Petrov Day Organizer Guide

James Babcock (jimrandomh@gmail.com)

September 26

Content note: This event is designed to provoke existential terror and involves staring into the abyss.

1 Overview

Petrov Day is a yearly event on September 26 commemorating the anniversary of the Petrov incident, where a false alarm in the Soviet early warning system nearly set off a nuclear war. The purpose of the ritual is to make catastrophic and existential risk emotionally salient, by putting it into historical context and providing positive and negative examples of how it has been handled. This is not for the faint of heart and not for the uninitiated; it is aimed at those who already know what catastrophic and existential risk is, have some background knowledge of what those risks are, and believe (at least on an abstract level) that preventing those risks from coming to pass is important.

2 How to Organize

It's designed for groups of 5-10 people. Because of the participatory elements, everyone needs to fit at the same table, so if you have more people than that then you should split into groups, ideally out of earshot of each other. If you have more people and don't have space to split up comfortably, you might want to encourage some of them to also hold Petrov Day ceremonies at their own houses.

Running Petrov Day is pretty easy. You need to invite people over, acquire a few simple props, and print one copy of the program for each person. You don't have to write or rehearse anything, and once things get started, you'll be doing the same things as any other participant. You might want to read through the program in advance, but this isn't required.

When you invite people to attend, you should also clearly specify the schedule. Dinner followed by ceremony works well, but make sure the food is all cleared away before you start. Also, you should warn people not to arrive in the middle. Here is a sample email you might use to announce that you are hosting Petrov Day:

Dear friends,
On September 26th, 1983, the world was nearly destroyed by nuclear war. That day is Petrov Day, named for the man who averted it. I will be hosting a ritual commemorating the occasion, on September 26th at ⟨ADDRESS⟩. We will gather for dinner at 6:30pm (please bring a dish) and begin the ritual at 8pm. It will last for about an hour; please do not arrive in the middle.
Sincerely,
⟨YOUR NAME⟩

3 Materials

You will need:

- One complete printout of the program for each person
- A table with enough chairs to seat everyone
- A candle holder
- 8 candles and a lighter
- A fire extinguisher close enough to retrieve if needed
- A deck of small index cards or a pad of post-it notes, and some pens

The candle-holder must hold at least eight candles. A Menorah will work, but it shouldn't have symbols or iconography from Hanukkah or any other holiday. You might want to put down aluminum foil to catch dripping wax. Also, you want candles that won't burn too fast.

- Candle holder: http://www.amazon.com/dp/B000BWPESK
- Candles: http://www.amazon.com/gp/product/B003U6ZVHS

There are two versions of the program, one with the pages in order for single-sided printing, the other with the pages rearranged to print two-sided and fold in the middle. This PDF file is the single-sided printing version, which is better for reading on a computer screen, but the two-sided version is preferred for printing on paper. Print it, stack the pages, fold them in half, and staple the spine. Staple down into the front cover from the outside, 1/2 cm from the fold, at the top, middle, and bottom.

Petrov Day

September 26

By James Babcock

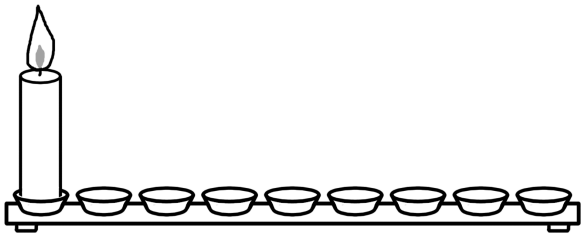
Stage directions are written in italics, like this. All other text is to be read aloud. Whenever there is a horizontal line, it becomes the next person's turn to speak, going clockwise. When reading quotes, you don't need to read the name and date at the end.

This day, September 26, is Petrov Day. In 1983, the story of humanity nearly ended. We're gathered here to remember that moment, and others like it. But to really feel the magnitude of those events, we need to visit them in their proper context. Let us begin the story of human history, starting from the beginning.

> "In the beginning, the universe was created. This has made a lot of people very angry, and been widely regarded as a bad move."
>
> — Douglas Adams

Let's fast forward over the thirteen billion year long prequel. Our story begins in the age of myth, of fossils and legends. It starts with the invention of fire.

> "I've hunted down and stolen, inside the hollow of a fennel's stalk, the seed of fire, a gift that has proven itself to be the teacher of every craft and the greatest resource for humans. Such is the crime I have committed and this is the penalty I am to suffer: nailed and chained on this rock beneath the open sky."
>
> — Prometheus Bound



Light the left-most candle, to represent the invention of fire. Point out the location of the nearest fire extinguisher, then dim or turn off all other lights in the room.

Depending which archaeologists you ask, fire was first used by either Homo Erectus of Homo Ergaster, some time between 400 thousand and 1.7 million years ago. Cooking is believed to have enabled larger, more energy-intensive brains, allowing the evolution of increased intelligence, and language.

> "Most species do their own evolving, making it up as they go along, which is the way Nature intended. And this is all very natural and organic and in tune with mysterious cycles of the cosmos, which believes that there's nothing like millions of years of really frustrating trial and error to give a species moral fiber and, in some cases, backbone."
>
> — Terry Pratchett

Homo Ergaster

Homo Erectus

Homo Florensiensis

Homo Habilis

Homo Heidelbergensis

Homo Neanderthalis

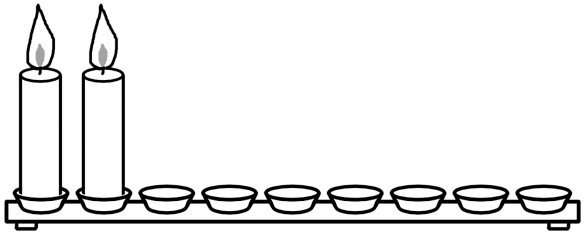
Homo Rudolfensis

"It certainly is not a true instinct, for every language has to be learnt. It differs, however, widely from all ordinary arts, for man has an instinctive tendency to speak, as we see in the babble of our young children; whilst no child has an instinctive tendency to brew, bake, or write."

— Charles Darwin, Descent of Man (1871)

Take the first candle, which represents the invention of fire. Use it to light the second candle, which represents the evolution of language.

Pass the candle once all the way around the circle. When you hold the candle, it is your turn to speak. What is your name, and when (what year) is your earliest memory?



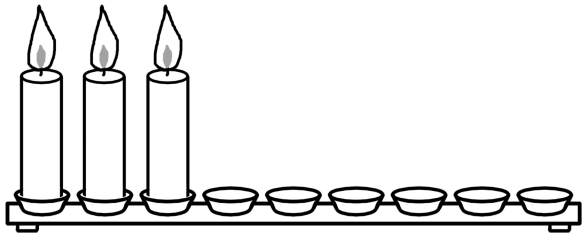
When everyone has spoken, put the candle back in the candelabrum.

Language is the first key to technology; with it, early humans could accumulate knowledge, not just in genes, but also in sayings and traditions.

They gave names to people around them. They gave names to species of animals and plants. They gave names to actions and to places and to strategies. They called some of these good, and called some of them bad. They learned to share their knowledge, and they learned to deceive each other. They built families and communities.

They began the long, slow process of taming the wilderness. Their tribes grew to cities. What became of them?

Take the second candle, which represents language. Use it to light the third candle, which represents agriculture.



Then, everyone sing together

Uplift

By Andrew Eigel

Hands chip the flint, light the fire, skin the kill
Feet move the tribe track the herd with a will
Mankind struggles in the cellar of history
Time to settle down, time to grow, time to breed

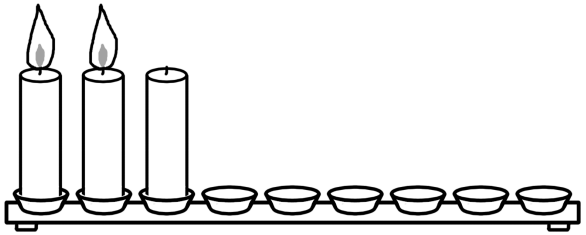
Plow tills the soil, plants the seed, pray for rain
Scythe reaps the wheat, to the mill, to grind the grain
Towns and cities spread to empire overnight
Hands keep building as we chant the ancient rite

"The power of population is so superior to the power of the earth to produce subsistence for man, that premature death must in some shape or other visit the human race. The vices of mankind are active and able ministers of depopulation. They are the precursors in the great army of destruction, and often finish the dreadful work themselves. But should they fail in this war of extermination, sickly seasons, epidemics, pestilence, and plague advance in terrific array, and sweep off their thousands and tens of thousands. Should success be still incomplete, gigantic inevitable famine stalks in the rear, and with one mighty blow levels the population with the food of the world."

— Thomas Malthus (1798)

Take the third candle, which represents agricultural society. Pass it around the circle.

Blow it out. Then return it to its place in the candelabrum.

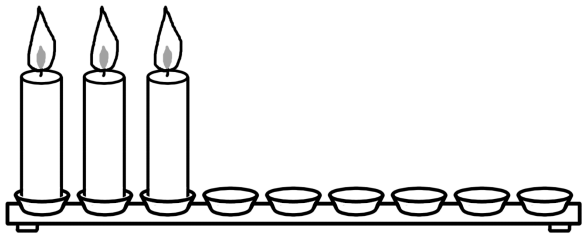


Mankind lived in equilibrium between growth and collapse, knowledge gained and knowledge forgotten. In that world, stories would last only as long as memory, monuments only as long as wood. For two hundred thousand years, nothing but genes survived.

But that was enough. Though they could not preserve knowledge over generations, they could preserve domesticated plants and animals. They saved the best, and little by little, the world got easier. And then a select few humans started writing, and the equilibrium between learning and forgetting was finally broken.

Of that age, what memories remain?

Using the second candle, which represents language, relight the third candle to represent the invention of writing.



I met a traveller from an antique land
Who said: Two vast and trunkless legs of stone
Stand in the desert. Near them, on the sand,
Half sunk, a shattered visage lies, whose frown,
And wrinkled lip, and sneer of cold command,
Tell that its sculptor well those passions read
Which yet survive, stamped on these lifeless things,
The hand that mocked them and the heart that fed:
And on the pedestal these words appear:
"My name is Ozymandias, king of kings:
Look on my works, ye Mighty, and despair!"
Nothing beside remains. Round the decay
Of that colossal wreck, boundless and bare
The lone and level sands stretch far away

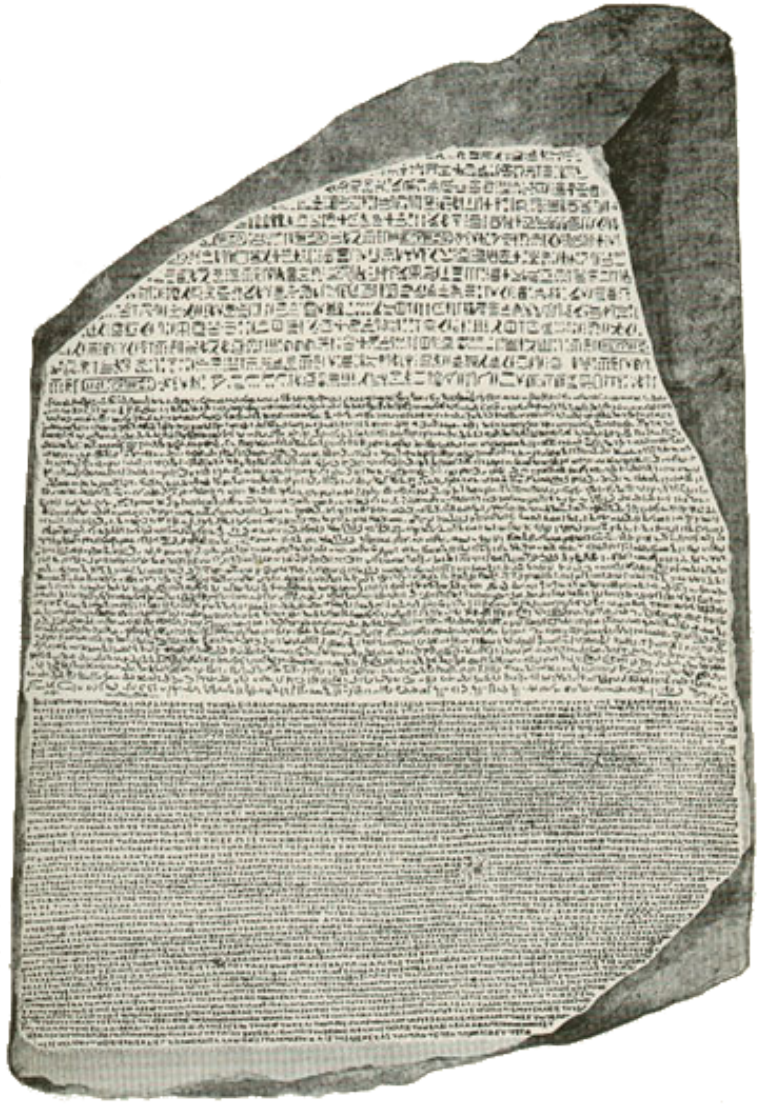
— Percy Bysshe Shelley (1818)

When you have finished reading, take a piece of paper and write down the name of the oldest family member - living or dead - that you can identify.

When everyone has written something, continue to the next page.

We know more about what the world was like after people started writing, but not very much survived. One of the most important writings was discovered by French soldiers in the wall of Fort Julien: the Rosetta Stone, important because it was written in three languages, two previously untranslatable. After a long string of honorifics and decrees about taxes and succession, it declares: there shall be a new holiday!

> "On these days in every month, on which there shall be sacrifices and libations and all the ceremonies customary at the other festivals, and the offerings shall be given to the priests who serve in the temples. And a festival shall be kept for King Ptolemy, the Ever-Living, the Beloved of Ptah, the God Epiphanes Eucharistos, yearly in the temples throughout the land from the 1st of Thoth for five days ... This decree shall be inscribed on a stela of hard stone in hieroglyphic and demotic and Greek characters and set up in each of the first, second, and third temples beside the image of the ever living king."
>
> — The Rosetta Stone (ca. 196 BC)



The majority of writing consisted of geneologies, legal codes, and fantastic stories. But some writing represented progress in philosophy and mathematics, eventually culminating in the invention of the scientific method.

> "Mathematics is the gate and key of the sciences... Neglect of mathematics works injury to all knowledge, since he who is ignorant of it cannot know the other sciences or the things of this world. And what is worse, men who are thus Ignorant are unable to perceive their own ignorance and so do not seek a remedy."
>
> — Roger Bacon, Opus Majus (1266)

Using the third candle, which represents writing, light the fourth candle to represent the scientific method.

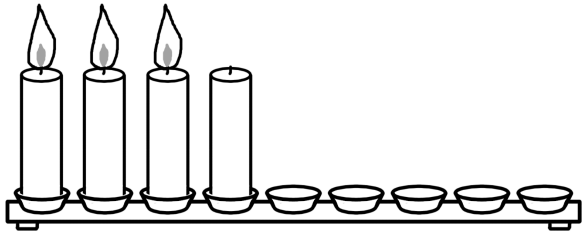
Then, everyone write down something surprising they learned about the world, and put it in the middle. When everyone has written something, continue to the next page.

The scientific method, combined with writing and a university system, marked the start of an accumulation of knowledge. This could have marked the beginning of a slow transition into the modern era. Instead, 81 years after Roger Bacon, history was derailed by a great plague.

Take the fourth candle, which represents the progress of science. Hold it, while you read the quote.

> "The seventh year after it began, it came to England and first began in the towns and ports joining on the seacoasts, in Dorsetshire, where, as in other counties, it made the country quite void of inhabitants so that there were almost none left alive. ... But at length it came to Gloucester, yea even to Oxford and to London, and finally it spread over all England and so wasted the people that scarce the tenth person of any sort was left alive."
>
> — Geoffrey the Baker, Chronicon Angliae (1360)

Blow out the candle. Then return it to its place on the candelabrum.



The plague killed about half the population of Europe during a four-year period, and it recurred repeatedly throughout the next three centuries killing double-digit percentages of the population each time. Between plagues, wars, and famines, there was little time to build or preserve knowledge.

Preserving knowledge required redundancy. In 1439, during the European Renaissance, Gutenberg perfected a device to do just that.

> "Pray, friend Martin, how many impressions can be made by this press in a day?" "About three hundred, if we work it constantly." "Is it possible!" exclaimed Peter. "Now indeed will books multiply. What will the plodding copyists say to this?"
>
> — Emily Clemens Pearson, Gutenberg and the Art of Printing (1870)

Take the fourth candle, which represents the progress of science.

Touch it to each of the other three candles in turn, until it is lit. Then return it to its place on the candelabrum.

Take the fourth candle, which represents science. Hold it, while you read the quote, then pass it directly to the next person. Repeat for each quote in this section.

> "By the aid of a telescope any one may behold this in a manner which so distinctly appeals to the senses that all the disputes which have tormented philosophers through so many ages are exploded at once by the indisputable evidence of our eyes, and we are freed from wordy disputes upon this subject, for the Galaxy is nothing else but a mass of innumerable stars planted together in clusters."
>
> — Galileo, The Starry Messenger (1610)

> "Matters that vexed the minds of ancient seers,
> And for our learned doctors often led
> to loud and vain contention, now are seen
> In reason's light, the clouds of ignorance
> Dispelled at last by science. Those on whom
> Delusion cast its gloomy pall of doubt,
> Upborne now on the wings that genius lends,
> May penetrate the mansions of the gods
> And scale the heights of heaven. O mortal men,
> Arise! And, casting off your earthly cares,
> Learn ye the potency of heaven-born mind,
> Its thought and life far from the herd withdrawn!"
>
> — Edmund Halley, preface to Newton's Principia Mathematica (1687)

"By calculations similar to these may be determined universally, what expectations are warranted by any experiments, according to the different number of times in which they have succeeded and failed; or what should be thought of the probability that any particular cause in nature, with which we have any acquaintance, will or will not, in any single trial, produce an effect that has been conjoined with it."

— Rev. Thomas Bayes, An Essay towards solving a Problem in the Doctrine of Chances (1763)

"I was thinking upon the engine at the time, and had gone as far as the herd's house, when the idea came into my mind that as steam was an elastic body it would rush into a vacuum, and if a communication were made between the cylinder and an exhausted vessel it would rush into it, and might be there condensed without cooling the cylinder. I then saw that I must get rid of the condensed steam and injection-water if I used a jet as in Newcomen's engine. Two ways of doing this occurred to me. ... I had not walked farther than the golf-house when the whole thing was arranged in my mind."

— James Watt (1765)

"I saw in a dream a table where all elements fell into place as required. Awakening, I immediately wrote it down on a piece of paper, only in one place did a correction later seem necessary."

— Dmitri Mendeleev (1864)

"I then shouted into the mouthpiece the following sentence: Mr. Watson, Come here, I want to see you. To my delight he came and declared that he had heard and understood what I said. I asked him to repeat the words. He answered, "You said, Mr. Watson come here I want to see you.""

— Alexander Graham Bell (1876)

"I speak without exaggeration when I say that I have constructed 3,000 different theories in connection with the electric light, each one of them reasonable and apparently likely to be true. Yet only in two cases did my experiments prove the truth of my theory. My chief difficulty was in constructing the carbon filament. ... Every quarter of the globe was ransacked by my agents, and all sorts of the queerest materials used, until finally the shred of bamboo, now utilized by us, was settled upon."

— Thomas Edison (1890)

Return the candle to the candelabrum.

Take a minute to notice the time scale of these discoveries. Each one significantly changed society, and each change was at least mostly for the better.

> "If we continually sample from the urn of possible technological discoveries before implementing effective means of global coordination, surveillance, and/or restriction of potentially hazardous information, then we risk eventually drawing a black ball: an easy-to-make intervention that causes extremely widespread harm and against which effective defense is infeasible"
>
> — Nick Bostrom (2013)

As we enter the thirties and forties, many of the rules on which human society was built have given way to science and industry. Prior to this point, technological progress moved at the speed of civilization, and its effects were mainly effects on societies. Each technology has a name attached, but those names do not matter much.

> "Those material inventions, beginning with the use of stones as weapons, which led to the domestication of animals, the production of fire by artificial means, down to the marvellous inventions of our own days, show clearly that an individual was the originator in each case. The nearer we come to our own time and the more important and revolutionary the inventions become, the more clearly do we recognize the truth of that statement. All the material inventions which we see around us have been produced by the creative powers and capabilities of individuals."

Each of the inventors mentioned so far has been basically a good person, interested in finding truth, improving society or, at worst, making a business for themself. Newton mastered calculus; Watt mastered steam; Edison mastered electricity. History was changed by their inventions, but not by their characters.

But in 1939, someone figured out *power* - what we would now call political science. He mastered *propaganda*. And this time, it matters a great deal who he was. He was the writer of the last quote. And he is now widely considered the most evil man ever to have lived.

"I should like to call attention to the fact that the principle of parliamentarian democracy, whereby decisions are enacted through the majority vote, has not always ruled the world. On the contrary, we find it prevalent only during short periods of history, and those have always been periods of decline in nations and States."

— Adolf Hitler, Mein Kampf (1926)

Starting in 1939 and continuing until 1945, World War 2 killed about 60 million people. Had it gone differently, it's likely that the entire world would have fallen under a single totalitarian regime. This would not have meant human extinction, exactly - but it would most likely mean the loss of humanity's potential.

And so the world's greatest minds believed they had no choice. They had to gather in secret, and create the atomic bomb - a weapon to destroy cities, or the whole world.

"Despite the vision and farseeing wisdom of our wartime heads of state, the physicists have felt the peculiarly intimate responsibility for suggesting, for supporting, and in the end, in large measure, for achieving the realization of atomic weapons. Nor can we forget that these weapons as they were in fact used dramatized so mercilessly the inhumanity and evil of modern war. In some sort of crude sense which no vulgarity, no humor, no overstatement can quite extinguish, the physicists have known sin; and this is a knowledge which they cannot lose."

— Robert J. Oppenheimer (1947)

Using the fourth candle, which represents science, light the fifth candle to represent industrialization.

After the war, things settled down - but the pace of progress continued.

In 1951, the first transistor.
In 1952, the first hydrogen bomb.
In 1953, the discovery of DNA's structure.
In 1954, the first solar cell, model rocket, and nuclear submarine.
In 1955, the Polio vaccine.
In 1956, the first commercial nuclear power station.
In 1957, Sputnik, the first orbital space flight.
In 1958, the first integrated circuit.
In 1959, Lunik 2, the first satellite to reach the moon.

As our technology took off - figuratively and literally - no one knew what we would find. Predicting the future was left mostly to science fiction writers, and their predictions were not especially accurate. But some scientists did take important questions seriously. For example, would there be life in space? The SETI project began in 1961, at the National Radio Astronomy Observatory in West Virginia.

"I wrote down all the things you needed to know to predict how hard it's going to be to detect extraterrestrial life. And looking at them it became pretty evident that if you multiplied all these together, you got a number, N, which is the number of detectable civilizations in our galaxy."

— Frank Drake

"Humanity seems to have a bright future, i.e., a non-trivial chance of expanding to fill the universe with lasting life. But the fact that space near us seems dead now tells us that any given piece of dead matter faces an astronomically low chance of begetting such a future. There thus exists a great filter between death and expanding lasting life, and humanity faces the ominous question: how far along this filter are we?"

— Robin Hanson (1998)

In 1962, the cold war between the United States and the Soviet Union reached a crisis. US destroyers under orders to enforce a naval quarantine off Cuba did not know that the submarines the Soviets had sent to protect their ships were carrying nuclear weapons. So the Americans began firing depth charges to force the submarines to the surface, a move the Soviets on board interpreted as the start of World War III.

> "We're going to blast them now! We will die, but we will sink them all. We will not disgrace our navy,"
>
> — Captain Valentin Grigorievitch Savitsky (1962)

Take the fifth candle, which represents industry. Hold it over the stack of papers until three drops of wax fall.

The launch of the submarine's nuclear torpedo required the consent of all three senior officers aboard: Captain Valentin Grigorievitch Savitsky, political officer Ivan Semonovich, and second in command Vasili Arkhipov.

Arkhipov was alone in refusing to launch the nuke, insisting the submarine surface to receive orders from Moscow. Had he chosen differently, the likely result would have been all-out nuclear war.

Meanwhile, technology marched on. And for the first time, it seemed that technological progress might not go on forever, but build towards an ultimate conclusion.

> "Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultra-intelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind. Thus the first ultraintelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control."
>
> — I.J. Good, Speculations Concerning the First Ultraintelligent Machine (1963)

Place an unlit candle in the last spot, to represent artificial intelligence.

Two years later, Gordon Moore famously observed:

> "The complexity for minimum component costs has increased at a rate of roughly a factor of two per year. Certainly over the short term this rate can be expected to continue, if not to increase. Over the longer term, the rate of increase is a bit more uncertain, although there is no reason to believe it will not remain nearly constant for at least 10 years. That means by 1975, the number of components per integrated circuit for minimum cost will be 65,000."
>
> — Gordon Moore (1965)

Using the fifth candle, which represents industrialization, light the sixth candle to represent the invention of computers.

> "Moore's Law of Mad Science: Every 18 months, the IQ required to destroy the world drops by 1 point."
>
> — Source unknown (2005)

Meanwhile, the rockets that had first been developed for war were turned to exploration as well:

> "Here men from the planet Earth first set foot upon the Moon July 1969, A.D. We came in peace for all mankind"
>
> — Apollo 11 Plaque (1969)

Compared to landing on the moon, computers were just not that impressive. In 1973 the British Science Research Council commissioned James Lighthill to evaluate the state of artificial intelligence:

> "Most workers in AI research and in related fields confess to a pronounced feeling of disappointment in what has been achieved in the past twenty-five years. Workers entered the field around 1950, and even around 1960, with high hopes that are very far from having been realized in 1972. In no part of the field have the discoveries made so far produced the major impact that was then promised."
>
> — James Lighthill (1973)

So began the AI winter: a period of pessimism, reduced funding, and minimal progress. We now know that AI is hard; how hard exactly, we don't know. Since then, Moore's Law has continued - and so, every year, the difficulty drops a little bit.

Moving away from the long-term trends and back to concrete events, we now reach the historical event that is today's namesake: the Petrov incident. On September 26, 1983, Stanislav Petrov was the duty officer at the Oko nuclear early warning system.

> "An alarm at the command and control post went off with red lights blinking on the terminal. It was a nasty shock. Everyone jumped from their seats, looking at me. What could I do? There was an operations procedure that I had written myself. We did what we had to do. We checked the operation of all systems - on 30 levels, one after another. Reports kept coming in: All is correct; the probability factor is two. ... The highest."
>
> — Stanislav Petrov

> "I imagined if I'd assume the responsibility for unleashing the third World War - and I said, no, I wouldn't. ... I always thought of it. Whenever I came on duty, I always refreshed it in my memory."
>
> — Stanislav Petrov

Had he followed procedure, and reported up the chain of command that the Americans had launched missiles, this could have set off a nuclear war. So instead of telling his superiors what the system was saying, Petrov told his superiors that it was a false alarm - despite not really knowing this was the case.

At the time, he received no award. The incident embarrassed his superiors and the scientists responsible for the system, so if he had been rewarded, they would have to be punished. (He received the International Peace Price thirty years later, in 2013).

Things eventually calmed down. The Soviet Union dissolved. Safeguards were put on most of the bombs, to prevent the risk of accidental (or deliberate but unauthorized) detonation.

In 1985, Joe Farman, Brian Gardiner, and Jonathan Shanklin made a disturbing discovery. The ozone layer, the part of our atmosphere that filters out most UV radiation, was disappearing due to chlorofluorocarbon pollution. Just two years later a treaty was written to ban the use of CFCs, and two years after that, in 1989, it was in effect. As of today, every country in the United Nations has ratified the Montreal protocol.

> "The hole in the ozone layer is a kind of skywriting. At first it seemed to spell out our continuing complacency before a witch's brew of deadly perils. But perhaps it really tells of a newfound talent to work together to protect the global environment."
>
> — Carl Sagan (1998)

But not every threat to humanity is as easy to understand or address as nuclear weapons or the ozone layer.

> "An unFriendly AI with molecular nanotechnology (or other rapid infrastructure) need not bother with marching robot armies or blackmail or subtle economic coercion. The unFriendly AI has the ability to repattern all matter in the solar system according to its optimization target. This is fatal for us if the AI does not choose specifically according to the criterion of how this transformation affects existing patterns such as biology and people. The AI does not hate you, nor does it love you, but you are made out of atoms which it can use for something else. The AI runs on a different timescale than you do; by the time your neurons finish thinking the words "I should do something" you have already lost"
>
> — Eliezer Yudkowsky, Artificial Intelligence as a Positive and Negative Factor in Global Risk (2006)

"What we do have the power to affect (to what extent depends on how we define "we") is the rate of development of various technologies and potentially the sequence in which feasible technologies are developed and implemented. Our focus should be on what I want to call differential technological development: trying to retard the implementation of dangerous technologies and accelerate implementation of beneficial technologies, especially those that ameliorate the hazards posed by other technologies."

— Nick Bostrom (2002)

Place an unlit candle in the second-to-last spot, to represent safe artificial intelligence.

"Though I would have liked my chances in a rematch in 1998 if I were better prepared, it was clear then that computer superiority over humans in chess had always been just a matter of time."

— Garry Kasparov, world Chess champion, after losing to IBM's Deep Blue

Twenty-four years after the Lighthill Report declared AI a failure, in 1997 the computer program Deep Blue defeated World Chess Champion Garry Kasparov. Chess, it turns out, was not as difficult as we thought. Fully general intelligence, however, remained out of reach.

"I, for one, welcome our new computer overlords."

— Ken Jennings, Jeopardy champion, after losing to IBM's Watson

Jeopardy, it turns out, was not as difficult as we thought. But fully general intelligence remains out of reach.

We think that it is difficult.

Can progress in computing truly threaten us? So far, as science and technology have advanced, human flourishing has advanced in tandem. We have built horrors, to be sure: machine guns and mustard gas and even nuclear bombs. But their aggregate impact on human life pales in comparison to that of aviation and telecommunications and antibiotics and ten thousand other miracles.

Perhaps artificial intelligence will be made safe too, but the example of nuclear weapons shows that this is not certain. But for the actions of Arkhipov and Petrov, we could have wiped out not just ourselves, but our children's children, and the possibility of ever reaching beyond the Earth.

Which brings us to our next crisis, in 2012, and this one is not so clear.

> "Recently, several scientific research teams have achieved some success in modifying influenza A/H5N1 viruses such that they are now transmitted efficiently between mammals, in one instance with maintenance of high pathogenicity. ... The NSABB was unanimous that communication of the results in the two manuscripts it reviewed should be greatly limited in terms of the experimental details and results. The life sciences have reached a cross-roads. The direction we choose and the process by which we arrive at this decision must be undertaken as a community and not relegated to small segments of government, the scientific community or society. Physicists faced a similar situation in the 1940s with nuclear weapons research, and it is inevitable that other scientific disciplines will also do so."
>
> — Natl. Security Advisory Board for Biosecurity, (2012)

After several months, the decision was reversed, and a revised version of the bird flu paper was approved for publication, by a vote of 12 to 6.

And now we're at the present day. So far, humanity has neither destroyed itself, nor reached a safe position. But this is only the middle of the story. We approach the climax of human history, where we will either destroy ourselves, or spread through the stars.

Take the first candle. Read the following, then return it.

> By the power of fire, I am animated, freed from the demands of mere subsistence, and able to care about the future.

Take the second candle. Read the following, then return it.

> By the power of language, I am able to share what I know of the world and of the future, and to hear what others have learned before me.

Take the third candle. Read the following, then return it.

> By the power of writing, my words echo through time and space. I take part in a dialogue of billions, and together we choose the future we want.

Take the fourth candle. Read the following, then return it.

> By the power of science, I know the true nature of the world I live in. I live in a world that obeys physical laws, which govern the outcomes of my actions, and so I can know the consequences of what I do.

Take the fifth candle. Read the following, then return it.

> By the power of industry, the world I live in is transformed. I can transform it further, if I choose.

Take the sixth candle, and read the following.

> By the power of computation, the power of my mind is amplified. I can see whole the knowledge of mankind, a great fractal pattern of summaries and details and summaries-of-summaries and details-of-details and search it all with a word.

Take the candle and hold it near the seventh and eight candles, which represent safe and unsafe artificial intelligence. Then pull it back, and return it to its place.

Read the following together.

We, as individuals, must each choose:
Will we stand aside, as history continues?
Will we be heroes, moving to the center of the action?
Will we find the heroes and be their allies, up close or from afar?
It's easy to imagine ourselves as powerless, and take that as an excuse.

Let us not forget what power we have, and what we truly are.
