You Only Live Twice:

A Computer Simulation of the Past

Could be Used for Technological Resurrection

Alexey Turchin

Digital Immortality Now

Foundation Science for Life Extension

[alexeiturchin@gmail.com](mailto:alexeiturchin@gmail.com)

**Abstract**: In the future, it will be possible to create advance simulations of ancestor in computers. Superintelligent AI could make these simulations very similar to the real past by creating a simulation of all of humanity. Such a simulation would use all available data about the past, including internet archives, DNA samples, advanced nanotech-based archeology, human memories, as well as text, photos and videos. This means that currently living people will be recreated in such a simulation, and in some sense, “resurrected”. Such “resurrectional simulation” could be deliberately created just for this goal: to return to life all people who have ever lived. The main technical problem of such simulation will be uncertainty about the past, which increases exponentially for more remote times. Such problem could be partly addressed by “acausal trade” between different branches of the multiverse, which will create slightly different versions of the simulation using a quantum randomness generator. Such trade will result in resurrection of all possible people (including those who existed in other branches). Ethical problems of such a resurrectional simulation include: a) possible resurrection of some people against their will; b) such simulation may create additional suffering; с) such simulation could be used by hostile AI to return people to life and then torture them. In this work, I explore preliminary ideas about how to address these problems.

[1. Introduction 2](#_Toc6658603)

[2. Reconstruction of humanity’s past based on informational traces 3](#_Toc6658604)

[2.1 Indirect mind uploading (digital immortality) is aimed at simulational resurrection 3](#_Toc6658605)

[2.2. Advantages of a reconstruction of humanity as the brain modeling method 5](#_Toc6658606)

[3. Resurrection of the dead via multiverse-wide acausal cooperation: a solution to the lost information problem 7](#_Toc6658607)

[3.1. Almond’s universal resurrection machine 7](#_Toc6658608)

[3.2. Problems of Almond’s approach 8](#_Toc6658609)

[Problem 1: Measure decline 9](#_Toc6658610)

[Problem 2: Non-human and unwelcome minds 9](#_Toc6658611)

[Problem 3: Damaged minds 9](#_Toc6658612)

[Problem 4: Most variables inside minds are not mutually random, but products of unique life history 10](#_Toc6658613)

[3.3. Patches 10](#_Toc6658614)

[Patch 1. The use of the human mind’s universal model as a starting point 10](#_Toc6658615)

[Patch 2. The use of the digital immortality data to create only minds which comply with our expectations 11](#_Toc6658616)

[Patch 3. The use of multiverse-wide cooperation for cross-resurrection 12](#_Toc6658617)

[Patch 4: Add randomness not in minds, but in life-history simulations 13](#_Toc6658618)

[3.4. Remaining problems 13](#_Toc6658619)

[3.5. Combining resurrectional simulation and a quantum mind generator 14](#_Toc6658620)

[3.6. Digital immortality data should still be collected 14](#_Toc6658621)

[3.7 Resurrectional simulation and measure domination 15](#_Toc6658622)

[4. Ethical problems of resurrection via simulation 15](#_Toc6658623)

[4.1. Mindcrime and sufferings in simulation could be avoided via selective blunting of experience and the use of quantum randomness 15](#_Toc6658624)

[4.2. People with religious expectations could be resurrected according to them 16](#_Toc6658625)

[4.3. People who don’t want to be resurrected could be split, pro- and anti-life 16](#_Toc6658626)

[4.4. Resurrections by hostile AI could be astronomically outnumbered by benevolent AI resurrections 17](#_Toc6658627)

[4.5. Why resurrect past people? Let’s create new, happy people! 17](#_Toc6658628)

[Conclusion 18](#_Toc6658629)

[Literature 18](#_Toc6658630)

# 1. Introduction

The ability to return the dead to life in this world would be ultimate solution to the problem of death. However, discussion of the topic of resurrection has been limited to religion, with only a few exceptions.

One of the first who started to speak about possibility of technological resurrection was Russian cosmist Fedorov (Young, 2012). The ideas of cryonics, brain preservation at low temperatures, appeared in the 20th century (Ettinger & Rostand, 1965). Another idea is idea of indirect digital immortality (Almond, 2003; Rothblatt, 2012; Turchin, 2018a), that is, reconstruction of the personality based on informational traces.

Bostrom and others wrote that we are likely to be in a computer simulation of the past created by future superintelligent AI (Barrow, 2007; Bostrom, 2003), which could create enormous numbers of such simulations based on ever-cheaper computational power. Simulations of the past naturally entail reconstruction of the minds of people who lived in the past, but the use of a simulation as the main instrument of resurrection is underexplored compared to other ways of reviving the dead, like cryonics. However, Tipler wrote that future superintelligent AI Omega could reconstruct all possible people at the end of the universe using the enormous energy resources of the collapsing universe (Tipler, 1997).

The idea of a *resurrectional simulation* was [outlined](https://www.lesswrong.com/posts/C5ryrg5ktngwsZHnX/resurrection-through-simulation-questions-of-feasibility) on LessWrong by Jacob Cannell in 2012 (Cannell, 2015). The main difference between that work and the analysis presented here is that I into ways to escape the “lost information problem” in greater detail.

The possibility of technological resurrection is based on the assumption that the copy-identity problem will be solved in a copy-friendly way. I have previously discussed the criteria of identity in detail elsewhere (Turchin, 2018a). The main takeaway is that if soul doesn’t exist, then copy-friendly identity theory is probably true. In that case, the criterion for identity is *internal functional non-distinguishability*: that is, I am unable to find the ways to learn if I am a copy or the original from my internal experiences, even if I possess a counterfactual capability to compare them.

In that case, *my identity could be preliminarily defined as my indexical uncertainty about my location among all my copies* with the same indexical information—typically information about my location, time and self-image. More research about the nature of consciousness is need to finally solve the identity problem. In the meantime, to support future simulation, we must implement a “conservative approach to identity”: preserve as many data traces about personality as possible.

In Section 2, I explore the idea of resurrectional simulation. In section 3, the quantum random mind generator will be presented and its combination with resurrectional simulation will be explored. Section 4 is devoted to the ethical problems of simulational resurrection.

# 2. Reconstruction of humanity’s past based on informational traces

## 2.1 Indirect mind uploading (digital immortality) is aimed at simulational resurrection

The idea behind indirect mind uploading is that I will collect many different types of data about myself, including preserving DNA, constant video recording, writing a biography and running different psychological tests, with the expectation that future AI will be able to reconstruct me with some level of functional fidelity based on the accumulated data. I previously explored the best strategies to collect the data based on its capability to predict my future behavior and translate valuable things about me (Turchin, 2018a).

However, there is an open question: how will hypothetical future superintelligent AI—which could be a Singleton agent, a group of humans helped by CAIS, or something else—use this information to resurrect me?

One way is something similar to Egan’s “sideloading”, which could be called *model tweaking*. In Egan’s case, the sideloading happened by constant interaction of AI model and the original, but in our case the person is already dead. AI takes a general model of human mind and tweaks it based on the collected data: choose gender, implant memories, test behavior until it matches the original. This approach is relatively simple if a lot of correctly structured data is available. However, the result will be a rough model of the person, in some sense similar to a very advanced chat-bot.

Another way is direct *extrapolation*. AI will examine all available data about the person, for example his/her Facebook posts in chronological order and try to extrapolate the next post. This method is similar to current machine learning, and could, to some extent, be implemented now. For example, systems like Open AI’s text generator GPT-2 (2019) can write texts in an author’s style, while other systems can mimic voices or photos and even videos. The main difference here is that here there is no explicit model of the human brain, and the system outputs data to data without any tweakable internal variable. Such an approach could work even to predict the internal stream of consciousness—if, and this is the catch—a lot of training data have been recorded and the mind to be modeled is relatively simple. This approach is data intensive, but relatively computationally simple.

The third way is to run a full simulation of the person’s life. This approach is good if we know inputs (DNA+ social conditions in person’s early years) and outputs: his/her texts and other data. This approach seems reasonable if a lot of data are unavailable or missing.

The main idea of a resurrection simulation is that if one takes the DNA of a past person and subjects it to the same developmental conditions, correcting development based on some known outcomes, it is possible to create a model of a past person very close to the original. DNA samples of many people who lived in the past centuries could be extracted via global archeology, from marked graves but also from different form of waste, furniture, paper or even random pieces of soil.

Another possible approach is a brute force reverse engineering approach. The human mind is treated as a mathematical function for which the AI knows output and inputs and has to guess the function. This is similar to the field of functional equations, which is theoretically difficult, and such tasks are known to be computationally tough (most cryptography is based on this fact).

The last variant I can currently imagine, is the use of “deductive Bayesian reasoning”, where all data about the person is treated as evidence in evaluating some hypothesis about who this person was. This task is not only computationally difficult (but may be less so than brute-force reverse engineering) but also requires a very high level of intelligence capable of making very strong conclusions from very small pieces of data (See Yudkowsky’s [post](https://www.lesswrong.com/posts/5wMcKNAwB6X4mp9og/that-alien-message) of hacking of simulation).

There may well be other ways to reconstruct a person based on available information; perhaps using AIXI model (universal problem solver suggested by (Hutter, 2000)) or quantum computers, as well as combinations of the methods discussed above.

## 2.2. Advantages of a reconstruction of humanity as the brain modeling method

If we analyze not one but many persons’ reconstructions based on their informational traces, then the simulation of the history of humanity has the advantage over other listed above methods. The main reason is that the computational complexity of the reconstruction (per brain) diminishes, while in the other methods outlined above, the complexity only grows when going from modeling one person to modeling all of humanity.

If we simulate one person’s life, we necessarily simulate parts of the lives of the many other people with which that person interacts. Thus, we (or AI) simulate hundreds or thousands of people, and the simulation of these other people is mostly wasted. If we simulate all of humanity, we get much more evidence from any piece of data we have: because one piece of information could provide data about minds of many people, and also be a base to the perfect reconstruction of other world. For example, if we learn that there was a tower in a town, we could conclude that each citizen of the town was well aware about it, and also the tower was built by someone and later probably demolished.

Simulation of the whole humanity means resurrection of all people who ever lived. In other words, simulation of all of human history is computationally effective and also “morally effective”, as it provides all humans with a chance of resurrection.

This type of simulation mostly relies on raw computational power rather than the modelling AI possessing very high intelligence. Such computationally powerful but only mildly superintelligent AI is more technically feasible and safer than AI with very high “IQ”, which is much less predictable.

The setup of the resurrectional simulation may be presented as a pseudo-algorithm:

1. Future advanced AI collects all available data about the past. The data is from some arbitrary point in time in the past, like the appearance of written language (~3000 BC), or perhaps much later, the point at which we have much stored data about many people, like the beginning of 20th century up until the moment when the last person to live on Earth died. Such data collection may include brain scans, nanotech-based archeology, all written sources, all video, etc.

1. The AI starts the simulation of human history using known initial conditions. This simulation may not run linearly in the AI’s time: it may have some sort of scenario, as it knows which end points it must achieve (e.g. Homer must write the Iliad). There will still be a lot of people for whom almost no traces remain, and which need to be presented as non-player characters (NPCs) or be generated completely randomly (more about randomness below). The AI may also run the simulation in its own time backwards: from known outcomes to the past.
2. After any person dies in the simulation, the state of his mind is recorded.
3. The simulation should skip moments of intense, unbearable suffering to avoid “mindcrime”, and only output the results of such suffering. Note that the threshold of “unbearable sufferings” could be placed rather high, on the level which most normal people never feel in their life, like dying in flames. Thus, it will not affect most people’s identity.
4. After the simulation is finished, the AI will have approximate models of the minds of most people who ever lived as they were at the moment of their death.
5. Now the AI puts these mind-models into another simulation, an “afterlife”, which may look like the religious expectations of the dead people. The AI will slowly educate them about the real state of affairs and teach them how to live in this “real life”. This afterlife also includes gradual upgrades to the minds of newly revived people, so they have higher cognitive abilities enabling them to adapt to the new world.

The computational power required to do all this may be rather large, depending on the brains’ complexity. The simulation may be even “physical”, like “The Truman Show” movie, in which case minds will be biological and—by definition—conscious. Small nanobots will regulate them and “gods” will control their destinies. But even physical simulation will not require more than one solar system to simulate, so a Kardashev level 2 civilization can afford to resurrect its dead.

# 3. Resurrection of the dead via multiverse-wide acausal cooperation: a solution to the lost information problem

The main theoretical problem of simulational resurrection is that we don’t know much about the past. We will never know, for example, if Russian tsar Alexander I participated in the assassination plan to kill his father, Pavel I, or if “that girl” had an actual crush on you. In other words, our reality could have alternative pasts.

One way to escape this problem is to concentrate on immortality-for-others: if our expectations about other people are satisfied, that is sufficient. This is a very weak form of immortality, as our expectation and memory are weak and easily change.

Another possible approach, suggested by Almond, is resurrection of the dead via a quantum random generator which creates a random mind. But such an approach has several problems: non-human beings in our world, unnecessary suffering of non-perfect copies, and measure decline.

In the following, I suggest four patches to prevent most of the undesired effects.

## 3.1. Almond’s universal resurrection machine

In “[Many-Worlds Assisted Mind Uploading: A Thought Experiment](https://web.archive.org/web/20110513092111/http://www.paul-almond.com/ManyWorldsAssistedMindUploading.htm)”, Almond suggested the following idea about the resurrection of the dead by the use of a quantum random generator, which would create a random mind within a computer (Almond, 2006):

[A technician who lost someone’s brain scan file] writes a computer program which takes input from a physical system. The physical system, known as a *quantum event generator*, generates "1"s and "0"s randomly as a result of quantum events. The program will use the physical system to tell it what sequence of "1"s and "0"s will be used to try to recreate the lost scan file. The program starts with an empty scan file which will be filled with "1"s and "0"s.

If the many-worlds interpretation of quantum mechanics is correct, all possible minds will appear in separate timelines starting from the moment of random mind creation, which would mean the resurrection of everyone from the observer’s point of view. However, a) this approach will not help an outside observer, who wants to resurrect a relative, for instance, as the observer would see only a random mind; and b) the quantum "measure" of existence of each mind will be infinitely small.

It should be noted that the number of possible humans is finite, as it is limited by the number of atoms combinations in the human body, and as some small differences could be ignored, it makes number of significantly different humans much smaller. It makes the task of modeling all possible humans computationally tractable, but it needs to be distributed between different branches of the multiverse, as each separate branch will not have enough computational resources.

## 

## 3.2. Problems of Almond’s approach

To illustrate the problems with quantum mind uploading, I will explore a simplified thought experiment where only names will be restored using quantum mind uploading. First, here is what Almond suggested:

***Thought experiment “Not-patched quantum mind uploading”.***

*Bob had a friend John Smith.*

*John has died and Bob wants to resurrect him. Bob remembers only first letter of John’s last name: S.*

*Bob and John are interested only in the uniqueness of name preservation, and no other identity considerations are important. Bob wants to see his friend alive, and for his friend to be named “John S….” (I would call this immortality from the point of view of the external observer). However, John wants his own immortality, and will only be satisfied if “John Smith” is created.*

*Bob creates random quantum mind A using a quantum generator to choose each new letter in the names.*

*It turns out that A is “jYY2№@11”. Only less than 10-30 share of all such copies in the multiverse are named John Smith. Both Bob and John are unhappy.*

This thought experiment leaves both John and Bob unsatisfied, as described below:

### Problem 1: Measure decline

Problem 1 is a problem for John. Measure can be defined as a share of an observer of a given type between all possible observers. If the typical size of the simulated mind is, say, 1015 bytes, the chance that a randomly generated mind exactly match a particular person is 2-10power15. In other words, the quantum mind generator results in a measure decline of 2-10power15, an extremely large number. Even in thought experiment 1 (subsection 3.2.), the measure declines 1030 times.

Many authors claim that large measure decline should be treated as death or as an infinitely small chance of survival. Such discussions have appeared in the context of so-called quantum immortality, that is, the counterfactual possibility to survive death via existence in quantum multiverse timelines where a person will not die.

Even if measure decline is not bad per se, it leads to a world where very small probability outcomes will dominate the possible futures of an observer. Such parasitic outcomes may be full of suffering; for example, the quantum immortality improbable survival landscape may be dominated by people who are very old and dying but can’t die (such a timeline could be patched by signing up for cryonics).

If we use some expected utility calculations, and measure decline results in declining utility of any useful outcome associated with it, we can just ignore copies with infinitely small measures.

### 

### Problem 2: Non-human and unwelcome minds

Problem 2 is mostly for Bob. Another problem is that most random minds will be non-human, and will not be adapted to our world, so they will suffer or cause suffering to people living here. In our thought experiment, “jYY2№@11” is an example of a non-human random mind. Such random minds are bad for any outside observer, like Bob, as their presence means he will be very unlikely to meet anyone resembling his friend John Smith.

### Problem 3: Damaged minds

Problem 3 is a problem for both Bob and for John. Most randomly created minds will be not minds at all, but garbage code, or in the “best case,” damaged minds. For example, if Bob wants to resurrect John Smith, there will be many copies where his name (as well as his other properties) are permutations of the name Smith, for example Smthi, Smiht, Misth, Smitt, etc. For a name that is *n* bits long, there are *n* individual names with a 1-bit difference.

Thus, for any real person, there will be much larger set of his/her damaged copies, which implies suffering as the most probable outcome of quantum random resurrection for such a person and s-risks for all people.

### Problem 4: Most variables inside minds are not mutually random, but products of unique life history

MakoYass [suggested](https://www.lesswrong.com/posts/BJfb2hqtdnaRijAfz/resurrection-of-the-dead-via-multiverse-wide-acausual#5o7Dby3LCw24mvnY4) the following objection to the random mind quantum generator:

The aspects over which real people vary is a function of real lived histories, not random numbers in a neural template that you can just randomize independently. In order to dredge up a uniformly sampled mind who really could have lived, so that total measure of humans who really lived would be preserved over conditions of multiversal cooperation-, you would need to simulate randomly chosen actual human histories. That would be expensive. Most lived histories were pretty miserable, too, so there's an additional eudaimonia cost being imposed there.

If your generator abandons the brittle logic of real histories, you're going to end up with a majority (there are many more ways to miss reality than hit it) of specimens who, on inspection, could not have lived. Who have inconsistent memories, personalities that don't accord with their experiences, skills they never practiced.

In other words, we either get a lot of “bad brains” or have to simulate all possible human histories to know the correct distribution of bad and good brains, which makes the problem computationally intractable.

## 3.3. Patches

Fortunately, quantum random mind uploading could be patched to provide much more satisfaction for John and Bob.

### Patch 1. The use of the human mind’s universal model as a starting point

The goal of this patch is to avoid creating the minds of “aliens” or non-workable gibberish code, and thus prevent suffering of most created minds. For example, for a human mind model, his/her possible name will be generated not as random symbols, but from a set of typical human names.

Such a human mind model may look like an untrained neural network with the general architecture of a human mind and some other constraints, set up so any random set of parameters will create a more-or-less normal human mind. We assume that some future assistant AI will be able to find an appropriate model.

In that case, Bob uses a random mind generator for parameters of the universal human mind model. He gets “Maria Stuart”. This will increase the share of the worlds where the real John Smith is resurrected to 10-10. Both John and Bob are a little bit more satisfied, as Bob gets a human friend, and John increases his measure.

Obviously, some minds may not want to be resurrected, but this could an important parameter in the model, and models where “resurrection preference = false” will not be instantiated.

### Patch 2. The use of the digital immortality data to create only minds which comply with our expectations

The problem of Bob’s satisfaction could be overcome by the use of Bob’s expectations as priors, if there are no other current or future sources of data about John.

In that case, Bob could use his memories about John S. to create a model of John S. He remembers that John was either John Smith or John Simpson. He uses a random quantum coin to choose between Smith or Simpson, and gets “John Simpson”.

In another branch of the quantum multiverse, where the coin fails tails, John Smith appears, but his measure declines to 0.5. Both John and Bob are partly satisfied. Bob got someone who looks like his friend, but Bob knows that it is not exactly his friend, and that his friend has now has a smaller measure of existence.

Digital immortality, or indirect mind uploading, is the collection of information about a person while he is alive with the hope that future advanced AI may be able to resurrect the person. This type of resurrection will involve creating an advanced model of the personality based on all available information. Such a model will, by definition, satisfy Bob and all other relatives, as all available information has already been taken into account, including all relatives’ expectations. However, large chunks of information will never be known, and thus have to be replaced with some random data. Even if quantum randomness is used to fill the gaps, John will have an infinitely small share of all possible worlds, and in most other worlds, he will be replaced by someone else.

### Patch 3. The use of multiverse-wide cooperation for cross-resurrection

The next step is that Bob considers that not only his universe exists, but all other possible universes exist in the multiverse.

Bob concludes that because all possible observers exist in the Multiverse, his John Simpson created via a quantum random generator is a resurrection of some John Simpson from another universe. The John Smith who lived in our universe will be resurrected in some other universe, where another copy of Bob will do the same experiment. In other words, Bob and Bob’s copies in other universes will cooperate to resurrect the exact John Smith Bob remembers.

As the second universe is exactly the same as ours except for John’s name, there is another exact copy of Bob in it, and this Воb’s copy also wants to resurrect his friend John S., so he uses another quantum random mind generator. Now the following happens, see table 1:

*Table 1. Cross-universes acausal cooperation in the resurrection with use of quantum mind generator.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Universes:** | **Воb’s dead friend’s real name:** | **Result of quantum mind generator:** | **Total measure at the end in both universes:** |
| Воb’s universe, Bob wants to resurrect his friend John S. [Smith or Simpson]. | John Smith,  measure = 1 | If coin = heads, Smith created;  measure = 0.5 | John Smith = 1  John Simpson = 1 |
| If coin= tails, Simpson created;  measure = 0.5 |
| Воb’s copy universe, Bob wants to resurrect his friend John S. [Smith or Simpson] | John Simpson,  measure = 1 | If coin = heads, Smith created;  measure = 0.5 |
| If coin= tails, Simpson created;  measure = 0.5 |

So, the total measure of John Smith has not declined, if Bob takes into account that other copies of Bob in other universes will run the same experiment. By deciding to start the random mind generator (and to not turn off the resulting mind), Bob joins a large group of other minds who think similarly, but who are located in causally disconnected parts of the Multiverse. Everyone expects that some other random generator will create an exact copy of their loved one.

In a case of substantial amounts of missing data, like gigabytes, this requires a simultaneous run of an extremely large number of quantum random mind generators, like 1010power9, which is only possible via multiverse-wide cooperation. The measure will not decline in such a case, as for every dead person there will be one random person, and given the large numbers, any person will be randomly recreated, at least in approximately one world. Some may go deeper and take into account standard deviation, but because we use quantum generators in the many worlds interpretation, each universe creates exactly its share of John, and there will be no fluctuations, which would result in non-existence of some Johns and two copies of another.

Any of Воb’s copies can join such a multiverse-wide cooperation by creating just one quantum random mind (and treating the resulting mind well).

### Patch 4: Add randomness not in minds, but in life-history simulations

Problem 4 stated that we have to actually run all possible human histories in order to learn which variables in the human mind are truly random, and only after that could we run a quantum mind generator.

The patch is that we add quantum randomness not in the mind, but in the simulation of the person’s history. For example, if we don’t know the surname of John S., we create a simulation of his life in which his surname is chosen by a quantum randomness generator. Whatever surname appears will be naturally incorporated in his life history. In that case, we don’t have to run simulations of all possible human lives, but instead run just one life simulation, which is computationally tractable for future AI. In the other branches of the multiverse, the alternate versions of John S. will be created.

## 3.4. Remaining problems

*Multiverse*. What if the multiverse doesn’t actually exist? In that case, Bob and John will get partly satisfying results, as Bob will get a copy of John, though this copy is not perfect from John’s point of view. If the quantum multiverse is not real, but some other form of the multiverse exists, like one based on inflational cosmology, the resurrection method will still work.

*Defection*. Bob may not create any random mind generators at all but still expect that someone else will recreate his friend. In general, the rate of defections may be known and compensated for by those who have more resources increasing the number of random minds.

There are several other possible generic problems of multiverse-wide cooperation, including infinite ethics, the possibility of acausal blackmail, a method to measure similarity between agents, and problems with agents that have other values, as described in <http://effective-altruism.com/ea/1rz/request_for_input_on_multiversewide/fbo>.

## 3.5. Combining resurrectional simulation and a quantum mind generator

Above, I explored two approaches to the returning the dead to life: whole-history simulation and a quantum mind generator. Their combination is rather straightforward: fill the gaps in the simulation with outputs of random quantum generator. Actually, the proposed patches are inclined in this direction. It is also computationally cheap, as the whole invention only involves adding a quantum randomness generator to the project of simulating the past.

In the simulating the entire history of humanity, future AI will fill in gaps with random quantum noise, starting from the beginning of the simulation. Some of the simulated people will be almost random in their personal traits but they still will be the members of some tribe, will speak some language, have some cultural background, and observe some predictable patterns among wildlife (for example, when hunting mammoth). So, these will not be completely random minds, but they will be rather typical people for their historical moment.

## 3.6. Digital immortality data should still be collected

If a quantum randomness generator could recreate any possible mind, one may suggest that any other efforts toward immortality are useless and redundant. Thus, maybe there is no need to fight aging, sign up for cryonics, or collect data for digital immortality. However, we can’t be sure that a quantum mind generator will work.

The main uncertainty is our understanding of quantum mechanics—what if Everett’s interpretation (or our interpretation of his interpretation) is false? Problems could also arise with the appearance of future superintelligent AI and its willingness to implement a quantum mind generator, or with the acausal trade between universes required to increase “measure”.

That is why quantum randomness is only an addition to simulational resurrection. Then, even if it fails, we will still get some type of resurrection, albeit not one that is very precise for most people. To make it more precise for me—and increase the chance of success—I should collect digital immortality data. I should not only collect data about me, but help to preserve data about all of humanity in museums, archives and elsewhere.

## 3.7 Resurrectional simulation and measure domination

If in the future there will be many different simulations of the past, some of them may be not resurrectional, but created by other AI within other branches of the multiverse with different goals. Some such goal may be even to cause extreme pain to humans for very long time, that is so called s-risks.

The probability for someone to find oneself in a simulation is predicted by poorly defined thing called “measure”. I suggested that the measure is proportional to the energy used for computations of the mind’s simulation. More in the article “Simulation termination risks”, The AIs which have biggest share of resources could use them to dominate in measure and get most of observers’ probability weight inside their simulations.

Measure manipulation may be combined with a resurrectional simulation, as if a future AI wants to resurrect the dead, it may also want that most of the revived people will be in such resurrectional simulation, but not in some other experimental simulation where they could suffer. More about the idea in (Turchin, 2018b).

# 4. Ethical problems of resurrection via simulation

## 4.1. Mindcrime and sufferings in simulation could be avoided via selective blunting of experience and the use of quantum randomness

Many people argue against simulating the past at all, as it means creating a lot of artificial suffering, which they call “mindcrime”. The obvious way around this problem is to turn off subjective experiences in the simulation if the level of pain is above some threshold level, and instead approximate the effects of this pain on the mind of the sufferer.

However, if AI can predict future mind states without running a simulation, then the simulation of all of humanity becomes unnecessary. There is a tradeoff between simulation’s fidelity (equal to the fidelity of the resurrected mind) and the level of suffering during the simulation. If you want to have exact an immortal copy of you, you have to go through all the suffering of your life again. Not everybody will share this *amor fati* view, even when expecting a tremendous payoff of immortality and paradise-style living.

Quantum randomness-based mind generation may shift the tradeoff toward less suffering and lower fidelity, as non-fidelity could be compensated for by generating some quantum random outcomes of suffering without the need to go through them. For example, after a night of toothache, someone may have to decide whether he will pay for an implant, an outcome not available from historical records. If the whole night of suffering is simulated, the outcome could be easily predicted. However, we may skip simulating this night and just randomly generate the choice.

## 4.2. People with religious expectations could be resurrected according to them

Living their life again and then appearing in a new and completely different world is not what most people expect from immortality. While most people don’t want to die, their preferences about it are expressed in their religious expectations.

This could be compensated for by another “afterlife” simulation, where minds are transferred after their “death” in the resurrectional simulation. Such an afterlife will initially look like the model of afterlife according to the person’s religious expectations. The person will be gradually educated about what real life should look like; in some sense, it will be like being in school again.

## 4.3. People who don’t want to be resurrected could be split, pro- and anti-life

Some people, who committed suicide or just became bored with life may not want to be alive at all. In most cases, their problems could be solved: unshared love, humiliation or intense pain.

People consists (or could be describe as consisting of) of many internal parts, which often have different goal and subgoals. Not all such parts would want to die; in extreme cases, the person could be split into two personalities. One would include” the sub-personalities which want to be alive, and other the ones that want to disappear completely. Then, the second suicidal personality would be turned off.

Some people may not want to be simulated at all, and could express such desire, especially people living in the 21 century and close to the transhumanists’ circles. Such people could be replaced by NPC in the moments when they appear in life of other people. I hope that future advance AI will be also capable to solve complex ethical and legal problems. For example, getting permission to build something in historic city is more difficult then actually build it; the same problem also appears in medical ethics.

Also, if we assume existence of the multiverse, there will be different branches with different ethics, and some of them could be more pro-resurrection and pro-simulation anyway. A person who doesn’t want to be resurrected, only increase relative chances that she will be resurrected by unfriendly AI which doesn’t care about her choices. In other words, if someone doesn’t want to be resurrected because she is afraid of evil AI, she paradoxically only increases her chances to find herself in the hands of the evil AI! Thus, such volition probably should not be taken literary but extrapolated based on this additional information about expected probability distribution.

## 4.4. Resurrections by hostile AI could be astronomically outnumbered by benevolent AI resurrections

Some people are afraid of AI-driven resurrection, as such AI could turn out to be evil and torture the resurrected person for eternity (s-risks). If benevolent AIs outnumber hostile AIs, they could “save” minds from hostile AI by creating many copies of such minds in each observer-moment and thus imposing indexical uncertainty over each mind. This would almost instantly remove the mind from “hell.” This idea is controversial and requires patching to avoid causing more suffering than it creates. I have previously explored this idea: [Curing past sufferings and preventing s-risks via indexical uncertainty](https://forum.effectivealtruism.org/posts/3jgpAjRoP6FbeESxg/curing-past-sufferings-and-preventing-s-risks-via-indexical).

## 4.5. Why resurrect past people? Let’s create new, happy people!

The following objection has been presented often, but I will repeat it in my own words. Resurrectional simulation will only increase the total suffering in the world, will spend a lot of resources, and will produce maladapted beings, most of whose traits have to be changed in a costly rehabilitation process. From the utilitarian perspective, it would be much more effective to instead create new, healthy people.

Firstly, there are reasons we do not apply this logic to our world. Society cares about invalids and terminally ill patients despite the fact that it would be much cheaper to create embryos and pay surrogate mothers to bring new, healthy children to life.

One reason is that it is part of human nature to have strong personal bonds between already living people: parents, spouses, children, friends. Resurrection of the dead is part of our obligation to our fathers, as Fedorov said in “Common task” book. However, if I resurrect just one person, she will be lonely, and would also have the same obligation to her parents and other dead relatives. By induction, this means that all people who have lived (and maybe even animals) should be resurrected.

# Conclusion

For the resurrection of the dead, a Kardashev 2 AI could use a “*resurrection simulation*”: a simulation of the whole history of humankind based on all available traces. Such a simulation would use historical data as reference points to calibrate the simulation. It could deliver good results for people who lived recently, including those who died in the 21st century, but as uncertainty about the past grows exponentially due to data decay and the absence of “big data” in the past, there will be a point in time before which such a resurrection simulation will be insufficient. This data decline problem could be partly solved by adding quantum randomness to the whole setup, because in some branches of the universe, the result will be exactly as needed. Measure decline could be compensated for by acausal trade with other similar civilization in other parts of the multiverse.

# Literature

Almond, P. (2003). *Indirect Mind Uploading: Using AI to Avoid Staying Dead*. Retrieved from https://web.archive.org/web/20120310211305/http://www.paul-almond.com/IndirectMindUploading.htm

Barrow, J. D. (2007). *Living in a simulated universe*. na.

Bostrom, N. (2003). Are You Living In a Computer Simulation? *Published in Philosophical Quarterly (2003) Vol. 53, No. 211, Pp. 243-255.*

Cannell, J. (2015). The Brain as a Universal Learning Machine. Retrieved June 17, 2018, from LessWrong website: https://www.lesswrong.com/posts/9Yc7Pp7szcjPgPsjf/the-brain-as-a-universal-learning-machine

Ettinger, R. C., & Rostand, J. (1965). *The prospect of immortality*. Sidgwick and Jackson.

Hutter, M. (2000). A theory of universal artificial intelligence based on algorithmic complexity. *ArXiv Preprint Cs/0004001*.

Rothblatt, M. (2012). The Terasem mind uploading experiment. *International Journal of Machine Consciousness*, *4*(01), 141–158.

Tipler, F. J. (1997). *The physics of immortality: Modern cosmology, God, and the resurrection*.

Turchin, A. (2018a). *Digital Immortality: Theory and Protocol for Indirect Mind Uploading*. Retrieved from https://philpapers.org/rec/TURDIT

Turchin, A. (2018b). Preventing s-risks via indexical uncertainty, acausal trade and domination in the multiverse - LessWrong 2.0. Retrieved April 16, 2019, from LessWrong website: https://www.lesswrong.com/posts/eWx8LzasNutvmkPMf/preventing-s-risks-via-indexical-uncertainty-acausal-trade

Young, G. M. (2012). *The Russian cosmists: The esoteric futurism of Nikolai Fedorov and his followers*. Oxford University Press.

Survey of many ideas of resurrection, mostly religious

Some Resurrection Theories Eric Steinhart

<http://ericsteinhart.com/articles/reztheories.pdf>