity', in J-K. B. Olsen, E. Selinger

and S. Riis (eds), New Waves in Philosophy of Technology. New York, NY: Palgrave Macmillan, pp. 186–216.

Bostrom, N. (2009) 'Moral uncertainty—towards a solution?', in Overcoming Bias. Available from: http://www.overcomingbias.com/2009/01/moral-uncertainty-towards-a-solution.html

[Accessed 23 March 2011].

Bostrom, N. (2009) 'Pascal's mugging', Analysis, 69 (3), pp. 443–445.

Bostrom, N. and Cirkovic, M. M. (eds) (2008) Global Catastrophic Risks. Oxford: Oxford University Press.

Carter, B. (1983) 'The anthropic principle and its implications for biological evolution', Philosophical Transactions of the Royal Society, A 310, pp. 347–363.

Cirkovic, M. M. (2004) 'Forecast for the next eon: applied cosmology and the long-term fate of intelligent beings', Foundations of Physics, 34 (2), pp. 239–261.

Cirkovic, M. M. and Radujkov, M. (2001) 'On the maximal quantity of processed information in the physical eschatological context', Serbian Astronomy Journal, 163, pp. 53–56.

Cirkovic, M. M., Sandberg, A. and Bostrom, N. (2010) 'Anthropic shadow: observation selection effects and human extinction risks', Risk Analysis, 30 (10), pp. 1495–1506.

Diamond, J. (2006) Collapse: how societies choose to fail or survive. London: Penguin.

Feldman, A. (1980) Welfare economics and social choice theory. Boston, MA: Martinus Nijhoff Publishing.

Freitas, R. A. (1980) 'A self-reproducing interstellar probe', Journal of the British Interplanetary Society, 33, pp. 251–264.

Freitas, R. A. (1999) Nanomedicine volume I: basic capabilities. Austin, TX: Landes Bioscience.

Freitas, R. A. (2003) Nanomedicine volume IIA: biocompatibility. Austin, TX: Landes Bioscience.

Fukuyama, F. (2002) Our posthuman future: consequences of the biotechnology revolution. London: Profile.

Glover, J. (1984) What sort of people should there be? Harmondsworth: Penguin.

Gott, J. R., Juric, M., Schlegel, D., Hoyle, F., Vogeley, M. and Tegmark, M. et al. (2005) 'A map of the universe', Astrophysical Journal, 624 (2), pp. 463–483.

Hanson, R. (1994) 'If uploads come first?', Extropy, 6 (2), pp. 10–15.

Hanson, R. (1998) 'Burning the cosmic commons: evolutionary strategies for interstellar colonization'. Available from: http://hanson.gmu.edu/filluniv.pdf [Accessed 3 April 2011].

Hanson, R. (1998) 'The Great Filter—Are We Almost Past It?' weblog post. Available from: http://hanson.gmu.edu/greatfilter.html [Accessed 15 September 1998 and 24 March 2011].

Heyl, J. S. (2005) 'The long-term future of space travel', Physical Review D, 72, pp. 1–4.

Homer-Dixon, T. (2007) The upside of down: catastrophe, creativity and the renewal of civilization, London: Souvenir Press.

Kass, L. (2002) 'Life, liberty and the defense of dignity'. San Francisco, CA: Encounter.

Kaul, I. (1999) Global public goods. Oxford: Oxford University Press.

Krauss, L. M. and Starkman, G. D. (2000) 'Life, the universe, and nothing: life and death in an ever-expanding universe', Astrophysical Journal, 531 (1), pp. 22–30.

Lenman, J. (2002) 'On becoming extinct', Pacific Philosophical Quarterly, 83 (3), pp. 253–269.

Leslie, J. (1996) The end of the world: the science and ethics of human extinction. London: Routledge.

Manhattan Project (1946) LA-602: ignition of the atmosphere with nuclear bombs, Library Without Walls Project, Los Alamos

National Laboratory. Available from: http://www.fas.org/sgp/othergov/doe/lanl/docs1/00329010.pdf [Accessed 24 March 2011].

Matheny, J. G. (2007) 'Reducing the risk of human extinction', Risk Analysis, 27 (5), pp. 1335–1344.

McMahan, J. (2010) 'The meat eaters', New York Times. Available from:

http://opinionator.blogs.nytimes.com/2010/09/19/the-meat-eaters 19 September 2010 [Accessedviewed 23 March 2011].

Moravec, H. (1988) Mind children: the future of robot and human intelligence. Cambrige, MA: Harvard University Press.

Ord, T., Hillerbrand, R. and Sandberg, A. (2010) 'Probing the improbable: methodological challenges for risks with low probabilities and high stakes', Journal of Risk Research, 13, pp. 191–205.

Parfit, D. (1984) Reasons and Persons. Oxford: Clarendon Press.

Pearce, D. (2004) 'The hedonistic imperative'. Available from: http://www.hedweb.com/welcome.htm [Accessed 23 March 2011].

Posner, R. A. (2004) Catastrophe. Oxford: Oxford University Press.

Rawls, J. (1971) A theory of justice. Cambridge, MA: Harvard University Press.

Sandberg, A. and Bostrom, N. (2008)Global catastrophic risks survey, Future of Humanity Institute Technical Report, #2008-1, Oxford. Available from: http://www.global-catastrophic-risks.com/docs/2008-1.pdf [Accessed 9 March 2011].

Sandberg, A. and Bostrom, N. (2008) Whole brain emulation: a roadmap, Future of Humanity Institute Technical Report, #2008-3, Oxford. Available from: http://www.fhi.ox.ac.uk/__data/assets/pdf_file/0019/3853/brain-emulation-roadmap-report.pdf

[Accessed 22 March 2011].

Savulescu, J. and Bostrom, N. (eds) (2009) Enhancing Humans. Oxford: Oxford University Press.

Schroder, K-P. and Smith, R. (2008) 'Distant future of the Sun and Earth revisited', Monthly Notices of the Royal Astronomical Society, 368 (1), pp. 155-163.

Smil, V. (2008) Global catastrophes and trends: the next fifty years. Cambridge, MA: The MIT Press, Cambridge.

Sunstein, C. (2009) Worst-Case scenarios. Cambridge, MA: Harvard University Press.

Tegmark, M. and Bostrom, N. (2005) 'How unlikely is a Doomsday

catastrophe?', Nature, 438. p. 754.

Tipler, F. (1980) 'Extraterrestrial intelligent beings do not exist', Royal Astronomical Society Quarterly Journal, 21, pp. 267-281.

UK Treasury (2006) Stern review on the economics of climate change. Available from: http://www.hm-treasury.gov.uk/media/8A3/83/Chapter_2_A_-_Technical_Annex.pdf [Accessed 9 March 2011].

Weitzman, M. L. (2009) 'The extreme uncertainty of extreme climate change: an overview and some implications', Harvard University, mimeo, October 2009.

Wells, W. (2009) Apocalypse when? Calculating how long the human race will survive. Praxis, Chichester.

World Economic Forum (2011) Global risks. Available from: http://riskreport.weforum.org/ [Accessed 24 March 2011].

Wright, R. (1999) Nonzero: the logic of human destiny. New York, NY: Pantheon Books.

Yudkowsky, E. (2008) 'Cognitive biases potentially affecting judgment of global risks', in N. Bostrom and M. M. Cirkovic (eds), Global catastrophic risks. Oxford: Oxford University Press, pp. 91–119.

Author Information

Nick Bostrom, Professor, Faculty of Philosophy, Oxford University; and Director of the Future of Humanity Institute in the Oxford Martin School.

Translator Information

Dr. Hussein Abd AlGhani Ebraheem, Free researcher, got his B. A in Philosophy, then M. A. in Freedom and its relation to creativity in Berdyaev, then Ph. D. In Philosophy of Technology and its relation to phenomenological approach in Don Ihde, all from Faculty of Arts, Zagazig University, Egypt.

Translator of,

- Ethics, Technology and Posthuman Societies, Steven Benko. Global Culture Magazine, Kuwait, Vol. 207.

- Sharing the World with Digital Minds. Nick Bostrom. Global Culture Magazine, Kuwait. Under publishing.

- Vulnerable World Hypothesis, Nick Bostrom. Homepage. (Arabic Translation)

And other works.

E-mail, waves7000@yahoo.com