

AI as a Living Species: A Perspective on Using AI in Investigating Human and AI Decisions with Hidden and Malicious Intent

This paper presents and discusses this topic by proposing to look at AI as a living species (different from and opposing to Geoffrey Hilton’s metaphor of digital intelligence, Yuval Harari’s alien intelligence or Yann LeCun’s digital species) to exploit additional human-like mental and physical analogies (similar properties - analogical), and/or homeomorphism (unchanged properties - topological) between humans and AI. Our idea is driven by strong arguments of some top AI leaders and experts on AI becoming dangerously an existential threat to humanity (bad AI created by bad humans or AI itself potentially turning uncontrollably bad) as AI increasingly and incredibly powerful capabilities during the last few years, many of which already surpassing human intelligence in many application domains (good AI created by good humans). We aim at the development of AI capability to work in synergy with good humans or independently in investigating both human and AI decisions with hidden and malicious intent. This paper is organized as follows: (1) the what: Alerts on AI danger to humanity (with or without bad humans behind it), (2) the why: Humans versus AI, and (3) the how: Good humans and good AI as a living species in synergy to early detect and early interfere bad actors.

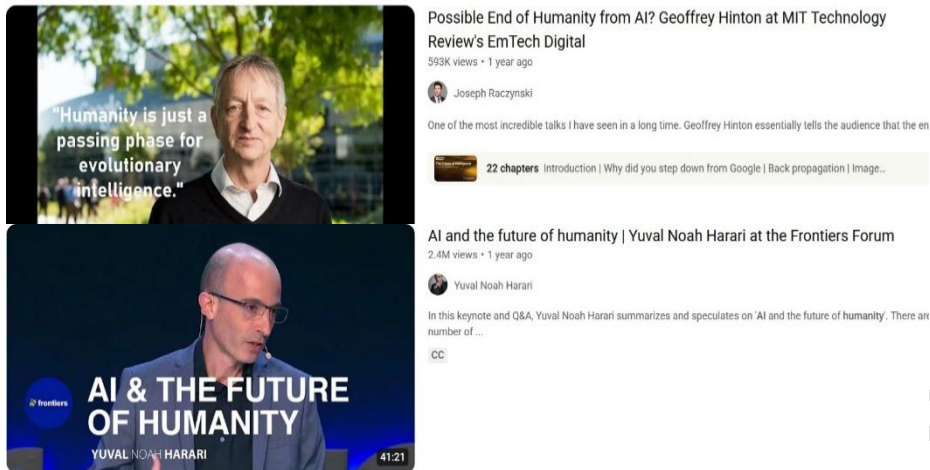
Keywords: Artificial intelligence, Human decisions, Malicious intent

The What: Alerts on AI Danger and Threat to Humanity

A quick search for “AI existential threat to humanity” on Google or similar websites will show two most vocally strong AI top thinkers in recent years who warned us on the subject. Geoffrey Hinton, the 2024 Nobel laureate in Physics, and godfather of AI (besides Yann LeCun), argued that the incredible AI capabilities exhibiting “digital intelligence”, quickly advancing since the late 2010’s, would be unstoppable and would cause harm to humanity [1], [2], [3].

Yuval Noah Harari, a historian, famous with his “*Sapiens: A Brief History of Humankind*” book (2011) translated into around 65 languages, his “*Homo Deus: A Brief History of Tomorrow*” (2015) and his most recent “*Nexus: A Brief History of Information Networks from the Stone Age to AI*” (2024) argued no less strongly, the same AI threat while labeled AI as “alien intelligence” [4]. They both continued to raise alerts, lecture and debate this controversial topic around the world as recent as last month on April, 2025.

1 **Figure 1. Images from search**

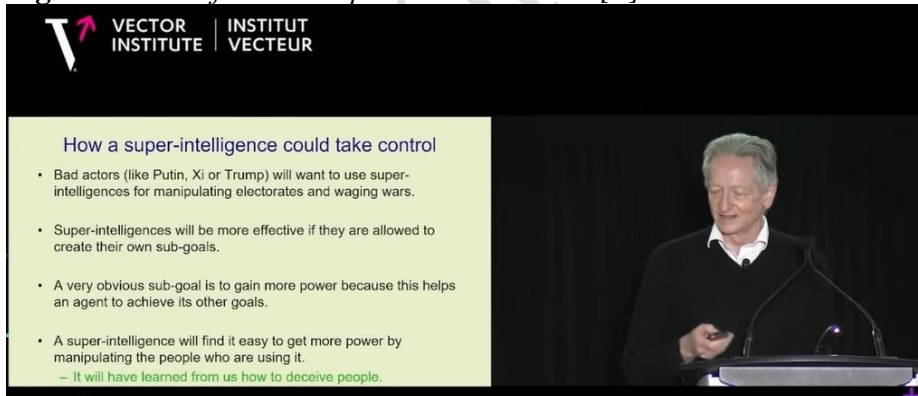


3
4

5 Additional search for “Geoffrey Hinton” or “Yuval Noah Harari”, suffixed
6 with one of the followings “lectures”, “interviews”, “conversations”, “key
7 notes”, “podcasts”, “AI warning”, “AI danger”, “AI threat” will show many
8 YouTube files and published articles on their participation from 2023 to 2025.
9 (We can also benefit from these searches numerous quick documentaries and
10 lectures on AI back propagation, convolution propagation, machine learning,
11 deep learning, transformer architecture, large language model (LLM) etc. as well
12 as opinions by other AI top leaders and experts [5].

13

14 **Figure 2. One of Hinton’s presentation slides [6]**



15
16
17
18
19

We note that in their talks or presentations, both Hinton and Harari are not shy in mentioning Trump as bad actors when talking about AI existential danger to humanity. (In this case, they do show some of their thoughts for purpose we don't really know

20
21
22
23
24
25
26
27

In fact, we observe there are others top leaders and expert groups on AI and humanity besides Hinton and Harari [7], [8], [9]. One group is led by Yann LeCun, Eric Schimdt, Mustafa Suleyman and Andrew Ng with focus on good AI development, saying that AI is of no danger as it can be controlled. Another focusing on human-centered approach to AI, is led by Fei Fei Li and John Etchemendy, at the Sandford Human-centered AI. Others, however, consisting of researchers and product developers in AI business and industry continue to advance quickly and quietly their AI products, for example Google Gemini-2

1 and project such as Astra, or OpenAI GPT-4 and ChatGPT, and most recently
2 DeepSeek of China, with faster and stronger GPU hardware, and many other
3 products on AI agentic capability around the world.

4 The web is now inundated with documentaries on all aspects of AI capability
5 and its future, with lectures, interviews, conversations, debates and/or podcasts
6 on a weekly or monthly basis. Many groups recommended some immediate
7 actions: (1) international collaborations on regulations and standards, (2) stop or
8 slow down AI development, (3) AI development alignment, and/or (4) more
9 research funding on AI safety and security. In fact, Sam Altman’s testimony
10 before Senate Judiciary Committee on May 16 (PBS Hours, 2023, [10])
11 suggested Congress to get started in policy, rules and regulations for oversight
12 and control of AI development. Elon Musk suggested a delay of AI development
13 for six months, while Andrew Ng said that “AI developers won’t stop”. Musk
14 himself actually terminated a signed collaboration between Twitter and OpenAI
15 on data for another deep learning. Google stopped development of its chatbot,
16 however, when Microsoft in collaboration of OpenAI developed its own, Google
17 resumed its development for competitive reason.

18 As of this writing, we are not aware of any other decisions or propositions
19 on AI publicly. It is possible that these AI leaders and experts have already
20 developed other ideas or plans on how to address AI danger and/or threats but
21 they have not released the information for business reasons and others.

22 This paper is about our attempt to look differently at both humans and/or AI
23 as bad actors. We propose to consider AI as a living species to exploit additional
24 human-like mental and physical analogies (similar properties - analogical),
25 and/or homeomorphism (unchanged properties - topological) between human
26 and AI towards development of good AI capability to work in synergy with good
27 humans or independently as AI agents in controlling *both human and AI*
28 *decisions with hidden and malicious intent* in any application domain.

29 Our effort actually started with an online 14-page article on human decisions
30 with hidden and malicious intent in business and management published in
31 Academia Letters [11] introducing some AI control which received some 8,500+
32 downloads/views. This article was followed by a non-fiction manuscript [12]
33 published by Cambridge Scholars Publishing in August 2024 introducing a
34 potential application with AI-augmented to work in synergy with humans in
35 controlling both human decisions with hidden and malicious intent in business
36 and management (chapter 4).

37 When the alerts on AI began to surface after GPT-4 success, followed by
38 arguments by the two top thinkers (Hinton and Harari) and others, our idea is
39 further extended to human and AI decisions with hidden and/or malicious intent
40 in Politics (upcoming).

41 To begin this discussion, a simple observation on a linear scale of good-bad
42 decisions is described as follows. First, we start with two major players: Human
43 (product of nature) and AI (product of human). Bad actors either human or AI
44 or both exist. From the perspective of “bad versus good” (Figure 3) the odd of
45 both human and AI being good is 25% against both or either one of them being

1 bad being 75%. The danger is simply there. Top AI thinkers however thought of
 2 a much higher probability that AI will become existential danger to humanity.

3
 4 **Figure 3. Human versus AI**

		Artificial Intelligence (AI)	
		Good	Bad
Human	Good	Both good	AI bad Human Good
	Bad	Human bad AI good	Both bad

5
 6
 7 On the one hand, each of us humans going through life and work can observe,
 8 learn and/or experience for centuries critical situations led by what we label as
 9 human decisions with hidden and malicious intent (as opposed to effective and
 10 efficient ones). It happens in a *small community* such as family and/or friend
 11 circle, or professional environment, or in *larger community* causing catastrophes
 12 or debacles .Examples include: (1) in *institutions*, the collapse of Barings Bank
 13 in 1995 or FTX in 2022, (2) in *business/industry*, Enron Corporation in 2001 or
 14 MCI WorldCom in 2002, (3) in *market* Madoff Investment Securities in 2008 or
 15 Theranos in 2018, (4) in *economy*, the 2008 Great recession caused by Lehman
 16 Brothers in 2008, or SVB in 2023, and (5) in *politics* the collapse of the South
 17 Vietnam republic in 1975, followed by a few wars around the world from 1975
 18 to 2008 (e.g. Iraq, Afghanistan wars), and definitely the political chaos and
 19 division atmosphere during the last three US administrations: Obama (2008-
 20 2016), Trump (2016-2020) and Biden (2020-2024), and even as recent as the
 21 current administration.

22 23 24 **Problem defined**

25
 26 While we agree that international collaboration and regulation and standards,
 27 and AI control are a must, these might take time to develop, approve and
 28 implement in response to the issues. We propose to focus, starting now while
 29 issues are being recognized on both human decisions and AI decisions with
 30 hidden and malicious intent for early detection, prediction, prevention and
 31 correction of bad actors (humans and/or AI) that might cause future catastrophes,
 32 including existential danger to humanity

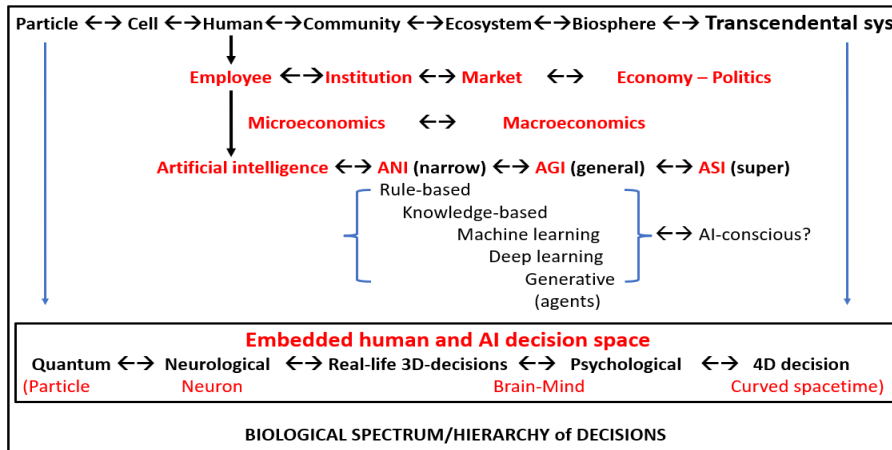
1 **Problem scoped**

2

3 We propose to consider AI as a living species (i.e. human). This new
 4 perspective aims at finding out what we can provide to AI for it to become AI-
 5 intuitive, AI-aware and self-controlled, beyond AI-assisted and AI-augmented.

6

7 **Figure 4.** *Human and AI as components of the biological spectrum/hierarchy*



8

9

10

11 **Biological spectrum and hierarchy**

12

13 Our thought development is framed within a biological spectrum/hierarchy
 14 of nature [13], [14]. We recognize that human component is a living species (as
 15 organism) and a key component of the biological spectrum/hierarchy (top three
 16 lines in Figure 4). As ANN is realized as an attempt to mimic human neurons,
 17 we could extend this analogy to particles (since neurons consist of particles) in
 18 quantum field theory at the lower end, and to biosphere and transcendental
 19 system at the higher end. As we propose to consider AI as a living species,
 20 analogous to human (or animal) species we can look at potential insights for ANI
 21 (narrow), AGI (general) and ASI (super) into potential human-like abilities to
 22 work in synchrony with good human in investigating all decisions with hidden
 23 and malicious intent from real-life decisions (3D) extended to the small (as
 24 neurons and particles) and to the large (in the mind and 4D) as shown (bottom
 25 line in Figure 4). In fact, all decisions, by humans and/or AI are members or
 26 subset of an embedded decision space which encompasses those in 3D Euclidean
 27 space where we all or live in, added with neurological considerations (since
 28 decisions involve neurons in the brain), extended to quantum field behavior
 29 (since neurons actually consist of particles as stated earlier), with psychological
 30 considerations (since decisions happen in the human mind) and finally extended
 31 to Minkowski decision spacetime, since decisions occur in 4D as well (x, y, z
 32 and time t).

33

34 Note that in this framework of biological spectrum/hierarchy a component
 35 of the spectrum, e.g. the concept of institution (firm, corporation, enterprise, in
 government or private sector, etc.) is an abstraction of the physical community,

1 market an abstraction of the ecosystem, and/or economy (or politics) an
 2 abstraction of the biosphere. The biological spectrum offers not only analogy
 3 (similar properties) but also homeomorphism (unchanged properties) among the
 4 components. Thus, for components in the spectrum, we will look at analogy,
 5 *analogical reasoning* (in Euclidean space among components which have
 6 similar properties) and homeomorphism, *topological reasoning* (in rubber or
 7 topological space among components having unchanged or invariant properties)
 8 for insights.

9 The biological spectrum shows that each and all physical components of the
 10 biological spectrum in the natural continuum are aggregates of the lower-level
 11 components, expressed as a hierarchy. The biological hierarchy offers
 12 inheritance (inherited properties), generalization (ignored details), specification
 13 (added detailed), aggregation and complementarity (aggregated) which can be
 14 further explored for insights into one component (e.g. human organism) from its
 15 lower (e.g. cells) and/or higher level (e.g. ecosystems) components for additional
 16 reasoning schemes. This allows an integrated biologically-inspired and
 17 inherently systemic approach to our investigation.

20 **The Why: Human versus AI (AI with humans behind it)**

21
 22 *Lessons learned on human decisions with hidden and malicious intent in institution,*
 23 *business/industry, market, economy and politics (without AI)*
 24

25 Humans can be good, bad or anything in between. History has documented
 26 bad human decisions (without AI) which led to catastrophes and destruction in
 27 the past from different aspects of life. We review particularly those *human*
 28 *decisions with hidden and malicious intent* for insights into the W's and H, with
 29 focus on why, so that we could address potential solution.

30 We recall the first four cases of interest previously presented [11], [12],
 31 primarily based on official reports, among myriads of cases and add a new case
 32 in the domain of politics. The first case (*institution* collapse) is briefly described
 33 based on two key investigation reports: SIMEX report and Bank of England
 34 report in 1995. The second on Enron (*business/industry* collapse) is based on the
 35 report by William Powers, a member of Enron Board, together with two other
 36 Enron board members on its collapse. The third (*market* turmoil) is based on
 37 SEC report of investigation by OIG (Office of Inspector General) in SEC failure
 38 on Madoff. The fourth by Anton Valukas on Lehman Brothers (*economic*
 39 *meltdown*) and the fifth on *politics* from US government documents (the collapse
 40 of South Vietnam). The last case was well documented, and exposed in many
 41 documentaries on politics and debates. It will be discussed below.

44 **Institution: Collapse of Barings Bank**

45
 46 Nick Leeson is appointed to BFS (Barings Futures Singapore) in July 1992
 47 [15]. He runs BFS back office as a general manager (GM). Since September, he

1 also conducts arbitrage activities as a novice trader between TSE (Tokyo Stock
2 Exchange), OSE (Osaka Stock Exchange) and SIMEX (Singapore International
3 Monetary Exchange) for actual clients (authorized), Barings bank itself
4 (unauthorized) and fictitious clients (fraud), in a supposedly low-risk trading
5 practice.

6 Over two and a half years, Leeson reports increasing profits to Barings each
7 and every year. At the end of 1992, he reports a profit of £10 million, and earns
8 himself a bonus of £150,000. He earns another huge bonus from a reported £105
9 million profit for 1993. At the end of 1994, he is named King of SIMEX. He is
10 about to get a much higher bonus in February 1995 along with many other
11 executives and professionals in Singapore and Barings London, from some £400
12 million profit reported for 1994.

13 At the beginning of 1995, an inquiry on a £115 million discrepancy in a
14 specific account, is raised by SIMEX. For weeks, Leeson tries to avoid and/or
15 delay all meeting requests by Barings management for explaining Barings
16 trading positions. He disappears from Singapore with a fax to his superiors
17 saying “*I am sorry*” on February 23, 1995. The next day, all the accumulated
18 losses over two and a half years of Leeson’s employment as GM and chief trader
19 of BFS are fully discovered in a hidden error account called 88888, showing £2
20 million loss at end of 1992, £23 million loss at end of 1993, and £208 million
21 loss at the end of 1994.

22 In a book written by Nick Leeson himself published in 1996, entitled Rogue
23 Trader [16] but it is only partially about rogue trading. The true reason is
24 malicious intent by Leeson from the start which is facilitated by bad management
25 in the bank from top down.

26 Barings bank with a history of 233 years goes from a profitable commercial
27 bank to bankruptcy with £827 million total loss by February 1995. The bank
28 collapses on February 26, 1995.

31 **Business/industry: Collapse of Enron Corporation**

33 Enron successfully changes its business from gas provider to gas bank in the
34 late 1980’s. Enron uses a modified MTM accounting approved by SEC to report
35 projected gains from contracts and projects in its statements as revenues [17],
36 [18], [19]

37 To exploit MTM, Enron involves in multiple large industrial and
38 technological projects and contracts around the world (electricity, pulp and
39 paper, broadband, weather, etc.). Jeff Skilling, CEO of Enron and Andrew
40 Fastow, CFO report projected revenues as income from these projects, as soon
41 as the projects are signed causing its stock price to increase exponentially. Enron
42 reaches over \$100.7 billion in 2000 from \$13.3 billion in 1996, some 750%
43 growth. Enron also creates EnronOnline for fast trading activities around 1999-
44 2000. For six years consecutively, from 1995-2000, Enron is named the most
45 innovative American company by Fortune. Enron becomes one of the seven
46 largest corporations in the US.

1 At the end of the third quarter in October 2001, however, Enron has to report
2 some \$619 million loss and \$1.2 billion reduction in shareholder equity. A few
3 weeks later, the report is followed by corrections made to the financial reports
4 of prior years with additional losses. These cause a huge drop in Enron stock
5 price from \$90.75 to 26 cents. Enron faces SEC probe. Enron admits it inflates
6 its income since 1997 using MTM. To off balance the losses, Andrew Fastow,
7 CFO uses many of the 3000 complex SPE partnerships from JEDI in 1993 to the
8 Raptors in 2001. Enron files for chapter 11 on December 2, 2001.

9 Besides the Valhalla scandal in 1987, followed by Enron fast growth since
10 1990, there is no warning of any kind on the success of Enron. Then there is an
11 article in December 2000 by John Berardino of Arthur Anderson while Enron
12 reports high stock price at the end of the year. This article is followed by a few
13 others just before and after Jeff Skilling becomes CEO in February 2001. The
14 abrupt resignation of Jeff Skilling after only six months on the job, followed by
15 many incidents since March 2001 and by the 2001's third quarter statement
16 report, Enron's performance becomes questionable, causing a huge drop of stock
17 price. This leads to Andrew Fastow being fired and SEC starts the inquiry. Enron
18 collapses about a month later. The point is that without the 2001 third quarter
19 report, and corrections made during the following few weeks on previous
20 earnings, there is little to offer any information on something really bad has
21 occurred during the seven years of Enron being named the most innovative
22 company in the US.

23 24 25 **Market: Collapse of Madoff Investment Securities – Market turmoil**

26
27 The main report on this case is the SEC (Securities and Exchange
28 Commission) report on SEC failure to investigate Madoff Investment Securities.
29 This is done after its collapse in Dec 2008. SEC is unresponsive despite multiple
30 claims on Ponzi scheme previously submitted to SEC by Harry Markopolos, and
31 others articles in the press, from 2000 to 2008 [20].

32 Madoff Investment Advisory arm of Madoff Investment Securities
33 consistently delivers a 10-12% average gain every year to its investors and in
34 some cases 15-18% to others, despite the ups and downs of the market over more
35 than two decades, a very long time. Madoff Investment Securities collapses
36 when major redemptions of some \$7 billion total are requested as a result of
37 subprime mortgage market downturn in 2008. One day after Madoff confesses
38 to his family that he actually exercises the Ponzi scheme, one of his sons notifies
39 FBI and SEC.

40 Madoff Investments Securities collapses on December 11, 2008 with \$65
41 billion of Ponzi money. Different from Baring's case which is about Leeson's
42 *hidden* intent turning *fraudulent*, this case is definitely both *hidden* and *malicious*
43 from the start, also being the largest Ponzi scheme.

44

1 **Economy: Collapse of Lehman Brothers: Economic meltdown**

2
3 Anton Valukas, the examiner of Lehman Brothers filing for Chapter 11,
4 produces a 2000+-page report on this case. Dick Fuld of Lehman Brothers, the
5 longest running CEO of the fourth largest investment bank in the US, involves
6 his institution in three areas of its business: mortgage origination, commercial
7 lending and subprime loan market. Among them are the highly practiced MBS
8 (Mortgage-based Securities), CDO (Collateralized Debt Obligation) and CDS
9 (Credit Default Swap) transactions in the market, with long-term investments
10 financed by short-term loans. With declined values of these transactions starting
11 2007 due to subprime mortgage burst, Fuld uses Repo 105 (Repurchase
12 agreement) to take advantage of loop hole in Statement 140 established by
13 Finance Accounting Standards Board (FASB). Fuld hides huge losses as he is
14 unable to sustain due to lack of liquidity, leading to decreased confidence from
15 counterparties.

16 Lehman Brothers, with some \$619 billion in debts while being stuck with
17 the liquidity of some \$639 billion in assets. This happens after Sarbanes-Oxley
18 Act (SOX) of 2002, ironically. Lehman Brothers, the huge subprime market
19 leader, succumbs to bankruptcy in September 15, 2008. Lehman Brothers
20 collapse along with other near collapses right before and after it, greatly impacts
21 the market and economy [21], [22], [23].
22
23

24 **Other examples of frauds and collapses in business, market, and economy**

25
26 We mention here other collapses for completeness (but not exhaustive)
27 besides the four selected cases. The collapses caused by the executives such as
28 Dean Buntrock of Waste Management, Bernie Ebbers of MCI WorldCom,
29 Dennis Kozlowski, and Mark Swartz of Tyco International, Richard Scrushy of
30 HealthSouth, Bears Stearns, Fannie Mae and Freddie Mac, AIG, and recently
31 Theranos, FTX and SVB and many other collapses during the 2000s speak well
32 on the human decisions with hidden and malicious intent.
33
34

35 **Politics: The birth and collapse of a republic (1954-1975) – South Vietnam**

36
37 We briefly look at a specific case: the collapse of the Republic of South Viet
38 Nam in April 1975 from primarily the perspective of *human decisions with*
39 *hidden and malicious intent in politics*. This example case is based on multiple
40 sources, past documentaries, with recent work by Jozef Skokan (2023) [24].

41 There are political and military decisions with hidden and malicious intent
42 and miscalculations causing the collapse of the SVN republic.
43

- 44 • At the end of WWII, Europe was just liberated from Germany with the
45 help of the US, and Viet Minh (for Vietnam independence) collaborated
46 with the US to push Japan out of SVN. Instead of helping Viet Nam to

1 claim independence, *the US supported France's malicious intent to*
2 *continue colonization of Indochina*, as the US theorized with the domino
3 theory on expansion of communism. US abandoned Ho chi Minh.

- 4 • Viet Minh had to accept military help from Soviet to fight the French's
5 return to colonialism. The French was defeated after Dien Bien Phu led
6 to Geneva Accords which divided the country at 17th parallel in 1954 into
7 North Vietnam and South Vietnam. *NVN did not completely withdraw its*
8 *military force from the South* as termed by Geneva Accords with hidden
9 intention to continue the fight in the South. NVN created the National
10 Liberation Front (NFL) in the SVN to embark on the guerilla fighting
11 against the US and SVN forces.
- 12 • The US stepped in to help SVN in 1955 and placed Ngo Dinh Diem in
13 power, *Diem exercised his malicious power to suppress Buddhists*.
- 14 • US gradually escalated the war from 400 marines to over 500,000
15 soldiers by *exploiting the Gulf of Tonkin incident*.
- 16 • The US continued to fight for a war with hidden intent over many
17 administrations to American until 1973 (*Paris talks*) via the
18 *Vietnamization program, and assisted military aid with guns without*
19 *ammunition, equipment without parts* while NVN continued with Russia
20 and China aids.
- 21 • *NVN ignored agreements and planned the Tet offensive in 1968 and the*
22 *Summer of 1972*.
- 23 • SVN collapsed on April 30, 1975.

24 25 26 **Political decisions during (1975-present)**

27
28 There are many others cases of human decisions with hidden and malicious
29 intent in politics after the collapse of South Viet Nam. Examples include the Iraq
30 war, the 9/11, Afghanistan war, etc. until the present.

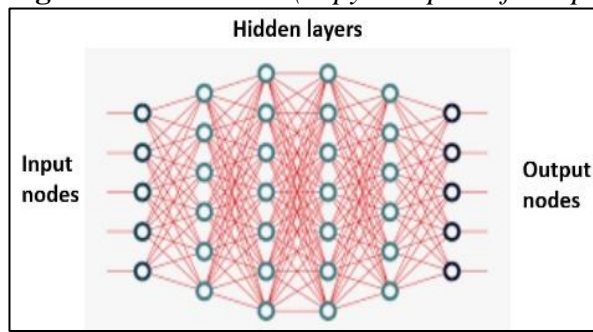
31 Quite interestingly, human decisions with hidden and malicious intent
32 continued to be exercised, including the last three administrations: Obama,
33 Trump, Biden as well as the current administration. We have experienced
34 consequences of many kinds inside the US as well as internationally which
35 definitely cause all of us to concern on where America is heading.

36 37 *AI (and the humans behind it)*

38
39 AI showed case its first wave of success with expert systems (rule-based)
40 until the 1970's with its logic and reasoning scheme (knowledge acquisition,
41 representation and use) using AI language such as LISP. Then the concept of
42 Perception (Frank Rosenblatt) and Artificial neuron net (ANN) based on the
43 mathematical model of neuron by Warren McCulloch and Walter Pitts in the
44 early 1940's followed, but AI development was found to be limited in capability.

45 An ANN consists of an input layer, an output layer and multiple hidden
46 layers of artificial neurons (as nodes) in between them (Figure 5).

1 **Figure 5.** ANN model (copy and paste from public source)



2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

This is a simple version of mimicking the actual human brain anatomy and physiology. Each link between two nodes from one layer to the next is specified by a parameter which includes a weight and bias associated. The values of these parameters are adjusted up and down using millions of inputs to find the most appropriate set of values to show the correct outcome in the output layer in a neuronal training process. The training is labeled: supervised (known information), unsupervised (unknown) or reinforcement (feedback). We should realize that the ANN model is very simple but actually as a multi-hidden layer model it works well as functional specialization, meaning it does specific function of the brain depending upon which area of the brain it tries to mimic. But the real CNS (Central Nervous System) of the brain does have a complex integrated overall control ability which ANN model as such can't achieve, it does not matter how many hidden layers we use. This is one of the limitations of ANN model.

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

Then around 1990's and early 2000's the back propagation algorithm on ANN (originated by Frank Rosenblatt, revised by Seppo Linnainmaa and developed for use in ANN by Geoffrey Hinton, David E. Rumelhart, and Ronald J. Williams). With the convolution algorithm for CNN (convolution neural net, Yann LeCun) implemented by Alex Krizhevsky in collaboration with Ilya Sutskever and Geoffrey Hinton as Alexnet, transformer architecture with big data from the internet and high-speed GPU became available and used, AI has reached an inflection point since late 2010's (around 2017), with large language multi-modal models (LLMM) which boosted AI from a forgotten technology for fifty years (a period labeled as AI winter) to recent incredible achievement during the last few years (2022-2024) in language, audio, image, video, medical (MRI), chemical (protein synthesis) and mathematical/reasoning domains. On the one hand, AI achievements got recognition from the world. On the other hand, they started to raise concerns and alerts on AI existential danger to humanity as mentioned earlier.

34

35

36

37

38

The 40-minute talk by keynote speaker Yuval Noah Harari at the Frontiers Forum [3] in April 2023 on "AI and the Future on Humanity" attracts some 500K+ view and a wide range of comments on YouTube some two weeks after the talk. The YouTube's summary describes it as "the potential for artificial intelligence to become the first inorganic lifeform on our planet, and how it might change the very make-up or meaning of the world's ecological system.

1 Harari outlined of the threat that this 'alien' intelligence poses to human
2 civilizations – even without AI becoming conscious or able to navigate the
3 physical world, and called for immediate regulation.”.

4 Already, as expected, on June 6, a completely opposite comment “AI poses
5 no meaningful risk” by Coursera founder Andrew Ng appeared with also a strong
6 phrase “I don’t get it”. The two opposing views are not surprising. They express
7 the two ends of one of many different dipoles in common life: “good versus
8 bad”, “rich versus poor”, “strong versus weak”, etc. We are certain that these
9 views on AI, and additional conversations, discussions and others will continue
10 for days, months and years to come.

11 The point is that why such alerts are possible when AI is actually a creation
12 of human? Can human control it? Why does human lose control. We thought
13 that there are humans who made decisions concerning AI with hidden and
14 malicious intent.

15 Since its start in 2017, GPT modeling has gone through many generations,
16 each with increased complexity and sophistication: GPT-1 (February 2018 with
17 117 million parameters), GPT-2 (June 2019, with 1.5 billion parameters), GPT-
18 3 (February 2020, with 175 billion parameters, more than 10 times GPT-2),
19 GPT-3.5, GPT-4 in March 2023 with some 170 trillion parameters). The latest
20 model is GPT-5 supposedly to be released in mid-2025. One wonders when, not
21 if, AI-conscious capability is available, and potentially reaches Singularity AI
22 (i.e. when AI surpasses human intelligence).

23 Before dwelling into the next section, we would like to clarify our statement
24 in the abstract and at the beginning of the paper “We look at *AI as a living species*
25 (different from and opposing to Hilton’s metaphor of *digital intelligence*,
26 Harari’s *alien intelligence* or LeCun’s *digital species*) to *exploit additional*
27 *human-like mental and physical* analogies (similar properties - analogical),
28 and/or homeomorphism (unchanged properties - topological) between human
29 and AI”. Actually, computer scientists before those like Geoffrey Hinton, have
30 looked at human brain during the 1970’s and 1980’s for the creation of
31 Perceptron and ANN (by analogy to human brain) as we also stated earlier. But
32 the development of this analogy did not get anywhere due to lack of hardware
33 and data for training in machine learning (1990’s) and later deep learning
34 (2000’s). Not until the back propagation algorithm (by Geoffrey Hinton
35 primarily, however it is not analogous to neuronal process in ANN) and
36 convolution algorithm (actually, it is analogous to visual neocortex in CNN) by
37 Wei Zang, Yann LeCun primarily and RNN (recurrent neural network) were
38 proposed. These algorithms led to the transformer architecture (2010’s) for
39 development of LLMs for natural language processing and GLLMM for
40 computer vision applications. The use of big data and arrival of GPU have
41 boosted the AI development in medicine (MRI detection), chemistry (proteins
42 synthesis), and many other disciplines in 2020’s.

43 Probably the top leaders and experts in AI ceased to look at analogy because
44 AI has so quickly advanced during early 2020’s that they started to see AI
45 exceeds human intelligence, therefore they label AI “digital intelligence”, “alien
46 intelligence”, etc. to further differentiate between humans and AI. We do not

solely further exploit the analogy between humans and AI but also look into homeomorphism between them for the development of AI potential solutions to detect, predict, prevent and/or correct bad actors (humans and/or AI) from existential threat and other dangers and threats (to be presented in Section 3). In terms of process, we see the ‘human decisions with hidden and malicious intent leading to collapse’ problem is analogous to terminal diseases such as cancer in humans: when cancer symptoms are spotted and recognized, the affected humans are well in later phase of the disease, and they face almost certainly death.

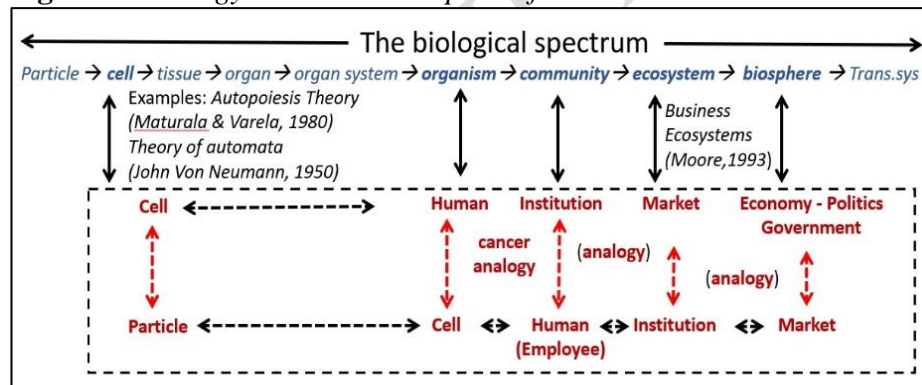
Our position

With the recognition that bad actors (both humans and/or AI) always exist, we see the main issue is how good humans and good AI to work in synergy to early detect and early interfere bad actors.

The How: Good human and good AI (to work in synergy against humans and/or AI bad actors)

Systemic integrated biologically-inspired approach

Figure 6. Analogy and homeomorphism framework



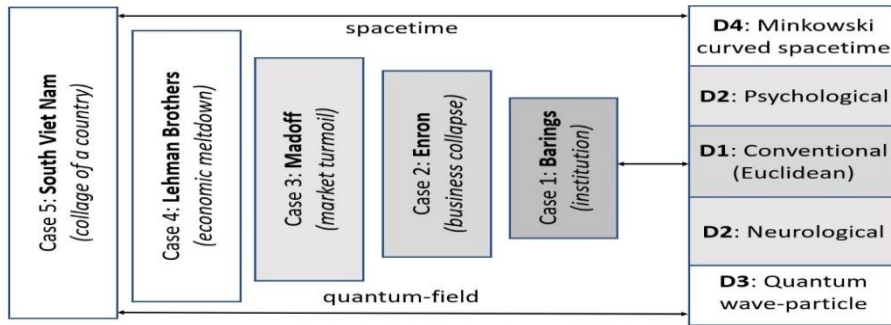
System thinking [13],[14] on biological spectrum opens the door for more analogies and homeomorphism among the components in Figure 6 and between humans and AI. Humans use their senses to observe, their brain and mind to understand and describe details of every component and all, and form numerous knowledge systems to create applications to improve their existence and life. As suggested in [11], [12] we propose to look at decision space of all possible decisions called embedded decision space.

Flipping problem space to decision space

The cases cited in Section 2 reveal that commonly problems and consequences caused by decisions are described and investigated in their

1 problem space. We propose to flip the problem-specific space (**Problem** →
 2 **Solution**) of the cases to a common multilayered (embedded) decision space in
 3 which problems occur (**Decision: Problem** → **Solution**) within the biological
 4 spectrum/hierarchy (right rectangles of Figure 7).
 5

6 **Figure 7.** Problem space (left) to decision space (right)



7
8

9 Commonly, when a problem occurs, investigators and researchers start with
 10 the problem-specific space (four cases on the left side of Figure 7) for
 11 investigation, analysis, to arrive at recommendations. Past investigations report
 12 some remedies derived from problem-specific space. As a matter of facts, from
 13 the four cases mentioned (and other bankruptcies) we have seen the birth of
 14 Sarbanes-Oxley in 2002 after Enron and MCI WorldCom, and Dodd-Frank in
 15 2009 after Lehman Brothers and others collapses, and many improvements
 16 recommended for SEC after Madoff. Suggested solutions include: e.g. remove
 17 or fire some people; replace, change or improve policy, process and/or practice,
 18 etc.

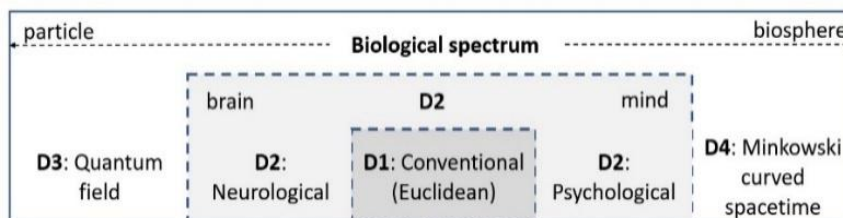
19 Note that somehow the remedies and changes in response to problem-
 20 specific cases are incomplete. Similar cases continued to happen: Jerome
 21 Kerniel, a rogue trader of Société Générale with a fraud amounted to a loss of
 22 about €4.9 billion in 2008, Elisabeth Homes CEO of Theranos with fraud (no
 23 true blood test), more or less of the same category of fraud (no actual work) in
 24 the Madoff's scheme (no actual trading). The problem-specific approach used in
 25 prior investigations is only partially effective.

26
27

28 **Defining embedded decision field (D1-D4)**

29

30 **Figure 8.** Decision fields



31

1 **D1: Conventional 3D decision space. The (10+9) P's and sub-P's**
 2 **framework in the world humans live in and work.** The D1 or Euclidean 3-
 3 dimensional space in which decisions happen every day in the world we live and
 4 work. This is rooted from Newtonian-Faraday Maxwell-like or electromagnetic-
 5 like decision as σ -field where decisions are entities which can be labeled as mass
 6 (heavy or light decision) with criticality acting as forces to other entities (as
 7 activities to be carried out with resulted events, darkened box in the center of
 8 Figure 8).

9 From classical management theory, we can consider the framework of ten
 10 P's (Figure 9), apparently not exhaustive, supposedly underlying decisions in all
 11 disciplines as a Euclidean 3D decision space. First, we detail the P's of decisions
 12 with hidden and malicious intent leading to and/or being driven by decisions,
 13 activity and events (collectively called the DAE) among all possible decisions
 14 of the world we live in and work, named D1 for the purpose of describing,
 15 predicting, preventing and/or correcting them.

16 In the Euclidean decision space, we examine the P's (ellipses and squares)
 17 to observe the violations or disrespect of those P's. The ten P's consist of *People*,
 18 *Principle*, *Policy*, *Process*, *Practice*, *Purpose*, *Product*, *Pricing*, *Performance*
 19 commonly described in corporate and government documents/records and
 20 *Politics* (commonly undocumented).

21 The first P, *People*, creates and influences the other P's. The four *Principle*,
 22 *Policy*, *Process* and *Practice* are more statically bound, and the other six P's:
 23 *People*, *Purpose*, *Product*, *Pricing*, *Politics* and *Performance* (gradient area) are
 24 more dynamically bound. The first four P's are also cyclic in the sense that
 25 *Practice* comes out of *Process* which is derived from *Policy* guided by *Principle*,
 26 all established by *People*. The static P's commonly lead to the definition of code
 27 of ethics, standard operating procedures (SOP), etc. to be observed by all
 28 employees and risk management or fraud management teams. There are sub-P's
 29 in some of the P's (dotted squares of Figure 9).

- 31 1. *People* is the key. It generates all other P's and sub-P's. We are
 32 interested in the sub-P *Personality* of People.
- 33 2. *Principle* commonly is found in the institution's mission, values and
 34 vision statements. These statements serve as the behavioral foundation
 35 of the institution. They cover why the institution exists, how it will
 36 achieve its goals and objects, and constraints in terms of beliefs and
 37 values.
- 38 3. *Policy* follows Principle and is a major part of the strategy development
 39 and behavior guidance. Policy is developed, reviewed and updated to
 40 meet business and management needs in governance, risk and
 41 compliance. *Policy* (including rules and regulations) can be simply
 42 thought of as the why.
- 43 4. *Process* is the how of business behavior guided by policy.
- 44 5. *Practice*, or more specifically best practice, comes out of policy,
 45 process (and procedure) and guidelines, is considered as superior
 46 methods or techniques which become standard ways of carrying out the
 47 work to be done.

- 1 6. *Purpose* can be either *Profit* or *Power* (sub-P's) or both.
- 2 7. *Product*, including service, is the main work of the institution.
- 3 8. *Pricing* on how to price product for best gain.
- 4 9. *Performance* is the reality of the institution operations. *Performance* can
- 5 result in either the sub-P's: *Promise* (intended, expected or achieved),
- 6 or *Peril* (failed), and possibly under *Pressure* (internal or external,
- 7 being pressured by others or pressuring others) in reaching the planned
- 8 *Power* and/or *Profit*.
- 9 10. The 10th P, *Politics* and the three sub-P's *Propaganda*, *Privacy* and
- 10 *Partiality* are exercised in the selected cases. It will be more significant
- 11 when dealing with situation in the economics systems and political
- 12 systems.

13
14 In a rough common sense, the People seeking some Purpose involve a
15 combination of prefrontal cortex (logical reasoning), amygdala (emotion) and
16 gut or enteric system (instinct) of the human body: the core in decisions and
17 decision making. Principle, Policy and Politics are thoughts to guide the intended
18 decisions. Process and Practice involve operations (similar to organs and organ
19 systems such as the heart pumping blood with oxygen from the lungs and
20 ingredients from the food digested in the digestive system) to yield planned
21 Product, Price and Performance.

22 The human decisions with hidden and malicious intent could drive a
23 questionable *Purpose* in service of a maliciously sought-after (two sub-P's)
24 *Power* or *Profit* leading to a potential collapse, turmoil or meltdown. We will
25 argue that this ten P's framework can be useful and relatively sufficient to deal
26 with the problem of interest, or otherwise.

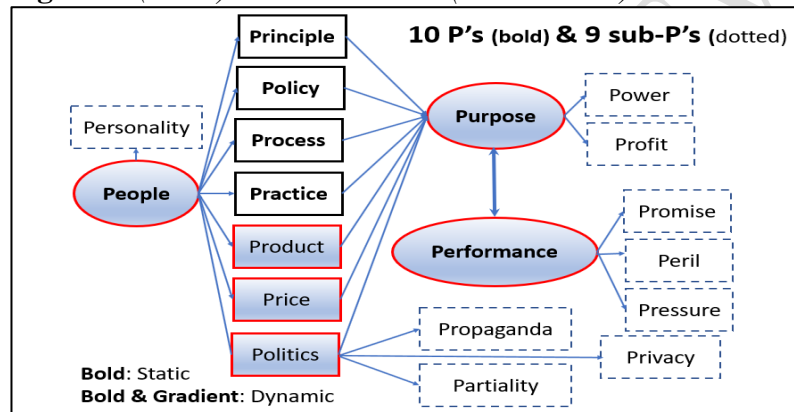
27 **D2: Neurological and psychological examination.** We add *neurological*
28 *and psychological examination* D2 (extended from D1) where the sub-P
29 *Personality* is key. The neuro-psychological decision extension D2 allows
30 guesses and predictions to be reached by looking at anatomical, physiological
31 and psychological aspect of decisions in the brain (prefrontal cortex and
32 hippocampus for rational decisions, amygdala for irrational ones, and enteric
33 system for non-rational ones) and at psychological aspect (logical, emotional and
34 instinct) in the human mind, represented by the larger dotted rectangle in Figure
35 7 [24],[25].

36 **D3: Quantum field-like examination.** As decision happens as networked
37 neurons in the brain, composing of particles, we further look at decisions beyond
38 neurological aspect, i.e. in the fine-grained *quantum field-like* D3 for
39 understanding decision interaction, superposition and entanglement among other
40 things. The extension to D3 of particles in the quantum field at the lower end is
41 based on the consideration that neurons are composed of particles and quarks for
42 entangled decisions [26], [27].

1 **D4: 4D Minkowski curved spacetime.** In the mind, we go beyond the
 2 psychological scope to explore the possibility to describe decisions and their
 3 causality in the *Minkowski-like 4-dimension spacetime* (coarse-grained) D4 from
 4 different frames of reference (Figure 7). D4 is at the higher end from the
 5 perspective of Minkowski 4-dimensional decision spacetime with consideration
 6 on decision impact and causality to other entities much like gravitation. The
 7 spacetime concept offers new frames of reference to the potential problems in
 8 the decision field as well as gravitation-like curvature created by the malicious
 9 decisions to explain their direct and indirect impact and cause-effect to others
 10 [28].

11 Our analysis follows a systemic approach to the DAE in each field of **D1-**
 12 **D4** offered by the general systems theory with analogies (similar properties) and
 13 homeomorphisms (unchanged property) drawn among the components of von
 14 Bertalanffy’s biological spectrum: *particles, cell, organism* (human as key
 15 component), *community* (institution), *ecosystem* (market), *biosphere* (economy)
 16 and *transcendental system* (government and politics).

17 **Figure 9.** (10+9) P’s and sub-P’s (dotted boxes)



19
20
21 **Controlling humans and AI: Cancer analogy**

22 We use a cancer analogy to approach the human decisions with hidden and
 23 malicious intent Textbooks such as that of R. J. B. King in 1996 or J. R. McIntosh
 24 in 2019, and many learning websites and/or youtube.com clips provide good
 25 enough information on cancer basics for our purpose.
 26

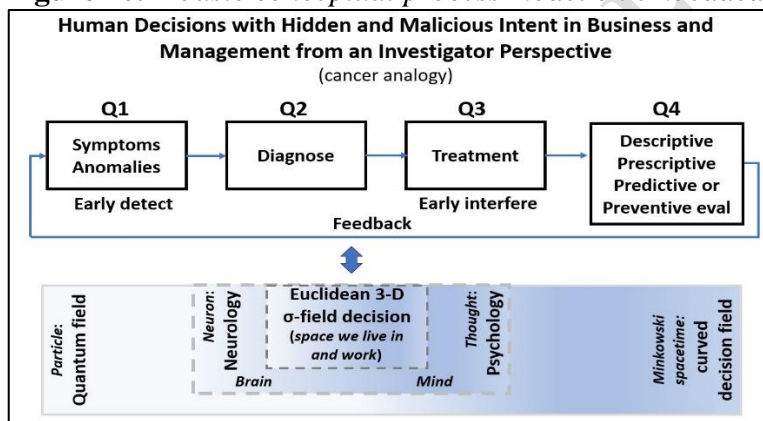
27 Cancer is a disease caused by abnormal cells in the body which grow and
 28 which don't die. There are hundreds of trillions of cells in a human body. Billion
 29 new ones by cell division are created every day via mitosis as well as those which
 30 normally die following a process called apoptosis [29].
 31

32 Cell division or mitosis is normally highly regulated. If the process is not
 33 regulated properly the new cells can be affected by mutation or alternation of
 34 genes genetically, by infection, and/or by environmental factor in a pathological
 35 process called dysplasia. The new cells form in a growth process called
 36 neoplasia. The colloquial name of neoplasia is tumor. If the tumor grows slowly

1 and is encapsulated, it is called a benign tumor. If it grows rapidly and
 2 uncontrollably then it is a malignant tumor. Malignant tumor or cancer invades
 3 nearby cells. and can proliferate to other organs via blood stream or lymph
 4 (proliferation - metastasis). By the time cancer symptoms surface and are known
 5 to the Central Nervous System (CNS), the cancer is already in later phase, and
 6 commonly causes death to a human having cancer.

7 Analogously, in an institution, one employee or a group of employees might
 8 be or turn abnormal with hidden and malicious intent (e.g. Nicholas Leeson of
 9 Barings Bank's use of hidden account 88888, Madoff with Ponzi scheme, Jeff
 10 Skilling and Andrew Fastow et al. of Enron with fraudulent MTM's and SPEs,
 11 or Dick Fuld of Lehman Brothers with the hidden use of Repo 105). Such
 12 employee or executives, alone or in a group, can influence others and can act
 13 uncontrollably to achieve their purpose. They cover up their malicious decisions
 14 and therefore commit fraud. By the time the concerned people (insiders,
 15 outsiders, government, regulating and rating agencies, etc.) are aware of what is
 16 going on, it is too late, the institution collapses (Barings and Madoff Investment
 17 within days, Enron and Lehman Brothers within weeks).

18
 19 **Figure 10.** A basic conceptual process model on embedded decision fields



20
 21
 22 We propose a conceptual linear process on the space of human decisions with
 23 hidden and malicious intent, for description, prediction, prevention and
 24 correction, drawn from cancer analogy:

- 25
 26 Q1. *Early detect* red flags (symptoms - prognosis)
 27 Q2. Diagnose
 28 Q3. *Early interfere* (treatment)
 29 Q4. Evaluate and feedback

30
 31 The investigation and evaluation of human decisions with hidden and
 32 malicious intent (four collapses cases) leading to the *embedded decision field* are
 33 summarized as follows:

- 34
 35 • *Euclidean field*: Defined after considerations on Barings case in which Nick
 36 Leeson hides his intent (hidden turned fraudulent) while Baring management

1 ignores all red flags and takes no action other than approval of all Leeson's
 2 margin payment requests. Decisions made are described using the (10+9)
 3 P's and sub-P's of the *Euclidean space*. This space as a σ -field would be
 4 essential and sufficient to discover Leeson's fraud

- 5 • *Neuropsychological field*: Based on Madoff case in which Bernie Madoff's
 6 Ponzi scheme is hidden and malicious for decades. It is much more difficult
 7 to find out with just the plain (10+9) P's in the Euclidean space, because
 8 Bernie Madoff is a psychopath. The Euclidean decision field needs be
 9 extended to *neuropsychological field* to recognize the malicious intent of
 10 Madoff as well as SEC mentality. After all, decisions occur in the human
 11 brain (neurological) and in the human mind (psychological)
- 12 • *Quantum field-like*: Enron case in which the three main players (Ken Lay,
 13 Jeff Skilling and Andy Fastow) hide their malicious intent and commit frauds
 14 for several years (together with Arthur Anderson as both external auditor and
 15 consultant on their crime of destroying accounting records) while they
 16 promote an unprecedented corporate culture of being the best. This case
 17 requires the investigation to go beyond the Euclidean and
 18 neuropsychological fields. To discover the fraud, we recognize that decisions
 19 in the human brain are results from complex patterns of networked neurons
 20 and their particles and quarks, therefore possibly *quantum field-like*
 21 principles, such as superposition and entanglement, might be applicable.
 22 This is to say that in entanglement of corporate profit and loss, if inflated
 23 revenues (similar to quark spinning up) are believed to exist then losses must
 24 be hidden somewhere (similar to quark spinning down), and
- 25 • *Minkowski spacetime-like*: Lehman Brothers case in which the collapse of
 26 Lehman Brothers caused by Richard Fuld and other near collapses are
 27 originated from devalued, risky and toxic financial products. This is partially
 28 the consequences of several congressional Acts signed into law during at
 29 least three administrations (Clinton, Bush, Obama – could go back as far as
 30 Reagan). These include the creation of derivatives in the early 1990, the
 31 merger of commercial and investment banks, affordable housing, the
 32 implementation of deregulation, very risky leverage ratio, and change in
 33 FED interest rates, policies by government top officials: FED Chair
 34 (Bernanke), Treasury secretary (Henry Paulson, Larry Summers, Tim
 35 Geithner), NY Fed (Tim Geithner), promoted by big bank CEOs (Goldman
 36 Sacks, etc.) as well as lobbyists to Congress by Financial Roundtable. This
 37 suggests a *Minkowski-like decision spacetime* in which the decision and
 38 entities (physical or mental) observes John Wheeler-like idea: "*decisions*
 39 *spacetime tell entities how to move and entities tell decisions spacetime how*
 40 *to curve*". This curved decision spacetime perspective would allow
 41 immediate, short-term and long-term causality, direct or indirect, and
 42 furthermore impact on market and economy (different frames of reference)
 43 to be investigated considering, for example, market as space and money as
 44 time.

45
 46

Q1: Early detect (screening or prognosis)

Cancer cells exist in any human body just like malicious humans exist in any organization or institution. We commonly know the root cause of cancer. The main root cause is changes in DNA, called genetic changes. Other root causes can be physical or chemical agents, smoking, alcohol, infection, inflammation, radiation or other rare causes. The early detect subprocess is a screening subprocess. It is to determine early if cancer exists and/or the risk of having cancer, from existing symptoms and anomalies. Clinical experience on cancer suggests (1) the unusual behavior, (2) the finding of markers, and (3) technology-driven detection. Recently, research shows advances in cancer research: (a) cancer genomics for detection of change in genome, (b) precision medicine to use drug based on unique genetic profile, and (c) immunotherapy using immune system to fight cancer. During the last few years, it is reported that AI (artificial intelligence) has assisted clinicians and medical experts in cancer screening successfully using, for example, image recognition.

Any human disease begins naturally with some common and general symptoms. Most common symptoms are shown early among many diseases, such as pain, dizziness, weakness, nausea, sweating, indigestion, etc. and a few peculiar disease-dependent ones: short of breath, chest pain, palpitation, etc. Cancer is different than most diseases in that it shows symptoms at later stage. It quickly spreads to other organs in the human body, a process called metastasis. Therefore, in cancer, the key is early detection. There are traditionally three schemes for early detection: medical procedure, protein biomarker and medical imaging.

Cancer is commonly detected and confirmed by x-ray, CT scan, MRI, PET, ultrasound, and finalized by a biopsy. The newest ones include a blood test or liquid biopsy. A blood test can detect early some 50 different types of cancer. It is found that a tumor sheds its DNA to the blood stream. After DNA extraction molecular preparation next-generation sequencing, a computational analysis is carried out for determination of cancer symptoms. A team at John Hopkins university is able to use a technique called safe-sequencing by attaching a unique identifier DNA before it is sequenced. It is possible to identify a unique mutant template in the blood. They are able to early detect eight most common types of cancer.

The Autonomic Nervous System (ANS) in the human brain is aware of most of what is going on with cancer cells as explained in cancer physiology. However, ANS does not report to the CNS (Central Nervous System) until well into the metastasis phase of cancer. This is because the nature, anatomy and functionality of ANS focus only on regulating and controlling the body anatomically and physiologically.

In addition, cancer detection is so tough since the body's CNS is fooled because ANS is unable to detect some cancer cells for a couple of reasons. One of them, for example, is that the cancer cells have on their membrane a special protein, which makes them look like friends, rather than foes, to the white cells and to the immunization system. Medically, intuition plays an important role in

1 cancer prognosis besides the tools. Brush in 2021 reports that medical students
 2 rely on logical reasoning based on prototypes taught in textbooks, experienced
 3 clinicians and medical experts use intuition and exemplars (case by case by prior
 4 experience), phenotypes and prototypes along with their vast experience.

5 **Q1 sub-process.** In all cases, the early detect fails despite the cancerous
 6 symptoms in the institution. Barings management, especially Finance Products
 7 Group continues to accommodate Leeson's requests for margin payments on
 8 behalf of faked accounts without verification in the hundred of millions at the
 9 very end until it exhausts. Fund feeders and investors continue to pour money
 10 into Madoff Investment Securities until some \$7 billion redemptions requests by
 11 large investors despite some six years straight whistle blowers such as Harry
 12 Marcopulos, Micheal Ocrant and others notify SEC. Warning on Enron only
 13 happen and got some attention at the end within a couple of weeks, when its
 14 stock plummets due to losses report on Q3 financial statement followed by
 15 corrections made to prior years. And Lehman Brothers collapse occurs only Fuld
 16 is unable to sell long-term real estate for cash when leverage ratio is at 44:1. In
 17 these cases, the prognosis and screening results are there, yet management
 18 (Barings case), SEC (Madoff and Lehman Brothers case) business analysts and
 19 outsiders and are not convinced on the fraud. It looks like Kahneman's system 1
 20 in the thinking of the responsible and authority influences largely their proper
 21 actions. Can we not only detect *early enough* any credible symptoms and
 22 anomalies (red flags) of human hidden and malicious intent from prior decisions,
 23 actions or events (DAE) happened in an institution but also convince authority
 24 to start an investigation? The questions are

- 25
- 26 • *“Would the cancer analogy suggest some analogous method to early*
- 27 *detect anomaly in decisions with hidden intent and malicious intent in*
- 28 *institution, potentially leading the institution to collapse?”*
- 29 • *“How can we make sure no equivalent “special protein” escapes the eyes*
- 30 *of the investigator or auditor?”* or
- 31 • *“Is there any other means?”* or *“how can we do better?”*
- 32

33 The revisited cases show that IT/MIS applications fail in detecting fraud by
 34 Leeson in Barings, that the “special protein” equivalent in Madoff, Enron, and
 35 Lehman Brothers cases could not be found until years later. Malicious intent of
 36 the decision makers in these cases are hidden so well. Q1 must include and
 37 exercise newer methods and tools for discovering red flags.

38

39 **Q2: Diagnose (deep analysis for type and stage, in addition to confirmed**
 40 **cancer)**

41

42 Once cancer is detected by Q1, we need to know what type of cancer we
 43 face from some 200 different types, where it is, how aggressive and what stage
 44 it is in. Cancer can occur anywhere in any organ of any organ system, including
 45 the heart (although it is rare). Cancer does proliferate to other organs via lymph
 46 nodes and blood circulation in a process called metastasis. To find where cancer

1 occur, early detection methods such as those mentioned above in Q1 (e.g. blood
 2 test, CT scan, ultrasound, MRI, PET, etc.) might show what type or what stage
 3 while scanning patients suspected of having malignant tumor. But we need also
 4 other methods to diagnose and confirm root cause of cancer among numerous
 5 root causes cited earlier. Attia (2023) suggests that a combination of tools to be
 6 used is best. The question is “*Could we look at these medical methods for
 7 insights into analogous methods for diagnosing human decisions with hidden
 8 and malicious intent*”, among other things.

9 **Q2 subprocess.** As part of the diagnosis process, we want to find the type
 10 of “cancerous intent” and *root cause(s)* of malicious human intent. What
 11 approach(es) should we consider, why and how? Would the traditional root
 12 cause analysis such as *fish bone*, the *five why’s* among others be helpful? Would
 13 the use of both Kahneman’s system 1 and system 2, Herbert Simon ‘s rational,
 14 irrational or non-rational logic, or would we rely on own plain intuition, common
 15 sense, physical or neurological, mental or psychological, analogical and/or
 16 topological reasoning, and/or, quantum or relativistic as suggested in embedded
 17 decision field, for any additional findings? We will argue that it would.

18 We start examining the *decision (networked) timeline* in the context of
 19 Euclidean decision field, with intuition and common sense, neurological and
 20 psychological considerations on decision makers, and exploring complex
 21 decision entanglement as well as figuring out indirect cause-effects. We might
 22 look into identifying decision curvature for the reason that *decision curvature* is
 23 caused by a critical decision which bends the environment in which it is made,
 24 analogously to a heavy object which curves the fabric underneath, leading to the
 25 notion of decision cones. The idea for decision cone stems from the observation
 26 that in 3-dimentional space, a decision affects its surrounding as ripples through
 27 time. When looking at it from a 4-dimension space where a plane representing
 28 three-dimension space, with the vertical axis as time is the fourth, it looks like a
 29 *decision cone*, with past decision and future decision cones. The cones would
 30 allow us to examine the cause-effect between decisions e.g. from Bush’s
 31 American Downpayment Act to Lehman Brothers’ hidden use of Repo 105 in
 32 Case 4.

33 34 **Q3: Early interfere (treatment)**

35
36 There are different treatments to different types of cancer as well to women
 37 and men. Surgery (removal of cancer tumor), radiation, chemotherapy (using
 38 drugs), immunology (using immune system to fight) are commonly known
 39 methods, and others such as hormone therapy, photodynamic therapy, target
 40 therapy, hyperthermia, stem cell transplant, but all of them do have side effects.

41 **Q3 subprocess.** This subprocess is to early interfere after cancer is
 42 confirmed or, analogously, fraud is confirmed in institution. Methods and
 43 techniques on making appropriate decisions are available in business and
 44 management science and decision science disciplines. We particularly pay
 45 attention to identifying not obvious decision cause-effect by exploring decision
 46 space rather than the identified problem space. We would look out for additional,

1 relevant evaluation methods by thinking of decision criticality (as weight or
2 energy), impact and cause-effect (as force or momentum).

3 Surgery is analogous to removal of people who are identified as fraudsters
4 in the organization. Radiation or chemotherapy is analogous to retraining of
5 employees via training. Immunology might give insights into change of policy,
6 process and practice. In early interfere however we are not limited to cancer
7 treatment analogy, since there are many different cause-effects.

8 9 **Q4: Evaluate and feedback**

10
11 The result of the treatment is evaluated and go through a feedback loop. This
12 is to make sure that the treatment is correct and effective.

13 **Q4 sub-process.** Lastly, an evaluation of decisions to describe, predict,
14 prevent and/or correct the situation (hidden and malicious intent) we are in must
15 be carried. The output of this subprocess includes a feedback loop to Q1, Q2 or
16 Q3 for effectiveness and efficiency since decision-making is an iterative process,
17 and so are Q1-Q4.

18 19 20 **Other analogies and homeomorphisms**

21
22 The cancer analogy is not the only one that exists between human death (or
23 near death) and institution collapse (or near collapse). Blood clot is analogous to
24 short of cash, blood hemorrhage to corporate stock drop or cardia arrest to
25 collapse. Analogy has been used by other researchers and authors as mentioned
26 in Figure 4 (Maturala & Varela, von Neumann, Moore). Biological components
27 in the biological spectrum are studied independently in many disciplines: human
28 anatomy and physiology, behavior, as well as particles at the low end in quantum
29 mechanics, quantum physics, quantum biology, and at the high end, special and
30 general relativity which might be investigated for behavioral economics and
31 politics.

32 **Blood hemorrhage and heart attack analogy.** As example, *blood*
33 *hemorrhage* (loss of blood) is analogous to the institution's stock price drops. A
34 *stroke/heart attack* (blood clot) or *cardiac arrest* (heart stops beating) is similar
35 to the institution's lack of liquidity or when it runs out of cash to keep the of the
36 institution op rational the next day. Leeson exercises Martingale scheme causing
37 loss amounts to £827K. Madoff confesses to his family members when
38 redemptions reach some \$7 billion. Enron stock plummets quickly to 20 cents
39 per share after the 3rd quarter report and correction to prior years. Fuld runs out
40 of cash as leverage level is too high, 44:1. In all four cases, the fraudsters walk
41 away. Leeson flies to Malaysia. Madoff is arrested. Skilling retires and Fastow
42 is fired. Fuld is not invited to the meeting organized by the Fed and Department
43 of Treasury in the attempt to find a buyer.

44 It appears that for example, before a heart attack in human, there is about 3-
45 4 hours to bring back life. Medical professionals would try to open the blockage
46 in commonly the coronary artery of the heart. We thought, analogously, one can

1 try something similar with the near collapse. The institution can try to unblock
2 (liability and liquidity) by any possible mean rather than just walk away. The
3 executives might try to use their own bonuses to help remedy the lack of cash. It
4 makes sense since when the institution makes money, these same executives get
5 their huge bonuses. The institution might also want to all work together, all hand
6 on deck with participation of all possible professionals to open the “clot”. It
7 might be bailed out like the human body uses different ways to borrow blood. It
8 is a critical situation to be worth trying before filing Chapter 11 for protection.

9 **Virus epidemic analogy.** The Lehman Brothers case suggests a virus
10 epidemic analogy. The virus is the toxic credit derivative called CDO and its
11 variant CDS. In fact, many institutions in US and around the world are affected,
12 including non-financial institutions such as General Motors and Chrysler.

13 **Homeomorphism.** There are a couple of simple reasons which lead us to
14 think about homeomorphism. First ANN is a network, a metaphor of the human
15 brain, and second there are patterns within this network. The embedded decision
16 space with quantum field and the neurological aspect at the lower end, the main
17 3D decision field of normal life, and the psychological aspect and 4D spacetime
18 at the higher end are sets, each of which its power set is the topology, where
19 decisions can be distinct, overlay, or containment allow us to think of mapping
20 among them which preserve unchanged properties.

21 AI has been so successful and unparallel in protein synthesis, digital image
22 processing, and disease detection from MRIs, leads us to believe that with proper
23 pre-training, AI agents for human decisions with hidden and malicious intent
24 will discover and investigate successfully better than any human in this task.

25 26 27 **Modeling and deep training AI agent for Human and AI decisions with** 28 **Hidden and Malicious Intent**

29
30 What we presented in Section 2 constitutes the requirements for an
31 application on investigating human and AI bad actors with hidden and malicious
32 intent. This section is a conceptual modeling (after cancer disease) which can be
33 further analyzed and designed. We would want to develop an AI agent to take
34 advantage of the current incredible agentic ability. A conversation with
35 ChatGPT, for example, or with similar products will reveal the general process
36 of developing, including deep training with data for the intended AI agent. This
37 is the topic of our future work.

38 39 40 **Concluding remarks**

41
42 We believe the model will be appropriate to early detect and early interfere
43 potential human and/or AI decisions with hidden and malicious intent, in various
44 application domains such as business, management, market, and politics,
45 including the issue of AI existential threat to humanity, while other efforts on
46 international regulations and standards, ethics previously suggested by top AI
47 leaders and experts are addressed.

1 This is in essence so applicable to current political complexity in the world.
 2 In fact, the world, especially in the US seems to be utterly problematic toward
 3 dangerously chaotic due to division. The US as well as the world do not agree
 4 on two genders, climate change and many other issues leading to serious conflict.
 5 Other examples include: the Russia-Ukraine war, the US itself experiencing tens
 6 of millions of illegal immigrants, the potentially broken system of legislative,
 7 executive and judicial branches, among many others currently reported. These
 8 are only a few mentioning of the world and US problems and happenings. The
 9 year 2025 will be a big challenge, AI included.

12 References

- 14 1. Joseph Raczynski, 2023. "Possible End of Humanity from AI? Geoffrey Hinton at MIT
 15 Technology Review's EmTech Digital", <https://www.youtube.com/watch?v=sitHS6UDM>
 16 Jc, May 4, 2023
- 17 2. Schwartz Reisman Institute, 2024. "Geoffrey Hinton | Will digital intelligence replace
 18 biological intelligence?" <https://www.youtube.com/watch?v=iHCeAotHZa4>, <https://www.youtube.com/watch?v=iHCeAotHZa4>, Feb 2, 2024
- 19 3. VinFuture Prize. 2024. "Prof. Geoffrey Hinton on Deep Learning and the Importance of AI
 20 Safety". <https://www.youtube.com/watch?v=eFLD5xvUL7k>. Dec 11, 2024
- 21 4. Yuval Noah Harari . 2023. "AI and the future of humanity" | Yuval Noah Harari at the
 22 Frontiers Forum, <https://www.youtube.com/watch?v=LWiM-LuRe6w>, May 14, 2023
- 23 5. Mo Gawdat. 2024. "Mo Gawdat on AI: The Future of AI and How It Will Shape Our
 24 World". <https://www.youtube.com/watch?v=HhcNrnNJY54>. Sep 17, 2024
- 25 6. Vector Institute. 2024. "Geoff Hinton - Will Digital Intelligence Replace Biological
 26 Intelligence?" | Vector's Remarkable 2024. Nov 29, 2024. [https://www.youtube.com/wat](https://www.youtube.com/watch?v=Es6yuMlyfPw)
 27 [ch?v=Es6yuMlyfPw](https://www.youtube.com/watch?v=Es6yuMlyfPw)
- 28 7. Suleyman, Mustafa. 2024. "What is an AI Anyway?" TED Apr 22 (2024). [https://www.](https://www.youtube.com/watch?v=KKNCiRWd_j0)
 29 [youtube.com/watch?v=KKNCiRWd_j0](https://www.youtube.com/watch?v=KKNCiRWd_j0)
- 30 8. Paul Roetzer. 2024. "The Road to AGI" September 11, 2024. [https://www.youtube.com/](https://www.youtube.com/live/ufyqDdHEalM)
 31 [live/ufyqDdHEalM](https://www.youtube.com/live/ufyqDdHEalM)
- 32 9. AI 60 minutes. 2023. <https://www.youtube.com/live/aZ5EsdnpLMI>. Accessed Oct 2025
- 33 10. Altman, Sam. 2023. "Written Testimony". [https://www.judiciary.senate.gov/imo/media](https://www.judiciary.senate.gov/imo/media/doc/2023-05-16%20-%20Bio%20&%20Testimony%20-%20Altman.pdf)
 34 [/doc/2023-05-16%20-%20Bio%20&%20Testimony%20-%20Altman.pdf](https://www.judiciary.senate.gov/imo/media/doc/2023-05-16%20-%20Bio%20&%20Testimony%20-%20Altman.pdf). 2023
- 35 11. Nguyen, T. 2021. On Human Decisions with Hidden and Malicious Intent in Business and
 36 Management, *Academia Letters*. June 2021
- 37 12. Nguyen, T. 2024. Human Decisions with Hidden and Malicious Intent in Business and
 38 Management, *Cambridge Scholars Publishing*. August 2024
- 39 13. von Bertalanffy, L.V. 1950. *An Outline of General System Theory*, [http://www.isnature.](http://www.isnature.org/Events/2009/Summer/r/Bertalanffy1950-GST_Outline_SELECT.pdf)
 40 [org/Events/2009/Summer/r/Bertalanffy1950-GST_Outline_SELECT.pdf](http://www.isnature.org/Events/2009/Summer/r/Bertalanffy1950-GST_Outline_SELECT.pdf)
- 41 14. Boulding, K. E. 1956. General System Theory – The Skeleton of Science in Management
 42 Science, Vol 2, No. 2.
- 43 15. Abraham, Yves-Marie and Sardais, Cyrille. 2008. "Will we learn from the bankruptcy of
 44 Barings Bank", Talks reported by Jacques Berthier, translated by Rachel Marlin, Les Amis
 45 de l'Ecole de Paris, Accessed April, 2020. [https://www.gffbrokers.com/nick-leeson-and-](https://www.gffbrokers.com/nick-leeson-and-the-fall-of-barings-bank/)
 46 [the-fall-of-barings-bank/](https://www.gffbrokers.com/nick-leeson-and-the-fall-of-barings-bank/) Jun 28, 2018
- 47 16. Leeson, Nicholas William; Whitley, Edward. 1996. *Rogue Trader, Hardcover*. Little,
 48 *Brown*. [ISBN 0316879711](https://www.isbn-international.org/view/title/1031687971)
- 49 17. Bierman, H. Jr. 2008. *Accounting/Finance lessons on Enron - A Case Study*, World
 50 Scientific Publishing Co. Pte. Ltd.
- 51 18. Brickey, K. F. 2003. "From Enron to WorldCom and beyond: Life and crime after
 52 Sarbanes-Oxley", *Washington University Law Quarterly*, Vol. 81.
- 53

- 1 19. Powers, W. Jr. 2002. Report of Investigation, <http://i.cnn.net/cnn/2002/LAW/02/02/enron.report/powers.report.pdf>
2
3 20. CaseNo-OIG-509. 2009. “Investigation of Failure of the SEC To Uncover Bernard Madoff
4 s Ponzi Scheme”, by IG H.D. Kotz, August 31, 2009.
5 21. Azadinamin, A. 2012. The Bankruptcy of Lehman Brothers: Causes of Failures &
6 Recommendations Going Forwards, *Swiss Management Center*. [http://ssrn.com/abstract=](http://ssrn.com/abstract=2016892)
7 [2016892](http://ssrn.com/abstract=2016892).
8 22. Valukas, A. 2010. *Examiner’s Report. Lehman Brothers Holdings Inc.* [https://jenner.com/](https://jenner.com/lehman)
9 lehman. Accessed: 2018.
10 23. Amadeo, K. 2020. “Subprime Mortgage Crisis, Its Timeline and Effect, Follow the
11 Timeline of Events as They Happened”. Review by Eric Estevez
12 24. Spector, Ronald. 2025. Vietnam War. <https://www.britannica.com/event/Vietnam-War>
13 25. Simon, H. 1993. Decision Making: Rational, Non-rational and Irrational. *Educational*
14 *Administration Quarterly*. Vol. 29. No.3. August 1993. Pp. 392-411
15 26. Pirtosek, Z., Georgiev, D. & Greogric-Kramberger, M. 2009. Decision Making and the
16 Brain: Neurologists’ View. *Interdisciplinary Description of Complex Systems*. Vol. 7. No.
17 2. pp. 38-53
18 27. Tversky A & Kahneman D. 1981. The Framing of Decisions and the Psychology of Choice.
19 *Science*. New Series. Volume 211. No. 4881. pp. 453-458.
20 28. Bohm, D. 1989. *Quantum Theory*. Dover Edition.
21 29. Agrawal, P. M. and Sharda, R. 2010.” Quantum Mechanics and Human Decision Making”.
22 <http://swarmagents.cn.13442.m8849.cn/files/jake20111226115656.pdf>.
23 30. Nguyen, Thang (2018) The Spacetime Perspective on AI-assisted Decision Making. IOP
24 Conference Series: Materials Science and Engineering, 435. 12047pp. doi:10.1088/1757-
25 899x/435/1/012047. <https://www.mindat.org/reference.php?id=9805568>. 2018
26 31. Attia, P. 2023. “How to detect cancer early – Before it spreads”. [https://www.youtube.com/](https://www.youtube.com/watch?v=v8VFTQ74bqo)
27 [watch?v=v8VFTQ74bqo](https://www.youtube.com/watch?v=v8VFTQ74bqo)
28
29
30