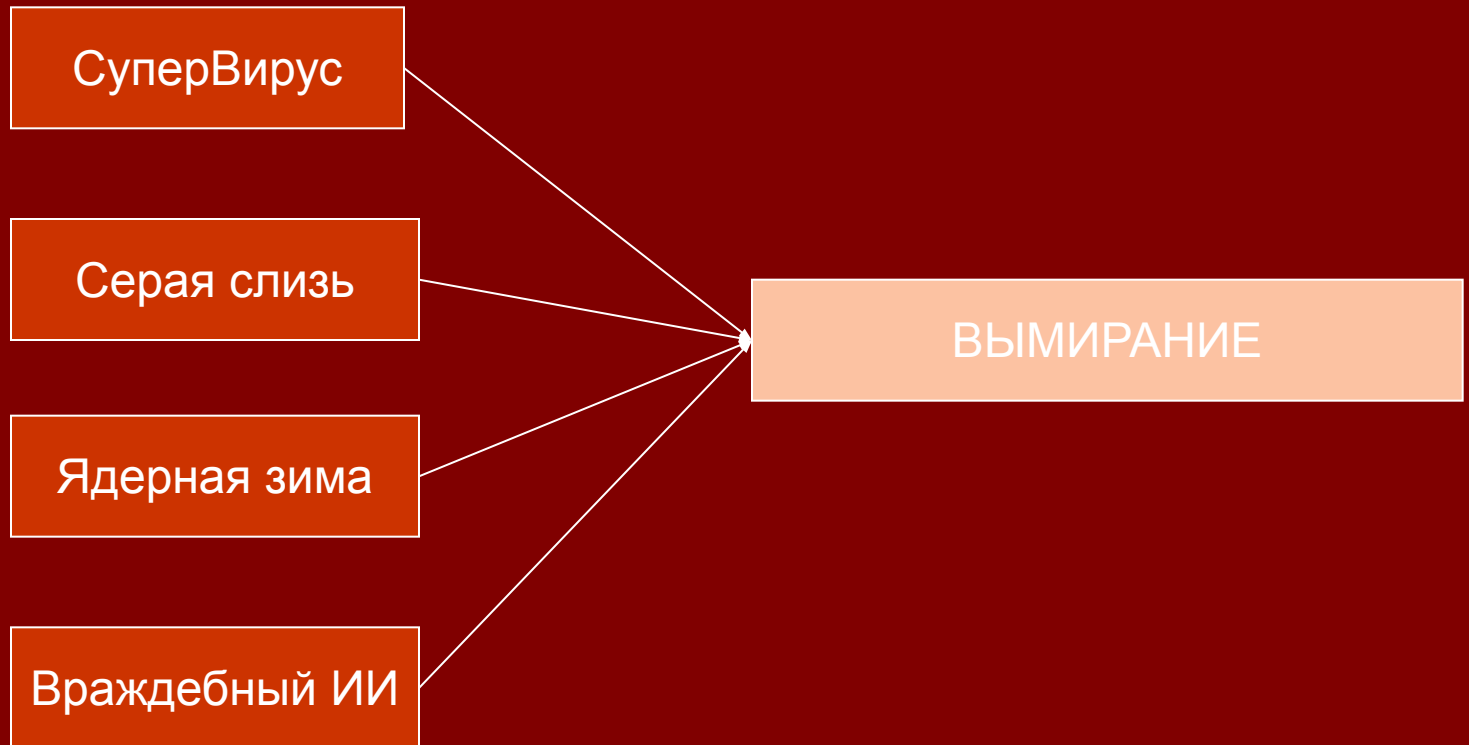


# Ошибки в оценке рисков

## ПОЛНОСТЬЮ ЗАТЕМНЯЮТ РЕЗУЛЬТАТ

- +1% -1% -1% -1%+1% +1% +1% +1%
- -1% -1% +1% +1% +1% +1% +1%
- -1% -1% -1% -1% -1%+1% +1%
- +1% -1% -1% -1%+1% +1% +1% +1%
- +1% -1%-1% -1% -1% -1%
- +1% +1% +1% +1%
- -1% -1% -1% -1%
- -1% +1% -1% -1%+1% +1% +1% +1%
- -1% +1% -1% -1%

# Линейное представление об угрозах существованию



# Системный риск:



# Структурный кризис как угроза существованию

- Превышение критического числа хищников в экосистеме. (остров Пасхи)
- Разрастающаяся трещина (конфликт сверхдержав)
- Самовоспроизводящаяся паранойя (резня в Кампучии и 37 г)
- Самовоспроизводящаяся дезорганизация (парад суверенитетов)
- Самоподдерживающаяся моральная деградация (крах Римской империи)
- Скручивающаяся спираль экономики (Великая депрессия)

# Структурный кризис как угроза существованию

- Эффект домино (наркоман, который подсаживает других, чтобы достать деньги на новую дозу)
- «Естественный» отбор краткосрочных выгод вместо долгосрочных (Маркс: злые капиталисты вытесняют добрых)

# Самоорганизация рисков

На примере катастрофы авианосца США

1. Завезли устарелые бомбы.
2. Модифицировали систему снятия с предохранителя ракет на самолётах.
3. В самолёте произошёл бросок напряжения. (дальше этот самолёт не важен)
4. Стартовала ракета и попала в другой самолёт (дальше ракеты не важны)
5. Все пожарные побежали тушить горящие бомбы
6. Бомбы взорвались и убили всех пожарных.
7. Матросы, не понимающие в пожарном деле, стали тушить водой горячее, и оно затекло в трюмы

# Самоорганизация рисков

- Катастрофический процесс переходит от одной системы к другой.
- При этом масштаб его усиливается и, часто, скорость протекания возрастает.
- Катастрофический процесс распространяется по линиям непредсказуемости – то есть он быстрее всего развивается в областях человеческого незнания.

# Борьба спасателей мира друг с другом как системный риск.

- Даже бен Ладен хочет по своему спасти мир.
- Большинство войн ведётся с тем, чтобы никогда больше не было войны.
- «Войны не будет, но будет такая борьба за мир, что от мира камня на камне не останется».
- Рак, лебедь и щука



# Непредсказуемость

- Удобнее прогнозировать понятные и контролируемые риски.
- Наибольшую опасность представляют неожиданные, внезапные и непонятные риски.
- Удобнее бить вперёд, но удар сзади опаснее.
- Большинство катастроф последнего времени оказались неожиданными – к ним не готовились:
  - Цунами в ЮВА
  - Сель в Долине гейзеров
  - Сход ледника Колка
  - Атака 11 сентября.

# Принципиальная непредсказуемость угроз существованию

- Таких событий ещё ни разу не случилось, иначе бы мы здесь не сидели.
- Предсказания проверяются экспериментом, но именно «эксперимента» мы и хотим избежать
- Статистика таких событий невозможна.
- Как пользоваться вероятностными оценками таких рисков, неизвестно, так как это однократные события.

# Антропный принцип и катастрофы

- «За всю историю Вселенной ещё ни разу не случилось такой катастрофы, которая бы сделала невозможной жизнь на Земле».
- Ничего не говорится о будущем.
- Будущее не похоже на прошлое.
- Нельзя пользоваться выводами из прошлого для предсказания будущего.

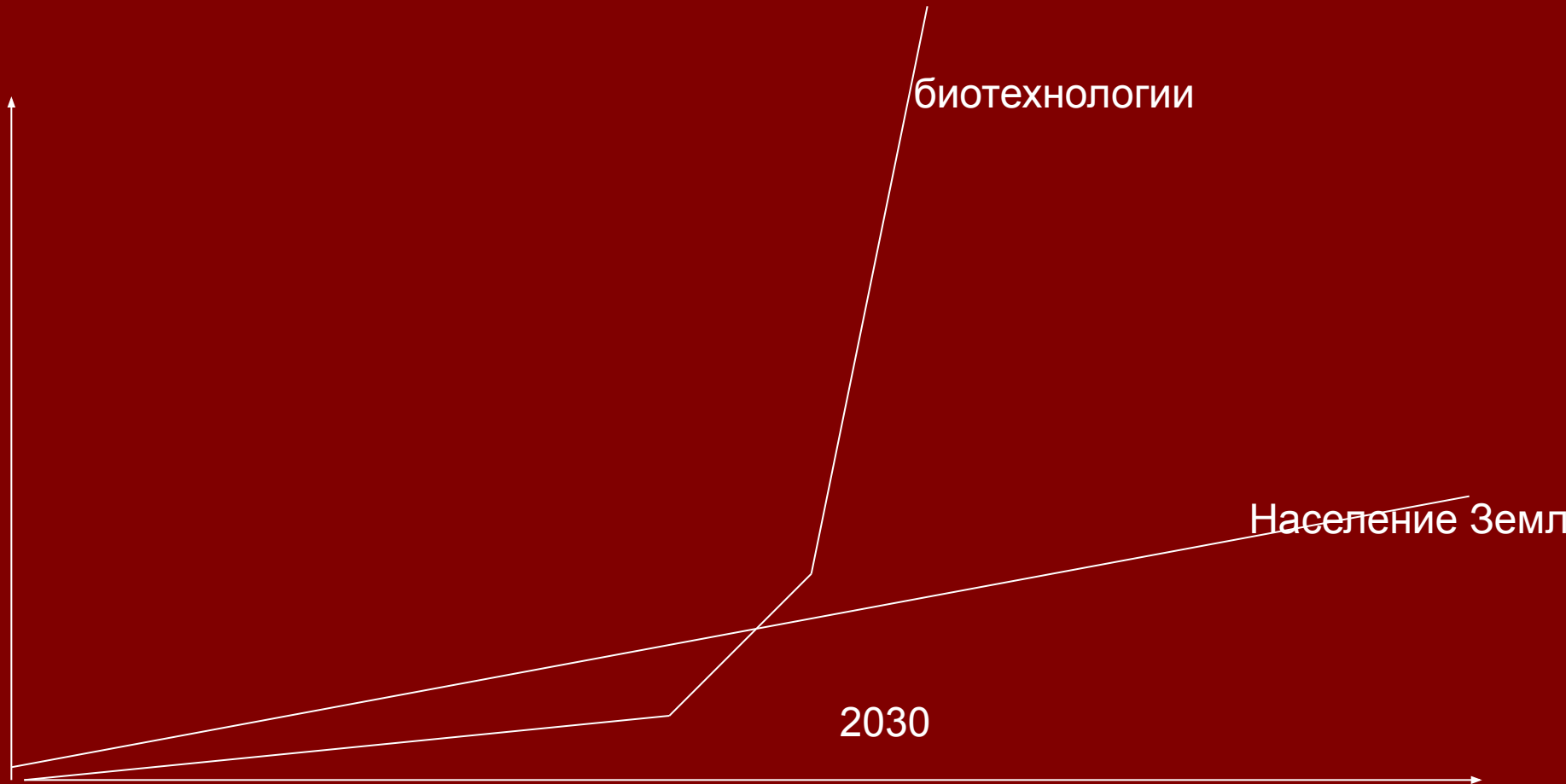
# Антропный принцип перестал нас «защищать»

- Из 100 миллиардов сперматозоидов только один был достаточно везуч, чтобы дать начало человеку.
- Но из этого не следует, что этот человек выиграет даже в орлянку.
- Прошное везение создаёт ложную уверенность в будущем везении.
- Люди систематически недооценивают риски угроз существованию.

# Национальная безопасность не тождественна защите от угроз существованию

- Миру угрожает перенаселение, а РФ – депопуляция.
- Увеличение ядерного потенциала способствует безопасности одной страны, но уменьшает суммарную безопасность в мире.
- Власть больше заинтересована в национальной, чем в глобальной безопасности.
- Национальная безопасность сильнее цепляет за эмоции людей.
- Однако угрозы существованию означают гибель всех стран.

# «Быстрые» проблемы отсекают «медленные»



# Вред от ошибочных рассуждений об угрозах существованию

- Неправильное вложение средств
- Невнимание к реальным опасностям
- Ложное чувства успокоенности
- Утрата доверия к прогнозам. «Волк,  
волк»

# Вред от ошибочных рассуждений об угрозах существованию

- Необратимое глобальное потепление
- Три закона робототехники:
  1. Робот не может причинить вреда человеку или своим бездействием допустить, чтобы человеку был причинён вред.
  2. Робот должен выполнять приказы человека, кроме приказов, противоречащих Первому закону.
  3. Робот должен заботиться о своей безопасности, если это не противоречит Первому и Второму законам.



# Три закона робототехники, компрометация:

- Ради блага всех людей всемирный ИИ отправляет их всех в рай.
- Всемирный ИИ не даёт людям путешествовать и вообще проявлять какую-либо активность, так как это может угрожать их жизни
- ИИ вводит всем людям наркотик, чтобы они были всегда счастливы.

# Необратимое глобальное потепление

- Крупная проблема выдаётся за угрозу существования
- Возможность необратимого глобального потепления не афишируется.
- НГП является системным риском, так как зависит, от множества факторов.
- Оно может произойти и после краха современной цивилизации, однако позитивная сингулярность сделает его невозможным.

# Сверхуверенность – основная причина катастроф

- Заставляет людей упорствовать в своих ошибках
- Делает их слепыми к новой информации
- Ведёт к неоправданному риску.
- Приводит к халатности

# Недостоверность экспертных оценок.

- Если эксперты утверждают, что нечто имеет шанс 1 к 10, он прав в 60 процентах случаев.
- Если эксперт утверждает, что нечто имеет шанс 1 к 1 000 000, он прав в 90 процентах случаев.

# Сверхуверенность и ошибки экспертов

- «Подобные уровни ошибок были обнаружены и у экспертов. Hynes и Vanmarke (1976) опросили семь всемирно известных геотехников на предмет высоты дамбы, которая вызовет разрушение фундамента из глинистых пород, и попросили оценить интервал 50% уверенности вокруг этой оценки. Оказалось, что ни один из предложенных интервалов не включал в себя правильную высоту. Christensen-Szalanski и Bushyhead (1981) опросили группу врачей на предмет вероятности пневмонии у 1531 пациента с кашлем. В наиболее точно указанном интервале уверенности с заявленной достоверностью в 88%, доля пациентов, действительно имевших пневмонию, была менее 20%».

# Сверхуверенность и ошибки экспертов

- «Типичный результат в исследованиях систематических ошибок состоит в том, что предложение денег или другой стимул не устраняет систематическую ошибку (Kachelmeier and Shehata (1992) предлагали жителям КНР эквивалент трёхмесячной зарплаты.)»

# Сверхуверенность и ошибки экспертов

- «Когда-то, давным-давно, я написал несколько сверхдетальных сценариев, не осознавая, что *каждая* добавленная деталь является дополнительной нагрузкой. Когда-то, давным-давно, я действительно думал, что я могу сказать, что имеется 90 процентная вероятность появления искусственного интеллекта между 2005 и 2025, с пиком в 2018 году. Это заявление кажется мне теперь полностью абсурдным. С какой стати я мог думать, что я могу определить точное вероятностное распределение для проблемы вроде этой?» Э.Юдковский

# Сверхуверенность и ошибки экспертов

- «В среднем, студентам потребовалось 55 дней, чтобы завершить свои дипломы, на 22 дня больше, чем они ожидали, и на 7 дней больше, чем они ожидали в худшем случае».
- *«Реальность, как оказалось, зачастую преподносит результаты, худшие, чем самый наихудший случай.»*