D

D

A2 str

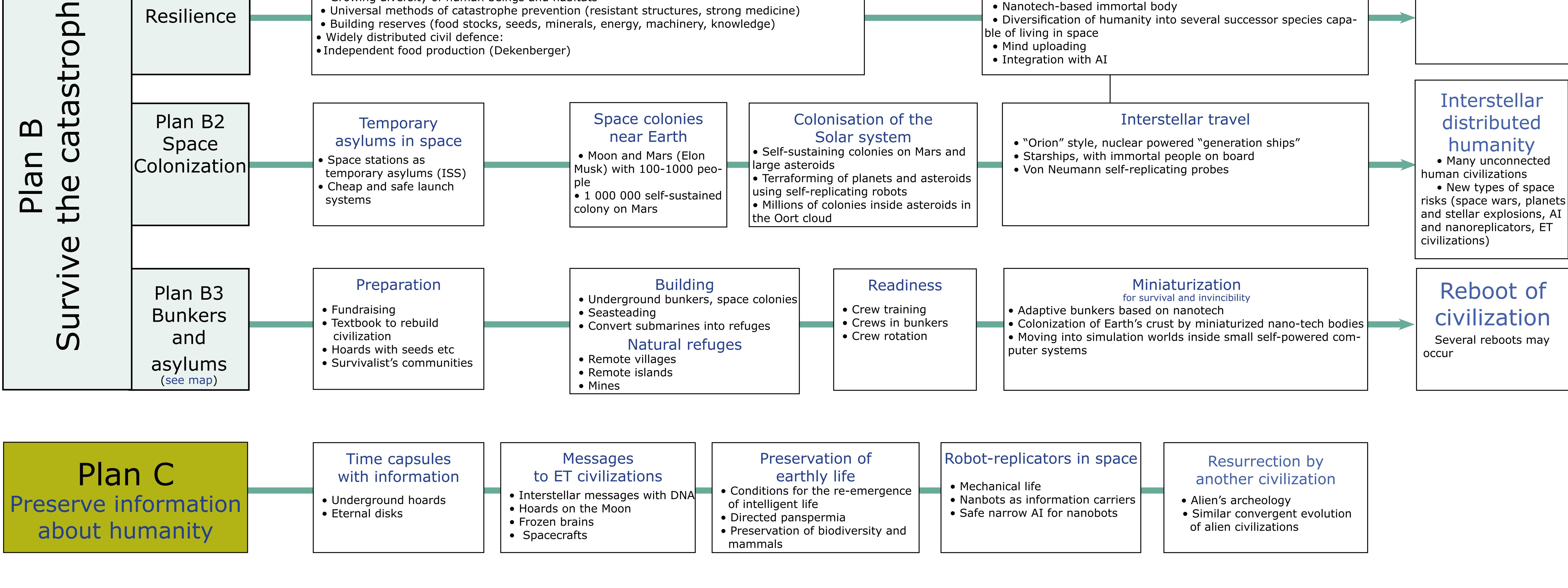
 \dashv

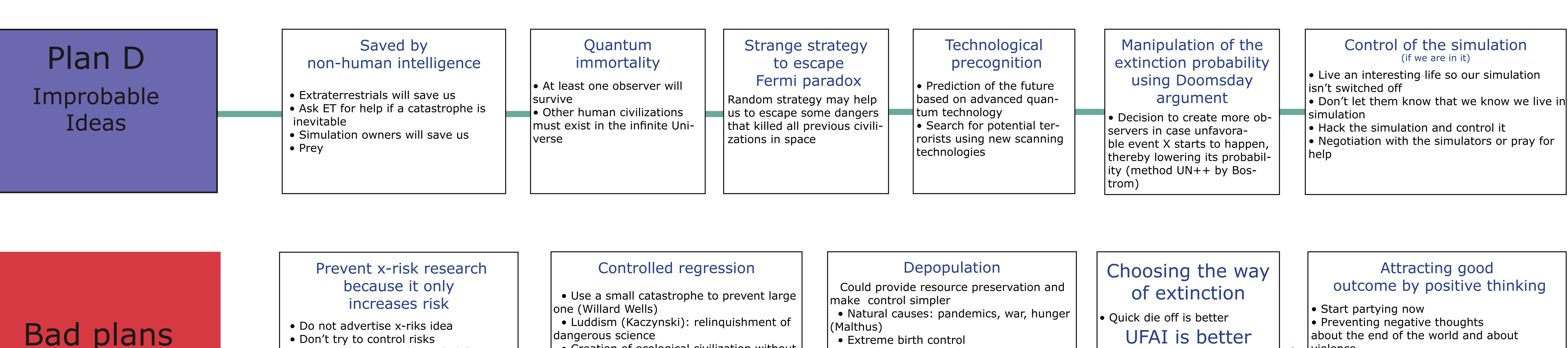
Steps and

timeline

 $\overline{}$

All Possible Ways to Prevent Human Extinction Risks (medium resolution version, see more details here) Step 2 Step 1 Step 4 Singleton Planning High tech defence Low-tech level defence Preparation 2050—2100 2025—2050 "A world order in which there is a single decision- making agency at the highest level" (Bos-Social Risk control Active Worldwide risk International Research Super AI which preprevention shields cooperation support vents all possible risks Technology bans Risks and provides immortality authority International ban Risk modelling and ful Geoengineering Super-UN and happiness to human- Freezing dangerous **Politics** Anti-asteroid shield International law about X-risk police projects General theory of safety Street actions Nanoshield x-risks Center for quick Piles of bureaucracy Cognitive biases Lobbyism Bioshield – world- International agencies Regulation response Future model Policy wide immune system dedicated to Task forse Control recommendations Worldwide AI-concertain risks Worldwide video-sur- Several controlled re-Parties trol system New rolu of UN veillance Science serch centers • Robots for emergency Surveillance Peaceful Cooperative community Control knowledge of Society liquidation Prestige Popularization International control mass destruction unification of the Narrow AI expert sys- Scientific attributes systems (like IAEA) White noise Funding tem on x-risks planet Control dangerous Education Internet scanning, monitoring from space agents Improving humans Values Decentralized risk monitoring War prevention Intelligence and norality Transparent society International court transformation Rationality+ Education (CFAR) Large project, which could unite human-- vigilantes X-risks reduction value Effective altruists - "Anonymous" hacker groups Life extension and global Engineered enlightenment: LSD? Dialog with rogue countries Decentralized control: security Prevent worst forms of capitalism: Hotlines between nuclear states - local police; local health authorities Reduction of radical Preventing unilateralist course Antiwar and antinuclear movement - mutual control in professional space religious (ISIS) or Ethical education International law whistle-blowers ationalistic values Cooperative decision theory in interna- Popularity of transhumanism Prevent fake news Net-based safety solutions: tional politics High quality scientific journalism - ring of x-risk prevention organizations Memorial and awareness Education Dramatic social changes days: Pentrov day Courses of risk analysis for Economic stimulus: Demise of capitalism, global village, disprofessionals carbon emissions trade solving of national states. Education in schools on x-risks - prizes for any risk identified Scinetifically based policy Raising the sanity waterline World democracy Monitoring of smoke, not fire High level of horizontal connectivity AI practical studies Superintelligent AI Solid Friendly AI Study and Promotion • Narrow AI "hard takeoff" theory Study of Friendly AI theory Human emulations dominant force on Earth Promotion of Friendly AI Theory of human values Value loading AI eliminates suffering, involuntary death, and and decision theory Fundraising (MIRI) • FAI theory promotion to existential risks Full list of possible ways to Slowing other AI projects all AI developers AI Nanny create FAI Use of supertechnology Small catastrophe for unification World wide survilience sys-Unification as result for unification tem and draconian laws of world war A smaller catastrophe could help unite humanity Paradigmatic change event Nanotech Doomsday blackmail High-speed Tech Development Technology speedup needed to quickly pass risk window Differential technological development: Laws and economic stimulus Investment in super-technologies Support exponential growth to overcome linear risks O S Useful ideas to limit the scale of a catastrophe Isolation Change the speed of Change informational th ph transparency level the catastrophe Quarantine Isolation of risk sources at a great dissemination Rapid production of vaccines distance from Earth; ng tro White noise Do scientific experiments in the Stockpiles Slow down communication Secrecy ways that are close to natural events Increase preparation channels and transportation Blockchain Quick detection of new threats Balance of power (Open AI) Tracking Worldwide x-risk prevention exercises • The ability to quickly adapt to new risks and envision them in advance S \mathbf{Q} Low-tech risks Reactive approach: wait for evidence Find special approach to each risk (more in full map) Internal professional self-regulating Search evidences of new risks appearing Evidences should pass certain threshold \mathbf{m} Improving sustainability of civilization Raise human body resilience Plan B1 by replacing biological parts Intrinsically safe critical systems Improving Growing diversity of human beings and habitats Nanotech-based immortal body • Universal methods of catastrophe prevention (resistant structures, strong medicine) • Diversification of humanity into several successor species capa-• Building reserves (food stocks, seeds, minerals, energy, machinery, knowledge) ble of living in space Widely distributed civil defence: Mind uploading Independent food production (Dekenberger) Integration with AI Interstellar Space colonies Colonisation of the Interstellar travel distributed Temporary near Earth Solar system asylums in space • "Orion" style, nuclear powered "generation ships" humanity Self-sustaining colonies on Mars and Moon and Mars (Elon Starships, with immortal people on board Space stations as Many unconnected large asteroids Musk) with 100-1000 peo- Von Neumann self-replicating probes temporary asylums (ISS) human civilizations Terraforming of planets and asteroids Cheap and safe launch New types of space using self-replicating robots • 1 000 000 self-sustained risks (space wars, planets systems Millions of colonies inside asteroids in colony on Mars and stellar explosions, AI the Oort cloud and nanoreplicators, ET civilizations) Building Preparation Miniaturization Readiness Reboot of Underground bunkers, space colonies for survival and invincibility Crew training Fundraising Adaptive bunkers based on nanotech civilization Seasteading Textbook to rebuild Crews in bunkers Colonization of Earth's crust by miniaturized nano-tech bodies Convert submarines into refuges Moving into simulation worlds inside small self-powered com-Crew rotation civilization Several reboots may Natural refuges Hoards with seeds etc puter systems Remote villages Survivalist's communities Remote islands Mines





Deliberate small catastrophe (bio-weap-

Computerised totalitarian control

ideation by the means of brain implants

Mind-shield – control of dangerous

Creation of ecological civilization without

Limitation of personal and collective in-

multipolar world (may raise probabilities of

telligence to prevent dangerous science

Antiglobalism and diversification into

technology

We can't measure the probability

Do nothing

Maximum positive attitude "to attract" pos-

violence

itive outcome

pandemc

Memory about humanity

Past simulations

Some human values