This database was created by Michael Aird, in relation to my work with Convergence Analysis. I intend for it to be a "living document", built up collaboratively by anyone interested in existential risk reduction.

Please add "comments" to suggest additions or mention places where you think I've misinterpreted or misrepresented an estimate, or where there would just be some other context worth noting. I'll check these comments regularly, and, where relevant, copy and paste them into the spreadsheet as additions.

See this post for context on this database

Some things worth bearing in mind:

Some limitations of existential-risk-related estimates, or explicit probability estimates in general, noted in this post (forum.effectivealtruism.org/posts/JQQAQrunyGGhzE23a/database-of-existential-risk-estimates), Beard et al. (https://www.sciencedirect.com/science/article/pii/S0016328719303313), and this post (https://forum.effectivealtruism.org/posts/KfqFLDkoccf8NQsQe/potential-downsides-of-using-explicit-probabilities).

These estimates are unlikely to be very independent; many estimators had probably seen each others' estimates, interacted extensively with other estimators, etc.

I don't provide much context or caveats for most of the specific estimates. It may often be worth checking the appendix of Beard et al. (https://www.sciencedirect.com/science/article/pii/S0016328719303313) and/or the original source.

It's often hard to be sure precisely what is being estimated, or what other conditions are perhaps being assumed. And I may sometimes misinterpret or misrepresent this; please make a comment if you think that that's the case.

I've organised these estimates by the broad categories of what's being estimated. Most estimates within each category are **not** of exactly the same thing; for example, they may differ in the timelines over which they're estimated, or in whether they're about existential catastrophe broadly or extinction specifically.

I've converted all estimates into percentages. Where the estimator expressed their estimate in another way, I've shown how they expressed it in brackets after the percentage. All bold is added by me.

This database was last updated on:

20 May 2020

This Sheet includes what I'm calling "existential-risk-level estimates", meaning estimates of things like the existential risk from something, the extinction risk from something, or how much something reduces the expected value of the future.

I'm not including here estimates of "less extreme" things, like major global catastrophes that don't cause existential catastrophe; some such estimates can be found in the Sheet "Estimates of somewhat less extreme outcomes". (Of course, such events may still be very extreme by regular standards, and I don't mean to imply otherwise.)

I'm also mostly trying to include "unconditional" estimates here, and "conditional" estimates in the next Sheet. For example, I'd include here an estimate of the chance of existential catastrophe as a result of nuclear war, but not **separate** estimates of the chance that a nuclear war occurs and the chance that a nuclear war, **if** it occurred, would cause existential catastrophe.

Who is the estimator?	When was the estimate made/published?	What is the estimator estimating?	What is their estimate?	Num. Years	% Probability	Annualized %	Century %	Subjective Quality Weighting	Source	Have I properly read the source myself?	Is this estimate included in Beard et al.'s appendix?	Other notes
"Total risk" (or similar)												
Toby Ord	2020	"Total existential risk" by 2120	~17% (~1 in 6)	100	17%	0.19%	17.00%	1	The Precipice	Yes	No	Ord writes: "Don't take these numbers to be completely objective. [...] And don't take the estimates to be precise. Their purpose is to show the right order of magnitude, rather than a more precise probability."
GCR Conference	2008	"Overall risk of **extinction** prior to 2100"	19%	92	19%	0.23%	20.47%	10	https://www.fhi.o	Yes	Yes	This is the median. The report about these estimates also plots the results for each question "with individual response distributions visible" in Appendix A.
Will MacAskill	2019/2020	Existential risk in the 21st century	1%	80	1%	0.01%	1.25%	1	https://80000hou	Yes	No	
Ben Todd or 80,000 Hours	2017	Extinction risk "in the next century"	Probably **at or above** 3%	?		#VALUE!	#VALUE!	1	https://80000hou	Yes	No	
John Leslie	1996	Risk of **extinction** over the next five centuries	At or above 30%	500	30%	0.07%	6.89%	1	Leslie, J. (2002).	No	Yes	"The probability of the human race avoiding extinction for the next five centuries is encouragingly high, perhaps as high as 70 percent"
Martin Rees	2003	Odds that our present civilization on earth will survive to the end of the present century	50% ("no better than fifty-fifty")	97	50%	0.71%	51.06%	1	Rees, M. J. (200	No	Yes	
Metaculus responders		"there be zero living humans on planet earth on January 1, 2100"	Median: 1%. Mean: 8%.	80	1%	0.01%	1.25%	3	https://www.meta	No	Yes	That median and mean is as of 3rd July 2019.
Nick Bostrom	2002	"existential disaster will do us in"	Probably **at or above** 25%	?		#VALUE!	#VALUE!		https://www.nickb	No	Yes	
Gott III	1993	The prior probability that "humanity will cease to exist before 5,100 years or thrive beyond 7.8 million years" (quote from Beard et al.'s appendix)	5%.	?		#VALUE!	#VALUE!		Gott III, J. R. (19	No	Yes	
Wells	2009	Annual probability **as of 2009** of **extinction**	0.3-0.4%	1	0.35%	0.35%	29.57%	1	Wells, W. (2009)	No	Yes	
Simpson	2016	"Humanity's prognosis for the coming century is well approximated by a global catastrophic risk of 0.2% per year."	0.2%	1	0.20%	0.20%	18.14%	1	Simpson, F. (201	No	Yes	Beard et al. seem to imply this is about extinction, but the quote suggests it's about "global catastrophic risk".
Toby Ord	2020	"Humanity avoids every existential catastrophe and eventually fulfils its potential: achieving something close to the best future open to us"	50% (~1 in 2)	?					The Precipice	Yes	No	
I will **not** report the number from Stern Review: The economics of climate change, as it was **not** actually meant as an estimate.												
					Mean	0.22%	18.20%					
					Median	0.19%	17.57%					
AI												
Toby Ord	2020	Existential catastrophe by 2120 as a result of "unaligned AI"	~10%						The Precipice	Yes	No	
Global Catastrophic Risk Conference	2008	Human **extinction** by 2100 as a result of "superintelligent AI"	5%						https://www.fhi.o	Yes	Yes	This is the median. Beard et al.'s appendix says "Note that for these predictions no time frame was given." I think that that's incorrect, based on phrasings in the original source, but I'm not certain.
Survey of "AI experts"	2017	"Extremely bad (e.g. extinction)" long-run impact on humanity from "high-level machine intelligence"	5%						https://arxiv.org/a	Yes	Yes	The report's authors discuss potential concerns around non-response bias and the fact that "NIPS and ICML authors are representative of machine learning but not of the field of artificial intelligence as a whole". There was also evidence of apparent inconsistencies in estimates of AI timelines as a result of small changes to how questions were asked, providing further reason to wonder how meaningful these experts' predictions were. https://web.archive.org/web/20171030220008/https://aiimpacts.org/some-survey-results/
Pamlin & Armstro	2015	"Infinite impact" from AI within the next 100 years, which "refers to the	0-10%						Pamlin, D. & Arm	No	Yes	
Rohin Shah	2020	Chance that AI, through "adversarial optimization against humans only", will cause existential catastrophe	~5%						https://www.lessw	Yes	No	This is my interpretation of some comments that may not have been meant to be taken very literally. Elsewhere, Rohin noted that this was "[his] opinion before updating on other people's views": https://forum.effectivealtruism.org/posts/tugs9KQyNqi4yRTsb/does-80-000-hours-focus-too-much-on-ai-risk#ZmtPji3pQaZK7Y4FF I think he updated this in 2020 to ~9%, due to pessimism about discontinuous scenarios: https://www.lesswrong.com/posts/TdwpN484eTbPSvZkm/rohin-shah-on-reasons-for-ai-optimism?commentId=n577gwGB3vRpwkBmj Rohin also discusses his estimates here: https://futureoflife.org/2020/04/15/an-overview-of-technical-ai-alignment-in-2018-and-2019-with-buck-shlegeris-and-rohin-shah/
Paul Christiano	2019	Amount by which risk of failure to align AI (using only a narrow conception of alignment) reduces the expected value of the future	~10%						https://aiimpacts.	Yes	No	He also says "I made up 10%, it's kind of a random number." And "All of the numbers I'm going to give are very made up though. If you asked me a second time you'll get all different numbers."
Buck Schlegris	2020	"the probability of AI-induced existential risk" (but from context, I believe this actually meant the probability of an AI-induced existential *catastrophe*)	50%						https://futureoflife	Yes	No	
James Fodor	2020	Existential risk from unaligned AI over the coming 100 years	0.05%						Critical Review o	Yes	No	This was a direct response to Ord's estimate. It focuses on one pathway to x-risk from AI, not all pathways (e.g., not AI misuse or risks from competition between powerful AIs). "These estimates should not be taken very seriously. I do not believe we have enough information to make sensible quantitative estimates about these eventualities. Nevertheless, I present my estimates largely in order to illustrate the extent of my disagreement with Ord's estimates, and to illustrate the key considerations I examine in order to arrive at an estimate." In comments on the source, Rohin Shah critiques some of the inputs to this estimate, and provides his own, substantially higher estimates.
Stuart Armstrong	2014	Chance of humanity not surviving AI	50, 40, or 33%						https://www.youtu	Yes	No	Stated verbally during an interview. Not totally clear precisely what was being estimatined (e.g. just extinction, or existential catastrophe more broadly?). "This number fluctuates a lot". He indicated he thought we had a 2/3 chance of surviving, then said he'd adjust to 50%, which is his number for an "actually superintelligent" AI, whereas for "AI in general" it'd be 60%.
Biorisk												
Toby Ord	2020	Existential catastrophe from "engineered pandemics" by 2120	~3% (~1 in 30)						The Precipice	Yes	No	
GCR Conference	2008	Human **extinction** by 2100 as a result of "the single biggest **natural** pandemic"	0.05%						https://www.fhi.o	Yes	Yes	This is the median. Beard et al.'s appendix says "Note that for these predictions no time frame was given." I think that that's incorrect, based on phrasings in the original source, but I'm not certain.
Toby Ord	2020	Existential catastrophe from "**'naturally' arising pandemics**" by 2120	~0.01% (~1 in 10,000)						The Precipice	Yes	No	
GCR Conference	2008	Human **extinction** by 2100 as a result of "single biggest **engineered** pandemic"	2%						https://www.fhi.o	Yes	Yes	This is the median. Beard et al.'s appendix says "Note that for these predictions no time frame was given." I think that that's incorrect, based on phrasings in the original source, but I'm not certain.
Millet & Snyder-Beattie	2017	The annual probability of an existential catastrophe arising from a global pandemic	0.008% to 0.0000016% (between 8 x 10-5 and 1.6 x 10-8)						Millett, P., & Snyd	No	Yes	The fact that there's a separate estimate from the same source for biowarfare and bioterrorism suggests to me that this is meant to be an estimate of the risk from a natural pandemic only. But I'm not sure. This might also include "accidental" release of a bioengineered pathogen.
Millet & Snyder-Beattie	2017	The annual probability of an existential catastrophe arising from biowarfare or bioterrorism	0.00019% (0.0000019)						Millett, P., & Snyd	No	Yes	
Pamlin & Armstro	2015	"Infinite impact" from a global pandemic within the next 100 years, whi	0.0001%						Pamlin, D. & Arm	No	Yes	The fact that there's a separate estimate from the same source for "synthetic biology" suggests to me that this is meant to be an estimate of the risk from a natural pandemic only.
Pamlin & Armstro	2015	"Infinite impact" from "synthetic biology" within the next 100 years, whi	0.0001%						Pamlin, D. & Arm	No	Yes	
James Fodor	2020	Extinction risk from engineered pandemics over the coming 100 years	0.0002%						Critical Review o	Yes	No	This was a direct response to Ord's estimate, although this estimate is of extinction risk rather than existential risk. "These estimates should not be taken very seriously. I do not believe we have enough information to make sensible quantitative estimates about these eventualities. Nevertheless, I present my estimates largely in order to illustrate the extent of my disagreement with Ord's estimates, and to illustrate the key considerations I examine in order to arrive at an estimate." In comments on the source, Will Bradshaw critiques some of the inputs to this estimate.
Nanotechnology												
GCR Conference	2008	Human **extinction** by 2100 as a result of "molecular nanotech weapons"	5%						https://www.fhi.o	Yes	Yes	This is the median. Beard et al.'s appendix says "Note that for these predictions no time frame was given." I think that that's incorrect, based on phrasings in the original source, but I'm not certain.
GCR Conference	2008	Human **extinction** by 2100 as a result of "the single biggest nanotech accident"	0.5%						https://www.fhi.o	Yes	Yes	This is the median. Beard et al.'s appendix says "Note that for these predictions no time frame was given." I think that that's incorrect, based on phrasings in the original source, but I'm not certain.
Pamlin & Armstro	2015	"Infinite impact" from nanotechnology within the next 100 years, which	0.0100%						Pamlin, D. & Arm	No	Yes	
Toby Ord	2020	Existential catastrophe from **"other anthropogenic risks"** (which **includes but is not limited to** nanotechnology) by 2120	~2% (~1 in 50)						The Precipice	Yes	No	See this post for some commentary: https://forum.effectivealtruism.org/posts/Z5KZ2cui8WDjyF6gJ/my-thoughts-on-toby-ord-s-existential-risk-estimates#_Unforeseen__and__other__anthropogenic_risks__Surprisingly_risky_
Nuclear												
Toby Ord	2020	Existential catastrophe from nuclear war by 2120	~0.1% (~1 in 1000)						The Precipice	Yes	No	
GCR Conference	2008	Human **extinction** by 2100 as a result of "nuclear wars"	1%						https://www.fhi.o	Yes	Yes	This is the median. Beard et al.'s appendix says "Note that for these predictions no time frame was given." I think that that's incorrect, based on phrasings in the original source, but I'm not certain.
GCR Conference	2008	Human **extinction** by 2100 as a result of "acts of nuclear terrorism"	0.03%						https://www.fhi.o	Yes	Yes	This is the median. Beard et al.'s appendix says "Note that for these predictions no time frame was given." I think that that's incorrect, based on phrasings in the original source, but I'm not certain.
Ben Todd or 80,0	2017	The chance of "a civilization-ending nuclear war" (or one that causes e	Probably **at or above** 0.3%						https://80000hou	Yes	No	It's also worth noting that Todd uses the phrases "permanently end civilization" and "civilization-ending nuclear war" here, but the article and some other estimates in it are focused on extinction, rather than other scenarios like permanent collapse. I'm thus not sure precisely what scenarios Todd is estimating (or broadly discussing) the risk of.
Dave Denkenberg	2018	"**Reduction in far future potential** per year from [the risk of a?] full sc	~0.29%						https://www.getgu	Sort-of	No	There's some relevant info/discussion in these three sources (especially the first two): https://forum.effectivealtruism.org/posts/CcNY4MrT5QstNh4r7/cost-effectiveness-of-foods-for-global-catastrophes-even https://80000hours.org/podcast/episodes/david-denkenberger-allfed-and-feeding-everyone-no-matter-what/ https://forum.effectivealtruism.org/posts/D8t5TtSemqcGa4g7Y/my-open-for-feedback-donation-plans#
Anders Sandberg	2018	"**Reduction in far future potential** per year from [the risk of a?] full sc	~0.051%						https://www.getgu	Sort-of	No	
Peter McIntyre	2016	Amount by which **"the future potential of humanity" is reduced** due	0.6-6%						https://80000hou	Yes	No	I multiplied 2-20% by 30%. I may be misinterpreting his statement. He writes "200 years", but 80,000 Hours seem to typically discuss risks either during the 21st century or over the next 100 years, so I'm not sure if "200" was intentional or a typo.
Alexey Turchin	2008	The risk of extinction due to the consequences of nuclear war, or as a	In the order of 1%						Turchin, A. V. (20	No	Yes	Beard et al.'s appendix makes it seem somewhat unclear what this is about or how to interpret. See that appendix for details.
Pamlin & Armstro	2015	"Infinite impact" from nuclear war within the next 100 years, which "ref	0.005%						Pamlin, D. & Arm	No	Yes	
Climate change												
Toby Ord	2020	Existential catastrophe from climate change by 2120	~0.1% (~1 in 1000)						The Precipice	Yes	No	
Roman Duda	2016	Amount by which "the future potential of humanity" is reduced due to the risk of "extreme climate change (>5ºC) in the next 100 years"	~0.1-2%						https://80000hou	Yes	No	I multiplied 0.5-10% by ~20%. I may be misinterpreting his statement.
Pamlin & Armstro	2015	"Infinite impact" from climate change within the next 100 years, which	0.01%						Pamlin, D. & Arm	No	Yes	
Toby Ord	2020	Amount by which existential risk till 2120 would decrease if "we could just somehow have the next century but make it so that climate change wasn't an issue"	0.1-1 percentage points						https://80000hou	Yes	No	"Toby Ord: [...] I think that if we imagine a world, if we could just somehow have the next century but make it so that climate change wasn't an issue. All of the dedicated altruists who are working on fighting climate change could then work on other things and global international tensions on this would go down and so nations could spend their "altruistic international corporation" kind of budget on something else. So I It's possible he means 0.1-1% "of the total risk level over the next 100 years", which would correspond to ~0.017-0.17 percentage points. But he estimates climate change itself/directly poses a 1/1000 risk, and says that it's bigger as a risk factor, which only makes sense if he meant 0.1-1 percentage points of risk from climate change as a risk factor.
Natural risks (excluding natural pandemics)												
Toby Ord	2020	Existential catastrophe from supervolcanic eruption by 2120	~0.01% (~1 in 10,000)						The Precipice	Yes	No	
Pamlin & Armstro	2015	"Infinite impact" from a super-volcano within the next 100 years, which	0.00003%						Pamlin, D. & Arm	No	Yes	
Toby Ord	2020	Existential catastrophe from asteroid or comet impact by 2120	~0.0001% (~1 in 1,000,000)						The Precipice	Yes	No	
Pamlin & Armstro	2015	"Infinite impact" from "a major asteroid impact" within the next 100 yea	0.00013%						Pamlin, D. & Arm	No	Yes	 Pamlin & Armstrc 2015 "Infinite impact" f 0.01% Pamlin, D. & Arm No Yes Pamlin & Armstrc 2015 "Infinite impact" f 0.01% Pamlin, D. & Arm No Yes
Toby Ord	2020	Existential catastrophe from stellar explosion by 2120	~0.0000001% (~1 in 1,000,000,000)						The Precipice	Yes	No	
Toby Ord	2020	"Total natural [existential] risk" by 2120	~0.01% (~1 in 10,000)						The Precipice	Yes	No	
Snyder-Beattie, Ord, & Bonsall	2019	The probability that humanity goes extinct from natural causes in any g	Almost guaranteed to be below 0.007% (1 in 14000), likely less than 0.001%, likely less than 0.0001% if we use th						https://www.natur	No	Yes	
Miscellaneous												
Toby Ord	2020	Existential catastrophe from "other environmental damage" (not climat	~0.1% (~1 in 1,000)						The Precipice	Yes	No	
Pamlin & Armstro	2015	"Infinite impact" from an "ecological catastrophe" within the next 100 y	0.5%						Pamlin, D. & Arm	No	Yes	
Toby Ord	2020	Existential catastrophe from "unforeseen anthropogenic risks" by 2120	~3% (~1 in 30)						The Precipice	Yes	No	See this post for some commentary: https://forum.effectivealtruism.org/posts/Z5KZ2cui8WDjyF6gJ/my-thoughts-on-toby-ord-s-existential-risk-estimates#_Unforeseen__and__other__anthropogenic_risks__Surprisingly_risky_
Toby Ord	2020	Existential catastrophe from "other anthropogenic risks" by 2120 (see	~2% (~1 in 50)						The Precipice	Yes	No	See this post for some commentary: https://forum.effectivealtruism.org/posts/Z5KZ2cui8WDjyF6gJ/my-thoughts-on-toby-ord-s-existential-risk-estimates#_Unforeseen__and__other__anthropogenic_risks__Surprisingly_risky_
Toby Ord	2020	"Total anthropogenic [existential] risk" by 2120	~17% (~1 in 6)						The Precipice	Yes	No	
Toby Ord	2020	Amount by which existential risk till 2120 would decrease if there definitely wouldn't be a great power war during that time	~1.7 percentage points ("something like a tenth" of the ~1/6 total risk over that time)						The Precipice	Yes	No	"Consider your own estimate of how much existential risk there is over the next hundred years. How much of this would disappear if you knew that the great powers would not go to war with each other over that time? It is impossible to be precise, but I'd imagine that an appreciable fraction would disappear - something like a tenth of the existential risk over that time."
GCR Conference	2008	Human **extinction** by 2100 as a result of "wars (including civil wars)"	4%						https://www.fhi.o	Yes	Yes	This is the median. Beard et al.'s appendix says "Note that for these predictions no time frame was given." I think that that's incorrect, based on phrasings in the original source, but I'm not certain.
Pamlin & Armstro	2015	"Infinite impact" from "an uncertain risk" within the next 100 years, whi	0.5%						Pamlin, D. & Arm	No	Yes	Beard et al. write: "The following estimate concerns risks associated with scenarios that at present seem either very unlikely or very unlikely to pose a significant risk, but where there is a possibility that these assessments represent a significant underestimate of the threat. This includes physics experiments, as described in the previous section, but also risks such as animal cognitive enhancement and the search for extraterrestrial intelligent life. Pamlin and Armstrong also discuss the existential threat posed by Global System Collapse and Future Bad Global Governance, but believe that the probability of even reaching infinite threshold from such risks cannot be quantified." Pamlin & Armstrc 2015 "Infinite impact" f 0.5% Pamlin, D. & Arm No Yes Pamlin & Armstrc 2015 "Infinite impact" f 0.5% Pamlin, D. & Arm No Yes
Dave Denkenberg	2018	"**Reduction in far future potential** due to [the risk of a?] 10% agricul	0.15%						https://www.getgu	Sort-of	No	There's some relevant info/discussion in these three sources (especially the first two): https://forum.effectivealtruism.org/posts/CcNY4MrT5QstNh4r7/cost-effectiveness-of-foods-for-global-catastrophes-even https://80000hours.org/podcast/episodes/david-denkenberger-allfed-and-feeding-everyone-no-matter-what/ https://forum.effectivealtruism.org/posts/D8t5TtSemqcGa4g7Y/my-open-for-feedback-donation-plans#
Anders Sandberg	2018	"**Reduction in far future potential** due to [the risk of a?] 10% agricul	~0.00023%						https://www.getgu	Sort-of	No	
How various actions may reduce certain risks												
Ben Todd or 80,0	2017	**One plausible** amount by which "$100 billion spent on reducing extinc	1 percentage point						https://80000hou	Yes	No	I'm not sure where this "1 percentage point reduction" lies on the continuum from "genuine guess" to "number used purely for the sake of illustration".
Paul Christiano	2019	Amount by which "really nailing" some portion of AI safety work could i	0.05						https://aiimpacts.	Yes	No	He says "All of the numbers I'm going to give are very made up though. If you asked me a second time you'll get all different numbers."
Paul Christiano	2019	Amount by which "a marginal person" doing some portion of AI safety	0.005% ("one in 20,000 or something")						https://aiimpacts.	Yes	No	He says "All of the numbers I'm going to give are very made up though. If you asked me a second time you'll get all different numbers."

This is for conditional estimates of existential risk (or similar), such as estimates of the chance that, **if** a nuclear war occurs, that would cause existential catastrophe.

Who is the estimator?	**When was the estimate made/published?**	**What is the estimator estimating?**	**What is their estimate?**	**Source**	**Have I properly read the source myself?**	**Is this estimate included in Beard et al.'s appendix?**	**Other notes**
"Total risk" (or similar)							
AI							
Survey of "experts" in the AI field	2016	The probability that the long-run overall impact on humanity of human level machine intelligence will be "Extremely bad (existential catastrophe)", assuming HLMI will at some point exist.	18%	Müller, V. C., & B	No	Yes	This is the mean. According to Beard et al., the question was "4. Assume for the purpose of this question that such Human Level Machine Intelligence (HLMI) will at some point exist. How positive or negative would be overall impact on humanity, in the long run?"
Rohin Shah	2019	Chance that AI, through "adversarial optimization against humans only", will cause existential catastrophe, conditional on there not being "**additional** intervention by longtermists" (or perhaps "**no** intervention from longtermists")	~10%	https://www.lessv	Yes	No	This is my interpretation of some comments that may not have been meant to be taken very literally. I think he updated this in 2020 to ~15%, due to pessimism about discontinuous scenarios: https://www.lesswrong.com/posts/TdwpN484eTbPSvZkm/rohin-shah-on-reasons-for-ai-optimism?commentId=n577gwGB3vRpwkBmj Rohin also discusses his estimates here: https://futureoflife.org/2020/04/15/an-overview-of-technical-ai-alignment-in-2018-and-2019-with-buck-shlegeris-and-rohin-shah/
Rohin Shah	2019	Chance that AI, through "adversarial optimization against humans only", will cause existential catastrophe, **conditional on "discontinuous takeoff"**	~70% (but with "way m	https://www.lessv	Yes	No	
Toby Ord	2020	Chance that we don't "manage to survive that transition [to there bei	~20%	https://80000hou	Yes	No	This may have been specifically if the transition happens in the net 100 years; it's possible Ord would estimate we'd have a different chance if this transition happened at a later time. "Basically, you can look at my [estimate that the existential risk from AI in the next 100 years is] 10% as, there's about a 50% chance that we create something that's more intelligent than humanity this century. And then there's only an 80% chance that we manage to survive that transition, being in charge of our future. If you put that together, you get a 10% chance that's the time where we lost cont [For people who would disagree, a question] is why would they think that we have much higher than an 80% chance of surviving this 'passing this baton to these other entities', but still retaining control of our future or making sure that they build a future that is excellent, surpassingly good by our own perspective? I think that the very people who are working on trying to actually make sure that artific
Biorisk							
Nanotechnology							
Nuclear							
Toby Ord	2020	Chance that "a full-scale nuclear war in the next century" would "be	~2%	https://80000hou	Yes	No	"I give existential risk over the next century from nuclear war at about one in a thousand. I initially thought it would be higher than that. That's actually something that while researching the book, thought was a lower risk than I had initially thought. And how I'd break it down is to something like a 5% chance of a full-scale nuclear war in the next century and a 2% chance that that would be the end of
Luke Oman	2012	"The probability I would estimate for the global human population of zero resulting from the 150 Tg of black carbon scenario in our 2007 paper"	0.001-0.01% ("in the ra	http://www.overc	Yes	No	I **think** that this is Oman's estimate of the chance that extinction would occur if that black carbon scenario occurred, rather than an estimate that also takes into account the low probability that that black carbon scenario occurs. I.e., I think that this estimate was conditional on a particular type of nuclear war occurring. But I'm not sure about that, and the full context doesn't make it much clearer.
Climate change							
Natural risks (excluding natural pandemics)							
Miscellaneous							

This Sheet is for estimates that I don't count as "existential-risk-level estimates". These may be estimates of, for example, the likelihood of Spanish-Flu-level pandemics, nuclear war, catastrophes causing over a billion deaths, etc.									**Sources from Beard et al.'s appendix that would be relevant here:**	**Other relevant sources**
Most of these events are of course still "extreme" by most standards.									1	https://www.metaculus.com/questions/2568/ragnar%25C3%25B6k-question-series-results-so-far/
I haven't yet taken the time to include many estimates of this type, because such estimates are further from the key thing I'm most concerned about (existential risk), and because it seems that it's much less hard to find those than to find what I'm considering "existential risk estim									12	Adam Gleave gave an answer to the question "what is the chance that advanced artificial intelligence poses a significant risk of harm?". But it's very unclear to me whether this means the chance that harm *occurs* vs the chance that a *risk is posed*, somewhat unclear what the latter would mean, and very unclear what level of harm is being discussed. So I haven't included the numbers here. https://aiimpacts.org/conversation-with-adam-gleave/
									13	Kevin Esvelt answered a relevant question at the 21:53 minute mark of this video: https://youtu.be/BbOHQLrVSX4?t=1313
Who is the estimator?	**When was the estimate made/published?**	**What is the estimator estimating?**	**What is their estimate?**	**Source**	**Have I properly read the source myself?**	**Is this estimate included in Beard et al.'s appendix?**	**Other notes**		14	
"Total risk" (or similar)									15	
									16	
									18 (for "infinite threshold")	
									20	
									21	
AI									22	
Adam Gleave	2019	Chance that AI safety is as hard as a (caricature of) MIRI suggests	~10%	https://aiimpacts.	Yes	No	"So, decent chance– I think I put a reasonable probability, like 10% probability, on the hard-mode M		23	
Adam Gleave	2019	Chance that "AI safety basically [doesn't need] to be solved, we'll just solve it by default unless we're completely completely careless"	~20-30%	https://aiimpacts.	Yes	No			24	
Rohin Shah	2020	"chance that the first thing we try just works and we don't even need	~30%	https://futureoflife	Yes	No	"There's some chance that the first thing we try just works and we don't even need to solve any so		25	
Buck Schlegris	2020	Chance we have "good competitive alignment techniques by the tim	~30%	https://futureoflife	Yes	No	"I haven't actually written down these numbers since I last changed my mind about a lot of the inpu		26	
Toby Ord	2020	"chance that we create something that's more intelligent than humar	~50%	https://80000hou	Yes	No	"Basically, you can look at my [estimate that the existential risk from AI in the next 100 years is] 10 Toby Ord: With that number, I've spent a lot of time thinking about this. Actually, my first degree wa		29	
									31 (for "infinite threshold")	
Biorisk									32	
									33	
									34	
Nanotechnology									35	
									36	
									37 (for "infinite threshold")	
									40	
Nuclear									42 (for "infinite threshold")	
Toby Ord	2020	"chance of a full-scale nuclear war in the next century"	~5%	https://80000hou	Yes	No	"I give existential risk over the next century from nuclear war at about one in a thousand. I initially t		43	
									44	
									45	
Climate change									46	
									47 (perhaps this in fact belongs in the main sheet; I haven't read the source)	
									48	
									49 (for "infinite threshold")	
Natural risks (excluding natural pandemics)									50	
									51	
									52	
									53 (for "infinite threshold")	
Miscellaneous									54 (for "infinite threshold")	
									55 (for "infinite threshold")	
									57 (for "infinite threshold")	
									58	
									59	
									60	
									61	
									62	
									63 (perhaps this in fact belongs in the main sheet; I haven't read the source)	
									64 (perhaps this in fact belongs in the main sheet; I haven't read the source)	
									65 (perhaps this in fact belongs in the main sheet; I haven't read the source)	
									66 (perhaps this in fact belongs in the main sheet; I haven't read the source)	
									67 (for "infinite threshold")	

Source	Description
https://wiki.lesswror	Apparently during (or from?) 2011, various people were asked questions including "What probability do you assign to the possibility of human extinction within 100 years as a result of AI capable of self-modification (that is not provably non-dangerous, if that is even possible)? P(human extinction by AI | AI capable of self-modification and not provably non-dangerous is created)". I haven't looked into these interviews yet.
https://forum.effecti	Gregory Lewis "[threw] together a guestimate as a first-pass estimate" of "the expected value of x-risk reduction by the lights of person affecting views". But I think he was just using "reasonable to conversative" assumptions in order to build an illustrative model of what the value of x-risk reduction would be under those conditions (see footnote 23 here: https://80000hours.org/problem-profiles/global-catastrophic-biological-risks/)
https://founderspledge.com/research/fp-existential-risk	John Halstead gives numbers related to total existential risk this century, existential risk from nuclear war, and existential risk from advanced machine intelligence. However, from a skim, I think these are a mixture of illustrative examples and just reporting on other estimates which I mention in this post.
https://www.getguesstimate.com/models/11762	In Denkenberger's model, he presents a modified version of an earlier model from others, which is focused on AI. I haven't included that here because I'm not sure whether Denkenberger actually endorses the values he provides in it, or if he was intentionally being "conservative" (i.e., using high estimates of AI risk) to make it harder for him to "make the case" for his own, non-AI intervention.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3155983	This paper may have relevant estimates; I haven't read it.
https://forum.effectivealtruism.org/posts/dFYDjr3spMn59zTwP/causal-network-model-iii-findings	This post may have relevant estimates; I haven't read it.
https://www.youtube.com/watch?v=tOSpj19eows and footnote 2 of this: https://80000hours.org/articles/extinction-risk/	Spencer Greenberg surveyed users of Mechanical Turk, and some EAs, and asked about their estimates of the chance of human extinction within 50 from the time the survey was taken (around 2017, I believe). I haven't included this as I'm less interested in the general public's estimates of existential risk. But perhaps I should include it anyway, or at least the EAs' estimates.
https://forum.effectivealtruism.org/posts/rmKwTzX5XLGs94fcB/public-opinion-about-existential-risk	"An MTurk study of people in the United States (N=395) found median estimates of 1%, 5%, and 20% for the chance of human extinction in 50, 100, and 500 years, respectively. People were fairly confident in their answers and tended to think the government should prioritize preventing human extinction more than it currently does." But I haven't read the post, and, as stated above, I'm not necessarily especially interested in the general public's estimates.
https://forum.effectivealtruism.org/posts/XXLf6FmWujkxna3E6/are-we-living-at-the-most-influential-time-in-history-1	MacAskill writes: "Quantitatively: These considerations push me to put my posterior on [the idea that we're in the hinge of history] into something like the [1%, 0.1%] interval. But this credence interval feels very made-up and very unstable."
https://forum.effectivealtruism.org/posts/frYYAAa5K4nHqRCPG/current-estimates-for-likelihood-of-x-risk?commentId=xobj9qdnGRoQqjyRS	"Anders Sandberg's Flickr account has a 2014 photo of a whiteboard from FHI containing estimates for [various relevant statements/questions". But I'm not sure how seriously these estimates were meant to be taken.
https://www.getguesstimate.com/models/11762	Denkenberger's model also provides estimates of how both ALLFED's work so far and alternative foods planning and R&D may protect against these reductions in far future potential. Perhaps that should be included in the section for estimates of "How various actions may reduce certain risks".
https://80000hours.org/problem-profiles/nuclear-security/	My interpretation is that McIntye suggests that a "major effort at the problem" could recover 30% of the 0.6-6% of the "future potential of humanity" that is currently reduced by the risk of nuclear war. In other words, I interpret him as effectively suggesting that such an effort would "increase" the "future potential of humanity" from 94-99.4% of what it would be if there was no risk of such a war to 95.8-99.58% of what it would be. But I'm unsure if he really meant that, and I'm very unsure how to interpret that anyway (among other things, what classifies as a "major effort"?). Thus, I haven't included this in the section for estimates of "How various actions may reduce certain risks". But perhaps I should.
https://80000hours.org/problem-profiles/climate-change/	My interpretation is that Duda suggests that a "major effort at the problem" could recover 50% of the 0.1-2% of the "future potential of humanity" that is currently reduced by the risk of extreme climate change. In other words, I interpret him as effectively suggesting that such an effort would "increase" the "future potential of humanity" from 98-99.9% of what it would be if there was no risk of such climate change to 99-99.95% of what it would be. But I'm unsure if he really meant that, and I'm very unsure how to interpret that anyway (among other things, what classifies as a "major effort"?). Thus, I haven't included this in the section for estimates of "How various actions may reduce certain risks". But perhaps I should.
The Precipice (Toby	He writes that his estimates already: "incorporate the possibility that we get our act together and start taking these risks very seriously. Future risks are often estimated with an assumption of 'business as usual': that our levels of concern and resources devoted to addressing the risks stay where they are today. If I had assumed business as usual, my risk estimates would have been substantially higher. But I think they would have been misleading, overstating the chance that we actually suffer an existential catastrophe. So instead, I've made allowances for the fact that we will likely respond to the escalating risks, with substantial efforts to reduce them. The numbers therefore represent my actual best guesses of the chance the threats materialise, taking our responses into account. If we outperform my expectations, we could bring the remaining risk down below these estimates. Perhaps one could say that we were heading towards Russian roulette with two bullets in the gun, but that I think we will remove one of these before it's time to pull the trigger. And there might just be time to remove the last one too, if we really try." One could perhaps interpret this as him saying that his estimate of existential risk by 2120 conditional on continued business as usual is 2/6, and his estimate of how much existential risk by 2120 we can reduce if "we really try" is ~100% of it (meaning ~17 percentage points). But it's not clear to me that these were actually meant as estimates, rather than somewhat poetic illustrations of certain points.
https://scripps.ucsd.	"Researchers identify a one-in-20 chance of temperature increase causing catastrophic damage or worse by 2050". In places, that page seems to imply this is a 1-in-20 chance of extinction, but I think that's just using ambiguous language in a dramatic way, and I was just skimming.
https://futureoflife.o	Andrew Critch: "In the same way you can predict an eclipse, you can predict that an asteroid is almost certainly not going to cause human extinction this century. I would bet, 99.99% chance, that we will not be extinct from an asteroid impact." But I think he just threw that number out there to make a point, and I wouldn't be surprised if his real estimate would be notably different. It's also unclear precisely what he's predicting (the chance of extinction from an asteroid this century? the chance of extinction from an asteroid this century conditional on us avoiding other problems, like AI catastrophe? the chance of extinction from an asteroid ever?).