Personal Digital Twins

A third evolution step for humankind?

By Roberto Saracco

The Personal Digital Twin
by Roberto Saracco • April 2022

This ebook is a collection of posts I wrote throughout January to March 2022 in preparation for an EU workshop on Personal Digital Twins where a group of experts were asked to share their views and steer the European policy in this area.

The last part of the ebook summaries the discussions and reports the conclusions of the workshop.
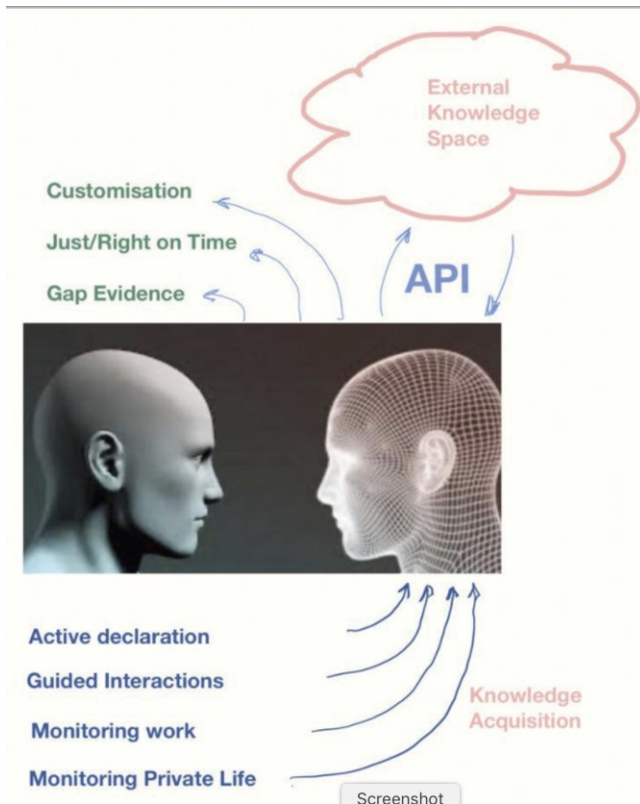
1. From "My" to "Mine"

Figure 1. My personal digital twin models some of my characteristics—in this graphic, the knowledge I have. It is also "mine" since I own it. I am the one using it to explore the external knowledge space and to highlight gaps in my knowledge that I need to succeed in and given situation I am (or will be) facing.

The concept of applying Digital Twins to model a person is already underway. Very few are attempting to model a whole person—the physical, emotional, cognitive, *and* behavioral aspects. Current Digital Twin technology aims to focus on a precise area and create a model that can be used to meet a very specific objective. It will take a few decades (this is my estimate) to create a Digital Twin of a person, mimicking every aspect of that person, but the desire for enabling such an omni-comprehensive model is highly debatable.

For quite some time, we have had companies modeling our interests: advertisement is all about knowing your target, anticipating wishes and needs, and steering consumer behavior. Profiling, in a way, has become a "science." and it is widely used. In the last few years, the adoption of increasingly sophisticated data analytics, artificial intelligence, and the growing availability of personal data has transformed advertisement from targeting a segment to targeting a single person.

I am pretty sure that Siri and Alexa have created a very good model of myself, including my voice, and are able to capture nuances of my emotional status by fine tuning their interactions with me.

In February 2022 I co-presented in an interesting webinar, with Derrick de Kerckhove and Patrick Henz, on the topic of "Social Digital Twins[1]." The social digital twin is a further facet of what can be represented by a Digital Twin.

I realized that there is a lot of work aiming at creating a digital twin of me (and you!) by third parties, but there are very limited resources that allow me, and you!, to create our PERSONAL digital twin (with personal in capital letters to emphasize the ownership, it is "my" because it mirrors me, and it is "mine" because I own and control it). The IEEE Digital Reality Initiative is working on a tool, called Knowledge as a Service (KaaS), that would allow just that—to create a digital twin of yourself, focusing on the unique knowledge you have, and any you might need or acquire in the future. This personal cognitive digital twin is yours to use.

Look at the graphic above:

- A digital model of my knowledge can be created (using the KaaS tool) through explicit (active) declarations, and through guided interactions;

---

[1] Social Digital Twins - webinar recording:
https://cmte.ieee.org/futuredirections/future-tech-forum/digital-twin-roundtable/

- The digital model can expand and be kept up-to-date by monitoring the interactions I have on the web (articles I read, searches, online courses, etc.);
- More fine tuning of the model, and tracking the evolution, can be achieved through shadowing (looking at my activities on the job and in my private life);
- My digital twin can roam cyberspace autonomously, obtaining new knowledge in areas that matter to me and in contiguous ones;
- Specific applications (data analytics, artificial intelligence) can determine gaps between my current knowledge and the required knowledge in respect to the activity/situation I am engaged in, or ones I might encounter. These applications can also be leveraged to score my competencies and compare them to people that are competing in my area. This will increase my awareness of my personal ranking on the knowledge ladder. This is something that is going to become more valuable as labor markets shift towards temporary, project-based hiring;
- My digital twin will also be able to act as a bridge towards executable knowledge, becoming an agent to link my knowledge to other people's (and machine) knowledge, as we are moving towards an executable distributed knowledge paradigm (in other words, making my knowledge relevant by associating it with knowledge owned by other entities, both human and machine).

Notice that the transformation of "my" into "mine" is also required to solve the inconsistency we are facing today with GDPR[2]: in principle, according to the European regulation, I have the right to remove all data related to my person, but on the other hand, quite a bit of this data is needed by third parties, such as Government Institutions, to interact with me, deliver services, and make me part of the social community.

Take health care data and the EHR (Electronic Health Record), for example. According to GDPR, I should have the right to delete my EHR, but at the same time, how would I get health services without it? By shifting ownership of my EHR from the Health Service to me, the EHR becomes part of my (mine—owned—controlled) Personal Digital Twin. There is no longer an issue of data ownership, and my Personal Digital Twin will be interacting with the digital twin of the medical doctor / hospital within a framework that allows the sharing of certain data.

I see this evolution as a crucial one to leverage Personal Digital Twins. They should not be considered "Personal" because they "know" about me, rather they are "Personal" because I own them. Notice the use of the plural: it is most likely that we will have several personal digital twins, each one mirroring a certain set of characteristics—for example, one mirroring my knowledge, another my health, another my interest in music, etc...

Eventually, but it will take time, they might coalesce into a single, over-reaching, Personal Digital Twin.

2. Neurotwins

The rapidly growing area of a Personal Digital Twin focusing on creating a model representing the physiology of a person, and how that specific body "works," is the first of its kind. This technology has clear, important use-cases in the healthcare realm—some examples include assessing the impact of drugs, detecting early signs of malfunction (and generating alert), monitoring general wellbeing, etc. In the latter, we are seeing an increased interest in using digital twins in athletic training[3].

---

[2] GDPR: https://gdpr-info.eu

[3] Use of DT for athletes' training: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7589903/
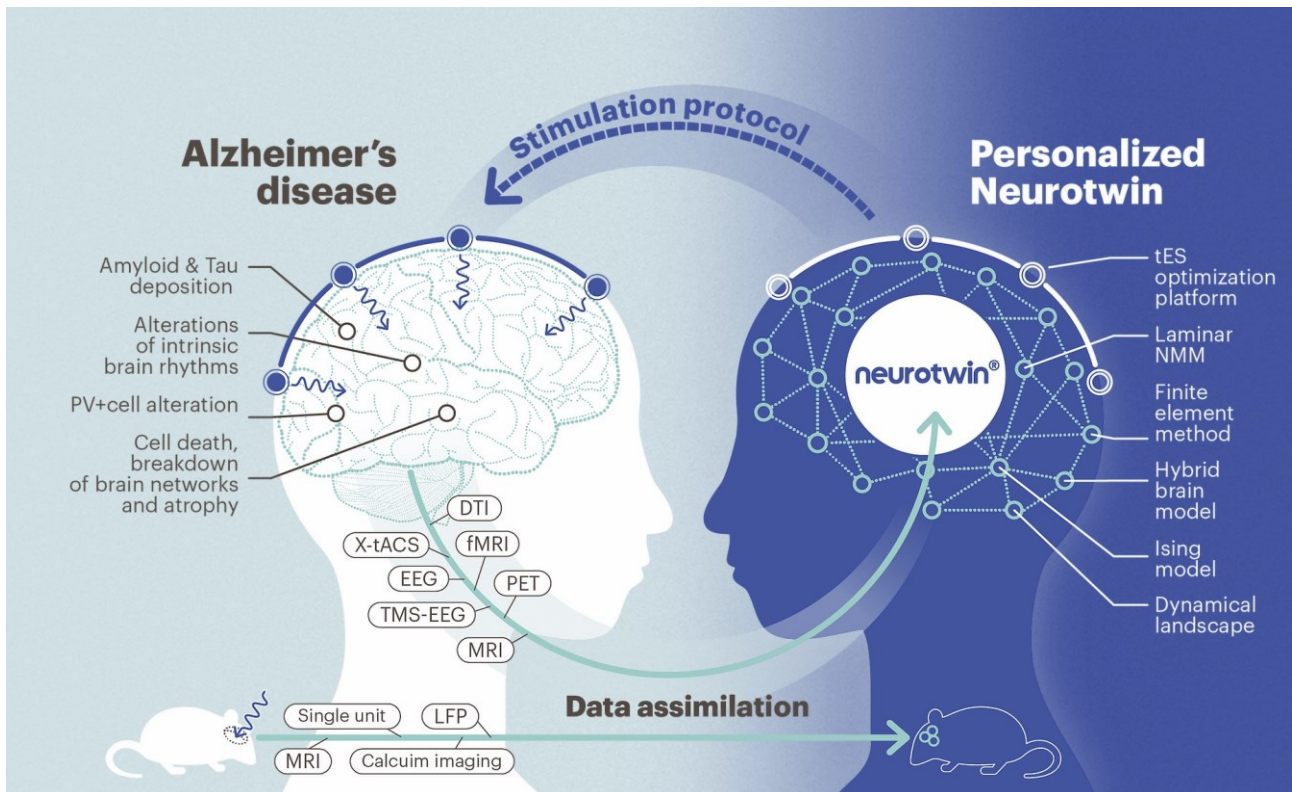
4

Figure 2: Is it possible to create a digital replica of the mechanisms of interaction of electric fields of a specific human brain? This is the goal of a European Project that hopes to use the digital replica to design personalized optimization strategies to treat Alzheimer's disease. (Image credit: Neurotwin)

The starting point of this evolution is the creation of a digital twin for a generic "body," or even more common, a part of it. Pharma is creating models of organs to test drugs in the virtual space. We have seen a growing application of Digital Twin technology in the vaccine production[4] for various tasks in the process—supporting design, manufacturing, and clinical tests all from within cyberspace. The urgent, global need to identify an effective vaccine for Covid-19 has been a powerful driver in perfecting this technology.

Studies are ongoing to create digital twins of various organs, watch the clip[5] showing the development of a digital twin of the heart.

Neurotwin[6] is another example of developing a digital twin of an organ. It is a project, funded by the EU, aiming to create a digital model of the electrical activities of a human brain to optimize Alzheimer treatment in patients. Notice that here we are seeing the shift from "digital twin" to "personal digital twin": Neurotwin aims at creating a digital replica of a very specific brain, the brain of a specific patient.

The project started in January 2021 and will continue for 4 years. It starts by looking at ways to create a replica of the electrical activity of the brain, and then aims to demonstrate instances that represent the electrical activity of a specific patient. The next step is to evaluate models of electrical stimulation and assess the impact of these stimulations on the patients' brain—first in vitro, or through simulation in the cyberspace, then by shadowing what is actually happening when the stimulations are given to the patient brain in-vivo. At this point we have all the characteristics of a Digital Twin (Model,

---

[4] Use of DT in Covid-19 vaccine design:
https://atos.net/en/blog/the-role-of-digital-twins-in-producing-a-covid-19-vaccine
[5] Development of a Digital Twin to mirror the human heart, video clip:
https://www.youtube.com/watch?v=BbgTw83z8AQ
[6] Neurotwin project: https://www.upf.edu/en/web/focus/enginyeries-i-tic/-
/asset_publisher/hPtUzEWt2uC7/content/id/243467700/maximized#.YhC5YC1h3xg

Figure 3. Graphic representation of the landscape for a Personal eHealth Assistant. A digital twin plays the role of orchestrator. Image credit: Golem IMS GmbH

Shadowing and Thread-record of events), and the ones of a Personal Digital twin (mirroring and shadowing a specific person).

The understanding gained in the design and use of the "Neurotwin" will be applied to the creation of electrical stimulators, and in the definition of the stimulation protocols, to slow down Alzheimer progress, hopefully blocking its evolution.

3. Personal caregiver

A strong drive in the consolidation of Personal Digital Twin (in the sense of digital model of a person) is coming from healthcare "in the large." In the previous section, I introduced the Neurotwin, a digital representation of the electrical activity of a person's brain and mentioned the use of Digital Twins to model organs, like the heart, created by Pharma in the design of drugs and vaccine to simulate impact (and possible side effect).

I also mentioned an example of using a Personal Digital Twin as support for athletic training, and there is a similar one used for re-habilitation[7].

Let's take a broader view: healthcare is both a personal matter (it has to have a beneficial impact on the person) and a societal matter (it involves cost, wellbeing of communities). These two aspects are connected, and digital twins can be very useful tools in establishing and supporting the mutual relationship.

---

[7] Digital twins to support rehab:
https://www.digitalfutures.kth.se/research/research-pairs/digital-twins-of-human-neuromusculoskeletal-system-a-new-paradigm-of-personalized-medicine-in-neuro-rehabilitation/

Additionally, healthcare and wellbeing are interconnected with a person's environment. A point in case is the use of digital twins in the context of home as a healthcare environment. This is shown in figure 3 above where we have the household, with all its appliances and sensors that can detect what is going on, the wearable (biomedical, but also plain wearables with embedded sensors that can generate data that may provide useful hints on the person's wellbeing), and the Digital Twin Model, in principle consisting of several digital twins (for example, the one representing the home[8], those representing the people living in the home). This Digital Twin can process data, keep learning from data (and from external interactions), and can be a bridge towards remote support provided by professionals (and even by digital twins of professionals).

The Digital Twin can provide a continuous assessment of the environment and of the individual community of people. Furthermore, the personal digital twin of each person living in that home (or more generally in that environment, it can be an office…) can provide a focused assessment on the person's health and raise red flags if something looks fishy. Notice that the personal digital twin, over time, creates a more and more accurate digital signature[9] of that person, including their gait, their tone of voice, their heartbeat… This digital signature is monitored to intercept any variation and determine if such a variation can be dependent on a change in the environment and/or early pathological signs.

Now consider the Personal Digital Twin as "my" digital assistant that can autonomously interact with the environment, sharing data (like my health status), but only to authorized parties that would be beneficial for me in that environment. Also, my Personal Digital Twin will autonomously roam the cyberspace to gather data and drive information that can make me aware of any adverse or critical situation. It can be a warning on excessive humidity foreseen in a location I have on my schedule for tomorrow that, given my health status, may raise a problem—or it may be a suggestion for some preventative drugs to decrease risk of contracting a disease in that location.

The advent of seamless augmented reality (and the video cameras that goes with it) will enable the gathering of much more data, including information on what I am eating, that can be processed to keep my personal digital twin up to date and better positioned to provide healthcare advice.

---

[8] Example of a digital twin mirroring a home, developed by the owner during the lockdown:
https://www.arq.group/i-built-a-digital-twin-of-my-home-in-lockdown/
[9] Digital Twin capturing a person digital signature:
https://www.mdpi.com/1424-8220/20/20/5936/htm
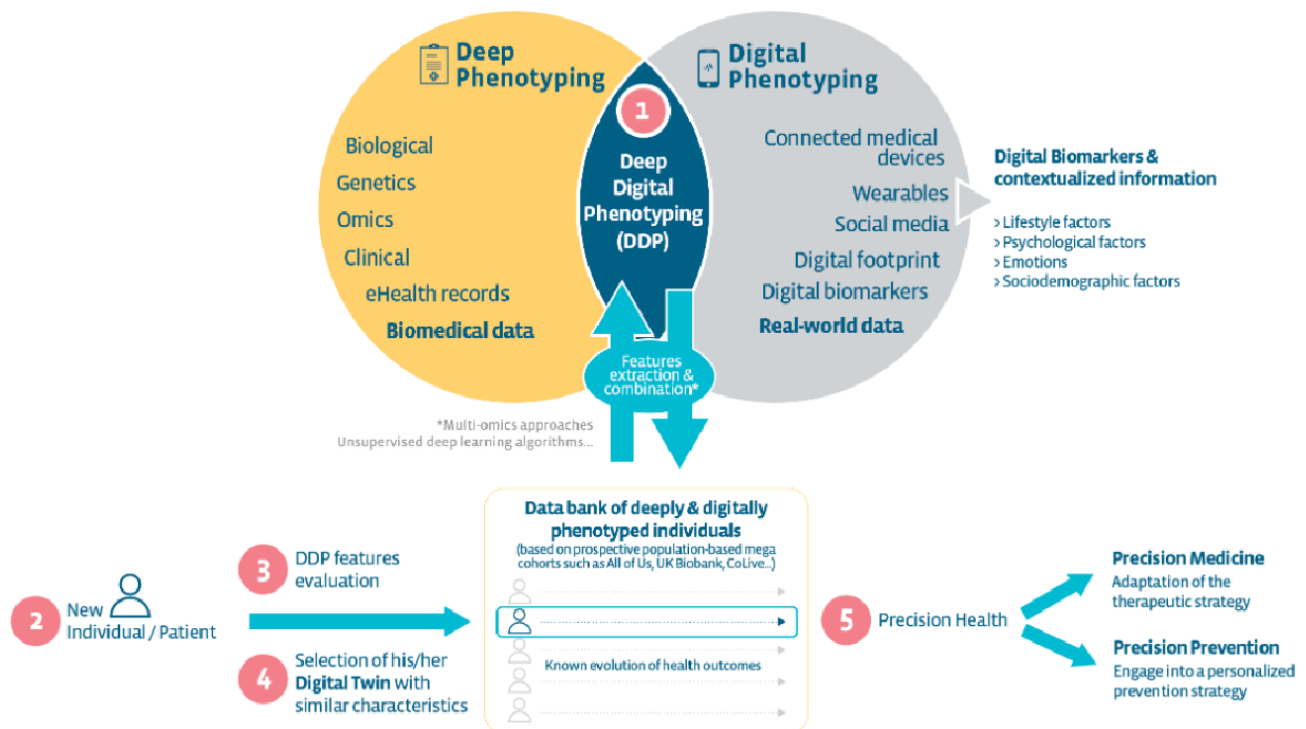
4. Digital Phenotype

Figure 5. Our phenotype is based on much more than the expression of the genotype. The whole set of "omics" is involved. Personal Digital Twins can be used to express the Deep Phenotype in a Digital Phenotype to support precision healthcare. Image credit: G. Fagherazzi et al

To what extent can we model a human body?

We have the capability to sequence a whole genome at an (almost) affordable cost, and that genome (some 3GB of data) is what we are…or is it? Well, it turns out things are much more complex. Think about the genome (DNA sequence) as a master plan that an architect has prepared to give the constructor the instructions of how to build the building. The actual masons will receive these instructions (handed over through mRNA). These are very basic instructions to build the bricks (proteins). The timing and interaction among the proteins is going to guide the construction.

To really work out the building (our body) we need the genome (genomics), the instructions sequence (transcriptomics), the proteins actually created (proteomics), and the way these are interacting "in vivo" (metabolomics), a seen in figure 5[10]. All these "omics" are required, and then some more (as I will address in the next (more futuristic) sections.

All this data, and their interpretations (information) can lead to a digital description of our body and its workings: the digital phenotype. One of the big challenges, and hurdles, that has been facing biologists is to understand the connection between the genome and the phenotype, how a difference in the genotype leads to a difference in the phenotype, and the other way around—how a given phenotype connects to a genotype.

---

[10] Creating a digital mirror of our physiology:
https://www.semanticscholar.org/paper/Deep-Digital-Phenotyping-and-Digital-Twins-for-Time-Fagherazzi/c6cb97152ec730c37c726809d0cf024163abef44

The first, genotype>phenotype, is important to:
- Understand from a parents' genotype the potential risk of an abnormal phenotype[11] (do I risk having a child with autism?);
- Detect from analysis of the foetus genome[12] the presence of an abnormal phenotype (if the newborn in question will have a genomic disorder—single gene mutation, chromosomal disorder, multiple genes mutations);
- Detect the root cause of a genetic disorder[13] from analysis of the genome and its expression.

The second, phenotype > genome, is important to:
- Find a cure by changing the genome[14];
- Personalise a cure[15] based on the specific genome.

We are seeing tremendous progress in all the above fields of application, with better screening tools (less invasive, more accurate, and more affordable). Artificial intelligence is playing a major role in this rapid evolution by allowing the (automatic) learning by sifting through millions of genomes and health records (soon to be in the billions).

Personal digital twins can become a crucial tool in the second part of this decade to mirror the phenotype (digital phenotype), thus enabling the person to be aware of specific risk, adopt a specific behavior, and enable the doctor to provide a personalized treatment, even simulating the cure on the personal digital twin before taking action on the physical subject.

It is no longer a question if we will be shifting from statistics-based healthcare to precision, personalized healthcare—the question is, when will precision medicine take the lead? Most likely in the next decade.

The road ahead is still long and winding. I explicitly mentioned autism because of its complexity. There are now some hundred genes that are likely involved in autism, and a subset of those have been recognized[16] as playing a major role. It is not just about the genes, it is about their expression, and when this is happening. New technologies allow the development in vitro of neuronal tissue and organoids, letting researchers discover and experiment the relation between genotype and phenotype. Most recently this has allowed the simulation, in vitro, of a genetic disorder leading to brain circuits found in ASD[17] (autistic spectrum disorder).

---

[11] Evaluating the genotype to assess risks genetic disorders:
https://www.nature.com/articles/s41598-019-51258-x
[12] Detection of genetic risk in the foetus:
 https://jmg.bmj.com/content/early/2021/09/20/jmedgenet-2021-107897
[13] Genome sequencing to detect root cause of a genetic disorder:
https://my.clevelandclinic.org/health/diseases/21751-genetic-disorders
[14] Gene modification to cure a genetic disorder: https://www.science.org/content/article/human-has-been-injected-gene-editing-tools-cure-his-disabling-disease-here-s-what-you
[15] Genome based personalised medicine:
https://www.wehi.edu.au/research/research-fields/personalised-medicine
[16] Autism related genes: https://www.clinmedjournals.org/articles/iacod/international-archives-of-communication-disorder-iacod-2-011.php?jid=iacod
[17] Neuronal circuits anomalies related to Autism:
https://www.wired.com/story/gene-edited-brain-organoids-are-unlocking-the-secrets-of-autism/
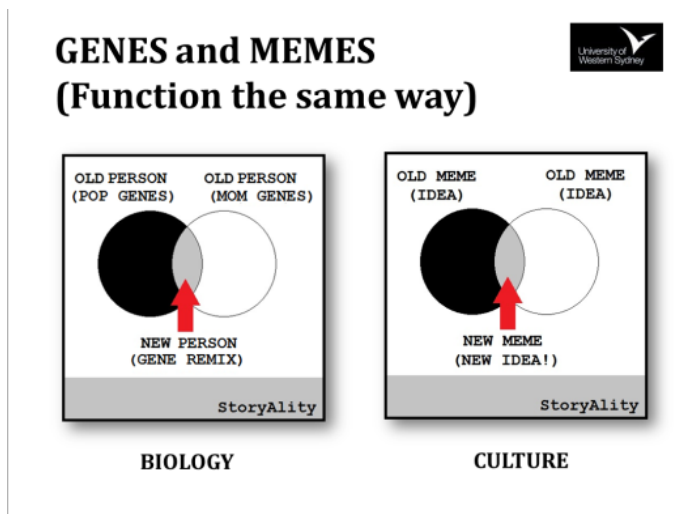
5. Memes 2.0

5.1 From Genes to Memes

Figure 6. Genes and Memes have interesting similarities in the way they evolve. They both inherit and fuse properties/characteristics of their parents. Image credit: Western University of Sydney

I described the evolution of Personal Digital Twins in terms of their capabilities to become an enhanced mirror of a person, taking into account more and more characteristics. We are seeing a lot of work in the area of mirroring physical characteristics—an effort that leverages on the growing sets of data available on our body physiology and on the possibility of receiving (quasi) real-time updates.

In the previous section, I discussed the personal digital twin as a digital representation of the phenotype (visible expression of the "omics"), and we are shifting to a grey area—one of representing not just the "functioning," but also the "behavior" of a person. From behavior to character, motivation, emotion, culture, the trend is visible.

Now, let me step into the future and try a bolder leap…

We, humans, are the result of 4 billion years of evolution—just like cats, flatworms, oaks, and the Covid virus. We may not like to be associated to flatworms, but that is the case.

However, what we see, looking back at our ancestry, is that we, humans, have evolved much more rapidly than cats, for example. We have very good testimonies provided by the remains of cats and pharaohs from Egyptian tombs that let us see how cats and humans were 5,000 years ago, and we can see that cats today are not that different in what they do and how they behave. Quite a different story for humans.

An Egyptian cat would be perfectly at ease in today's world, not so for an Egyptian pharaoh. We (our phenotypes) have evolved, and significantly so, cats have not.

Yet, the genotype of ancient Egyptians is not that different from ours. The solution (explanation) to this riddle was proposed by Dawkins (The selfish gene, The extended Phenotype) introducing the concept of a "meme" in evolution.

Memes are similar to genes, but the latter is based on DNA, and the former on Culture. When we procreate, our offspring will inherit part of our genes and genes from the other parent. Similarly, when we have a meme, or an idea, this is passed on and mixes with the memes of the receiver (ideas) taking a new shape, i.e. generating a meme that is the confluence of the two.

Evolution happens when three conditions are present:

- An abundance of a multiplicity of elements
- The elements have the capability to replicate themselves (create copies of themselves)
- The copies are not exactly the same (clones), but may differ slightly and are, in a context, where they can have different degrees of fitness, hence of success.

Notice that, as remarked by Dennett in his (very nice) book "Darwin's dangerous idea." these three conditions do not mention "gene" anywhere. Any entity and an ambient that satisfies these conditions is bound to (foster) evolution. Memes are, and we are witnessing the proof of their evolution every day.

The growing connectivity provided by the modern infrastructures multiplies the interchange of memes (ideas), and this generates further ideas in a never-ending cycle that is accelerating ever more.

Memes are the fuel for the rapid changes (in evolution scale) we are seeing in human societies. Now, this is my claim, we are on the brink of a further acceleration driven by personal digital twins and artificial intelligence. This is what I call memes 2.0.

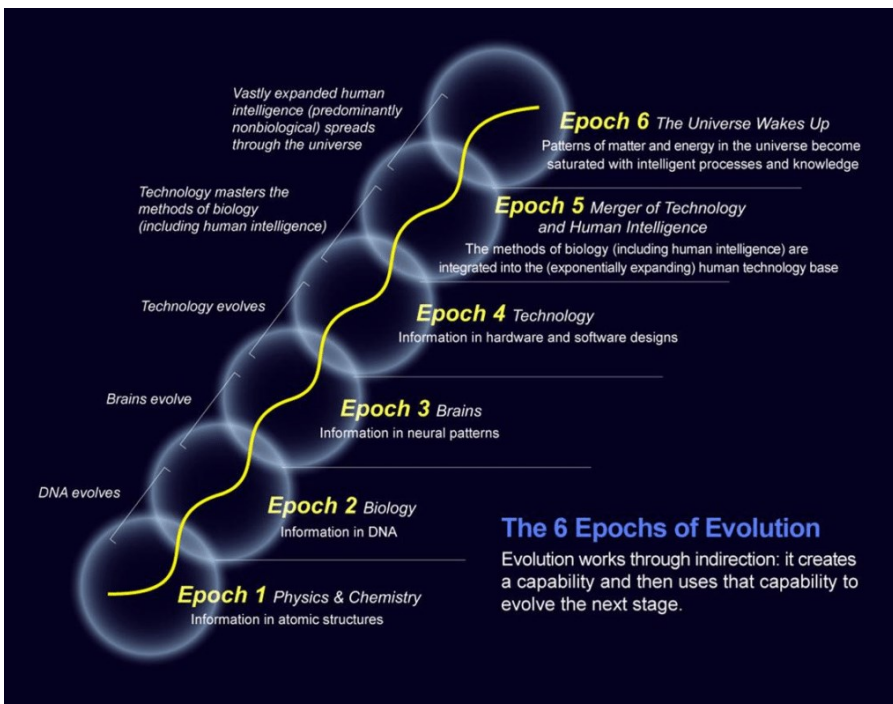## 5.2 Artificial Intelligence and PDTs Fueling Memes 2.0

Figure 7. The 6 Epochs of Evolution: from chemical and physics to DNA, to the evolution of Brains to the evolution of Technology (this is where we stand today). The 5th Epoch, we are quickly moving to, will see the merger of technology and human intelligence. Image credit: Ray Kurzweil

Ray Kurzweil has observed that evolution on Earth progressed first through the creation of ever more complex molecules. The newly created molecules interacted with the environment, changed it, and led to the creation of further molecules. This process is still going on (both naturally and artificially as I will point out in a moment). At a certain point, some 3.5 billion years ago, some molecules acquired the property of directing other molecules to assemble in specific ways (RNA and DNA). This led to ever more complex structures where organization took the lead. Cells were formed. After more than 2 billion years, this second epoch led to the creation of organized structures that can react to external conditions through the processing of information, that is neural circuits and brains (look at the overall evolution of these epochs in figure 7 as drawn by Kurzweil[18]). This marked the third epoch.

The fourth epoch, initiated about 6,000 years ago but going full steam in the last few centuries, has been characterized by the evolution of technology, with the last decades seeing an acceleration brought by electronics and software.

---

[18] The 6 Epoch of evolution as identified by Ray Kurzweil:
https://transcendentman.com/mind-boggling-technological-singularity-defined/

It is interesting to notice that there has been a continuous acceleration, with new information processing paradigms taking over the previous ones, although the previous ones did not disappear. We may no longer perceive their effect because they progress at a much lower speed, and by far, most of the impact we are seeing today is caused by the new ones (the rate at which fossil fuel is created is *much* slower than the rate at which we use (or consume) it. Hence what we perceive is just the depletion of fossil fuel resources. Also, it is notable that new epochs are influencing the processes set in motion by previous ones. The creation of molecules has been accelerated by DNA and RNA—think of all the white beaches made of shells remains … Also, CO_2 in the atmosphere is increasing (unfortunately), and new artificially produced molecules are now an ever growing presence (think of plastic -polymers- for example) due to epoch 4 (technology side-effect).

This is where we are.

However, notice that this classification in epochs is not a clear-cut partitioning. I already noticed that processes dominating in previous epochs are still present, they are just less impactful (or noticeable) in the shorter timeframe imposed by the new epoch. Hence, although we are in the 4th epoch, we are still experiencing diluted impacts from the previous 3 epochs, and most important for this discussion, we are already seeing the first signs of the next epoch on the horizon—one where we'll witness and experience the merging of technology and human intelligence (epoch 5).

Artificial Intelligence is a by-product of our current epoch. It is going to take the upper-hand leading us, and everything, into the 5th epoch—the one where humans and technology merge into a symbiotic entity.

My assertion is that Personal Digital Twins will become the glue in this symbiosis: today they are already a bridge between the physical and the cyberspace. With the embedding of artificial reality, the growing presence of IoT, and the seamless interfaces connecting our brain to the cyberspace, they will become the glue of the symbiotic relationship.

As it happened in the past, this transformation will further accelerate evolution, and the drive will be the fusion of culture and technology. We are already seeing the cross fertilization of both—our culture directs the application of technology, and technology becomes a tool to express culture and to change it through its expression (McLuhan[19]). When we look in small instances, it is very difficult to pinpoint a breaking point. We are already using plenty of technology to express our culture and our culture is evolving in step with tech possibilities (think about television, cinema… and coming soon VR, AR,….).

The ideas, or memes, that we exchange are exploiting technology as a medium and are influenced by technology in their creation and shape. It took a Julius Verne to create a meme of a rocket going to the Moon, and nowadays, that meme has become part of the global culture and has evolved into space stations, rockets ferrying goods back and forth, satellites constellations, etc.
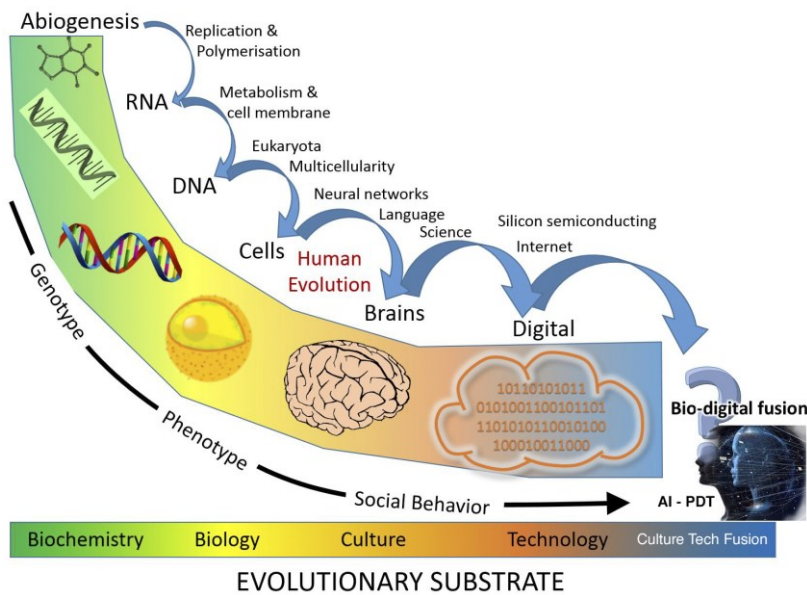
What we are starting to see is the capability of technology itself to create memes, and this is something new on one hand, but something that represents a step forward if we think about the many ways technology can be used to represent information in such a way to stimulate the creation of new information in our brain(s).

The creation of memes through and by technology can occur at a much faster pace than the ones generated by our collective brains.

This acceleration is a characteristic of Memes 2.0.

---

[19] McLuhan : https://ischool.utoronto.ca/research/institutes-labs/the-mcluhan-centre-for-culture-and-technology/

5.3 Digital Intelligence fuelling the evolution

based on: Gillings, M. R., Hilbert, M., & Kemp, D. J. (2016). Information in the Biosphere: Biological and Digital Worlds. *Trends in Ecology & Evolution*, 31(3), 180–189. http://escholarship.org/uc/item/38f4b791

Figure 8. The evolutionary substrate that has fueled evolution on our planet. From biochemistry to biology, from biology to culture, from culture to technology. We are now on the brink of a fusion of bio/culture and technology, Artificial Intelligence and personal digital twins can become the glue for a new symbiotic human. Image credit: Wikipedia modified by me

Evolution is taking place in an ambient, and it is based on specific building blocks that can interact with one another, and the result of their interactions being "evaluated" by the ambient: the results that are a better fit within that specific ambient (at that specific time) are more likely to go on fostering and directing further evolution.

The "building blocks" that have been involved in the evolution on our planet (resulting in the very many forms of life we have today, including ourselves) are represented in figure 8 (and are reflecting the classification presented in the previous section):

• Abiogenesis: generation of new molecules, and ever more complex ones, through casual encounters of molecules—giving rise to the first amino acids (organic molecules based on carbon) eventually leading to the creation of the first RNA strings;

- RNA, ribonucleic acid, that in conjunction with other molecules can direct the creation of proteins, thus creating order in the manufacturing of new components. This has given rise to the first proto-cells separated by a membrane from their environment, encapsulating the replication inside the cell;
- DNA, deoxyribonucleic acid, that can code the sequence of proteins and, in conjunction with RNA, direct the construction of proteins. The separation of tasks among DNA and RNA has sped up evolution possibilities;
- Organisms based on multiple cells, such as ourselves, and sexual reproduction that has fostered "forced evolution," what I define as the sense that any new offspring, by inheriting from their parents, has a mixture of genes (DNA) that is unique and "new";
- Brains, and more specifically, the evolved brain capable of abstract thinking and language as a tool to support thinking. This is what we mostly find in humans. The abstraction has given us the capability to change our relationship with the ambient, and even to change the ambient, thus affecting the rules of the evolution games. From this point on, the gene is now flanked by the meme;
- The creation of a digital ambient, and software that can be seen as a basic component, with self-replication/adaptation capabilities is introducing a new branch in the evolution landscape. This branch is not disconnected from the previous one—both because we (humans) are generating it, and because it can affect our environment—hence our evolution landscape.

What's next? My hypothesis is that we will not be seeing, as someone fears, this digital evolution branch taking over—rather that the evolution landscape is now consisting of both genes, memes (our culture), and memes 2.0 (the ones that are resulting from the interplay with the digital world).

I prefer the term "Digital Intelligence" over "Artificial Intelligence" to mark the difference between the intelligence of machines that is evolving towards a specific type of intelligence that is not a copycat of ours. This is flanking our intelligence, becoming a tool of the trade in our everyday lives. As it happens. those tools that are integrated seamlessly in our lives leads to a change in what we are. There is no longer a separation between myself and the tool, the tool is an integral part of me and changes the way I behave.

We have seen this happening over and over. However, so far, the tools were an extension, or an augmentation, of our physical capabilities/ Now, digital intelligence is becoming an extension and augmentation of our mental capabilities.
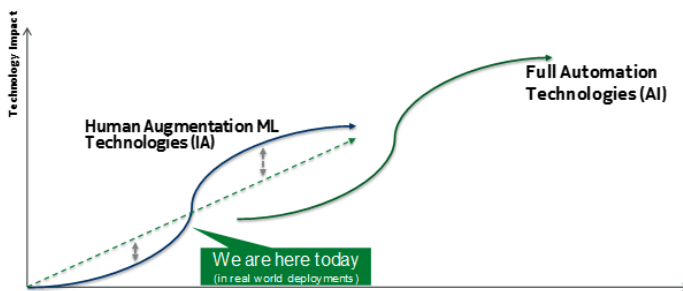
5.4 Personal Digital Twins becoming Memes 2.0

Figure 8. This diagram was included in an article 6 years ago to show the impact of AI. It is known as the Amara's law and aims at showing the fact that we overestimate the impact of technology in the short-term, and underestimate its impact over the longer term. This applies to AI as well. Image credit: CB Insights

How could Personal Digital Twins be connected to the idea of memes 2.0? When will they have an impact?

Well, look at figure 8: that curve is representing the Amara's law[20] (or, as I would say, "Amara's observation").

When we look at new, emerging, technologies, we usually perceive them as more impactful in the short term and underestimate their impact in the long term. In the short term, we are under the influence of novelty—we like it and expect everybody else will like it too, so we are foreseeing a rapid deployment/uptake. However, in reality, any new technology has to go through several gates and overcome many hurdles before becoming widespread. At the same time, we are usually unable to appreciate the big disruptions that are a consequence of a wide adoption of a technology—we tend to focus on what the technology affects directly, missing the overall picture. Hence, we underestimate its impact in the long term.

In the case of artificial intelligence, we can say we are at the green point in the graphic—we have already passed the overestimation phase of its impact (that's because we are seeing its many practical applications), but (my feeling at least) we are considering that the impact will be linear over the coming years (dashed line), more applications, better outcomes… We are not gauging the overall impact of handing over many activities to AI, hence likely underestimating its impact in the coming (5) years.

As AI becomes "the norm," it will disappear from our perception (it already has, you no longer give image recognition capability a thought, or natural language recognition of your devices (smartphone, television…). Just wait a few more years, and it will become embedded everywhere.

One place where it will be embedded is Personal Digital Twins (PDTs—represented in the graphic by the second curve), starting lower than the AI curve, but overtaking it in a few years (indicating a stronger impact).

---

[20] Amara's Law on foresight: https://www.avivainvestors.com/en-gb/views/aiq-investment-thinking/2019/10/amaras-law-applying-an-old-adage-to-new-technology/

14

It is most likely that devices that have embedded AI, like our smartphone, will also be hosting our personal digital twin. I do not see this happening in the next few years, but I can see that happening by the end of this decade. In the short term, we are going to see our PDTs provided, hosted, and "owned" by third parties (most likely they can start as an evolution of profiling by Netflix, Amazon, Walmart, Google…). At this time, it is not clear to what extent these third parties will be willing to share our PDTs with us so that we can use them. I bet that as soon as one starts, the rest will (have to) follow. Among those third parties, I would include insurance companies, healthcare institutions, and (a little bit later) educational institutions. As PDTs expand their reach and increase their value, you can expect your employer to start using them (exploit your PDT).

At this point I can envisage service companies offering us the capability to create our PDT, manage it, and, most importantly, exploit it. This is likely to happen in the second part of this decade (my estimate, but I would love to get yours…).

By the end of this decade, I expect many people will have their PDTs—starting with college/university students and young professionals. They will constitute an interesting (and growing) market that will stimulate both the generation of services based on the PDTs, and their further adoption.

In particular, and this is crucial to my point, this will stimulate the evolution of the PDTs and of their relationship with us.

5.5 Personal Digital Twins as Humans 2.0

We have artificial intelligence, and it keeps growing in capability and breadth. It will reach a point (through basic extrapolation) where the AI we are creating, and using, will match human capabilities: Artificial General Intelligence (AGI).

There are already areas where current AI matches human (expert) capabilities and even exceeds it, like in image recognition. It used to be an area where computers lagged, but now, they have the upper-hand—like in mammography and CT scans (image recognition used by self-driving cars is sometimes ahead of human drivers, sometimes behind...).

For sure AI has been making progress in understanding what an image is about, recognizing its various components, and deriving meanings (for example, there is a dog barking, and the little kid is scared).

The next step, and most likely a quick one to take after AGI, will be Artificial Super Intelligence (ASI), an intelligence that goes beyond our human intelligence. When will this happen? It is difficult to say because the boundaries are not defined, there are a lot of grey areas in the definition of intelligence, and there is also a big question mark in relating machine intelligence with our brain intelligence. For us, understanding is a crucial part of an intelligent behavior— understanding what is happening, why, and understanding what might happen depending on what we do (or independently of what we do). A machine (known as a Turing machine) can behave in ways that are very similar to what we would do, and even in smarter ways, but that does not mean that the machine understands what is going on... nor that it is using the same reasoning processes we would use. You can use a machine to translate from one language to another without the machine understanding what it is being translated (just by using string matching). That is not the case for a human translator.

Because of these differences between machine intelligence and human intelligence, (some) researchers have abandoned the idea of creating an artificial intelligence that mimics ours—rather they are aiming to get machine intelligence that is useful and that, in some cases, may substitute ours (because it has an equivalent "output"), and in other cases complement ours.

Figure 9. This diagram represents a possible way of creating ASI. It would result from the integration of intelligence at the edges (millions of intelligent nodes) as a basic knowledge base, a system knowledge base, and an experiential knowledge base. An interesting proposal that would go hand in hand with the idea of autonomous Personal Digital Twins. Image credit: Matjaz Marussig

Now, take the personal digital twin. We can imagine our PDT evolving to the point that it is almost an exact replica in terms of our experiences, our way of analyzing the world, and our way of making decisions. That PDT would show an equivalent level of intelligence (since it is mimicking our behavior). A perfectly matching PDT would be indistinguishable in its behavior from our behavior, hence it will demonstrate an AGI mimicking ours.

This can be mapped onto figure 9 above, taken from an interesting article [21] by Matjaz Marussig describing a possible way to achieve ASI by integrating intelligence at the edges (in the article, that intelligence is provided by IoTs). Please note that I am not claiming that the article is making my point, just that it is proposing a way to create ASI that could be used as a reference in my discussion.

In our case, we could imagine millions of PDTs, each one equipped with AGI, interacting with one another, thus generating an emerging intelligence that is greater than an individual AGI, an intelligence that falls under the ASI definition since it would be greater than our intelligence.

This emerging intelligence would not be separated from ourselves (from my own intelligence) because our PDTs are seamlessly connected with (each of) us and they are part of our intelligence. My claim is that my PDT will be living with a symbiotic relationship with me (which calls for a symmetrical statement of myself living in symbiosis with my PDT). Therefore, my behavior and the level of intelligence that I demonstrate is influenced by my PDT: this means that the ASI being created by the ensemble of the PDTs becomes part of myself. My behavior will become, in a way, Super Intelligent. This of course raises plenty of ethical and societal questions, but that will be saved for part of a different discussion.

It is also crucial that the ASI intelligence has less constraints than an intelligence that is tied to the physical world, and to analytics applied to and in the physical world. This kind of intelligence can evolve much more rapidly, creating an avalanche effect.

This is the memes 2.0—a rapid (exponential) evolution of knowledge and executable knowledge—meaning that our capability to interact, smartly, with the world will increase at an accelerated pace. We are already seeing the first sign of this accelerated executable knowledge mediated by machine intelligence: the time required to develop a vaccine, or a drug, is sharply decreasing, the time to improve material characteristics (the search for better alloys) is decreasing, etc.

I am offering these possibilities in a positive manner: we are going to be smarter because we will be able to leverage on smarter machines that will be in a symbiotic relationship with us. Symbiosis is

---

[21] A possible architecture generating ASI: https://medium.datadriveninvestor.com/artificial-super-intelligence-asi-could-it-rise-spontaneously-d0df6fd110fc

the crucial point: machines are not going to substitute us as they get smarter, nor are we going to become obsolete because of the increasing capabilities of machines. We will keep, as we did in the past, leveraging on smarter machines to become smarter ourselves.

Humankind has accelerated its evolution by flanking memes to genes. In the (near) future, humankind will further accelerate evolution by having memes 2.0 (resulting from a flanking of AI to our brains leading to IA - Intelligence Augmentation) flanking memes and genes.

Memes are characterizing our (different) culture, memes 2.0 will have a deep impact on culture, and, because of this, on society and societal relations.
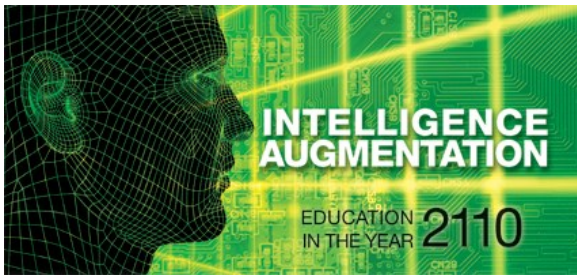
5.6 Intelligence augmentation

Figure 10. What if AI is not an independent entity, but becomes embedded in our brains? i.e. it becomes Augmented Intelligence. The Alberta Teachers' Association is exploring its implications. Image credit: Amara D. Angelica

Personal Digital Twins, as I envisage them, have the power to transform our progress in Artificial Intelligence into Intelligence Augmentation (IA). This is a very bold claim, but I am not alone in my thinking. I took figure 10 from an interesting, very forward looking article [22] discussing education in 2110. Of course, talking about 2110 is a nonsense, it is just an excuse to project our thinking in the future, overcoming all "practical" barriers (you could forecast ten years into the future, but you would have to take into account current constraints that are resolved/nonexistent looking 100 years ahead—you can well that today's constraints will no longer apply).

The basic point made by the article is that we are moving towards an augmentation of our intelligence through an artificial one. This means that there will be a seamless continuity between our brain intelligence and the artificial intelligence to the point that, from a perceptual point of view, the artificial one will not be seen as a separate entity. Think about looking inside a microscope (or a telescope). What you see is what you perceive as reality. The tool in between your eyes (brain) and the object disappears, and you feel you are looking "directly" at that object.

Likewise, this is what we expect will happen with Augmented Intelligence. We will be smarter, and it will be seamless. We HAVE Augmented our Intelligence.

In order to achieve that, artificial intelligence must be readily available. In this case, the issues are not in the future existence of an AI, but pertain to the trust we have in that AI and in the access to it. The trust involves many societal and ethical issues (from the very basic question "what does it mean to be intelligent?"). Intelligence is related to both the analysis of context and to the actions taken to reach a goal. For example, we need to cut healthcare costs—since most of the cost is related to elderly care let's stop treating the elderly.... (JUST TRYING TO MAKE THE POINT!). The access involves technology, and today we are still far from having technologies that can provide seamless, continuous access to artificial intelligence.

As I proposed in previous sections, we could embed AI into our PDT and then connect them in cyberspace to augment AI to ASI level, but how can this PDT become part of ourselves (of our brains "thoughts")?

---

[22] Education at times of Intelligence Augmentation: https://www.teachers.ab.ca/News Room/atamagazine/Volume-91/Number-1/Pages/Intelligence-Augmentation.aspx

A (wireless) Brain Computer Interface (BCI) may be a solution, but so far, and for quite some time in the future, we don't have seamless and effective BCI.

Senses-mediated BCIs seem closer, more like Augmented Reality (Virtual Reality in this context can play a role, but a much more limited one since, in general, we want to live in the physical world and not be decoupled from it).

Smart glasses and, even better, smart contact lenses, supplemented by a yet to be invented effective visual interface can pave the way towards Augmented Intelligence by providing additional stimuli to our brain (by far most of our thinking and decisions involve visual stimuli—seeing, reading ...). This is most likely the channel that will be used by the end of this decade, most of the time, to connect our brain to the cyberspace, and to merge AI with our brain, resulting in intelligence augmentation.

There are, of course, many questions not just about the feasibility, but also on the implications, and many more will be arising as we get closer. The following are things to consider:

- The adoption of IA will be in niches, and there will be plenty of "have-not." Therefore, there will be a gap, and because of the few benefitting from IA, that gap will likely widen rapidly.
- If we don't know exactly what intelligence means, how can we define augmented intelligence?
- We can communicate with one another because we are on the same "platform" (brain), and part of our interaction is based on the implicit understanding of the other persons "reasoning" process. Is the adoption of IA going to hamper bi-directional communications (I can assume that an augmented brain can understand a plain, vanilla brain, but would the reverse be true)?
- Would the use of Intelligence Augmentation leave me at loss in "understanding"? I could, for example, use IA to converse in Chinese, but I may not be able to actually appreciate what I am saying... (this already applies to real-time translators that you can only trust!)
- Assuming seamless, wireless BCIs become a reality (we are probably talking more than 20 years from now), would I be aware of how much my reasoning is "internal" and how much is "external."? In other words, would I be able to distinguish brain supported thoughts from "cloud" supported thoughts? Notice that already today we can feel a multitude of thoughts floating in parallel in our brains (in our conscience), and several times we are playing one against the other. Would this kind of interaction or competition be possible among thoughts generated outside my brain, but to all effect, perceived as part of my brain?
- My (augmented) intelligence will evolve very rapidly, beyond my capability to follow it. Today, I must digest new knowledge, tomorrow it will come to my fingertips in a flash. I can very well become inconsistent in my behavior since tomorrow I will be acting based on a knowledge set that may be significantly different from the one I have today.

The list is really long, I chose the above as examples that may not be coming to mind immediately.
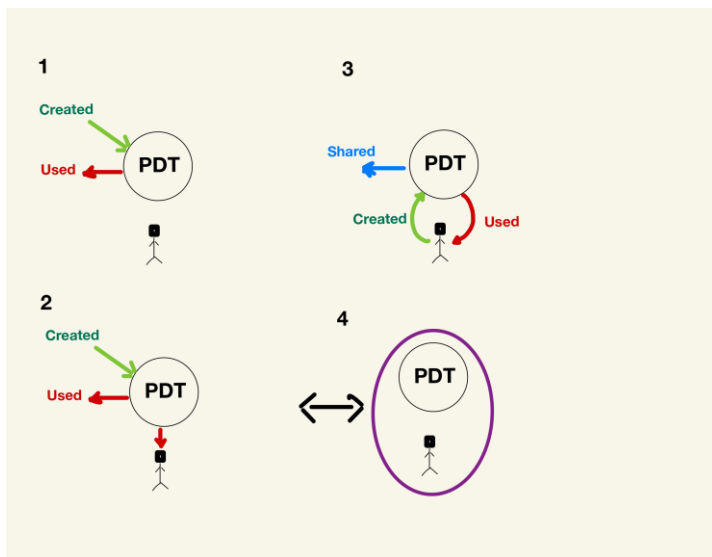
5.7 Roadmap

Figure 11. Main evolutionary steps for PDTs

Time to take a look at the expected overall evolution. For that, I tried to write down the main evolutionary steps and the leading players in figures 11 and 12.

Today, we are basically at stage 1—the Personal Digital Twin is an evolution of profiling, and it is created by companies that have an interest in "mirroring" some of our characteristics. They create "our" PDT using the data they have available (and usually they have quite a lot of it—much more than we think!), and "our" PDT is theirs (!) for them to use. Anytime they have to interact with us (or push some offer to us), they do so by looking at our PDT.

Hopefully, in the conning years, a few of these companies will start giving us access to "our" PDTs (the European Data ACT is, at least partially, addressing this). My personal take is that we will start accessing and using the PDT created by medical institutions (our EHR and the prescriptions and counsel provided by our doctors…). In a few years, we will begin feeding this PDT data that we will directly provide, for example from wearable sensors, thus updating our physiological status and sharing behavioral data (how much I walked, exercised, eaten…). This first "timid" extension of the PDT under our control may open the door to the next stage, stage 3, where my PDT (the one that represents part of my being…) is created by me and it is under my control.

This transition could be accelerated by:

- Technology: Providing easy to use tools (and in most cases automatic tools) to create and update my PDT (I expect a significant contribution will come from consumer electronic companies, Apple is probably well-advanced in this area—think about the health app and how it connects to the smartphone and Apple watch);

- Regulators: May enforce the sharing of personal data collected by third-parties, and possibly share some of the functionality that is being developed from my data.

Stage 4 is immediately after stage 3, of course, but it is far down the line. At stage 4, we will be so intimately connected with our PDT that our behavior (how we are perceived AND the way we perceive the world) will be the result of the seamless interactions continuously flowing between us and our digital twin. It is probably so far in the future that it might be a moot point to discuss now. Yet, some of the ethical and societal implications would be better addressed well in advance. Also, there won't be a pivotal moment when the shift from stage 3 to stage 4 happens, it will be a grey area that won't be noticed, similarly to the smooth transition from the moment we lived without a cellphone and the moment we can no longer live without it.
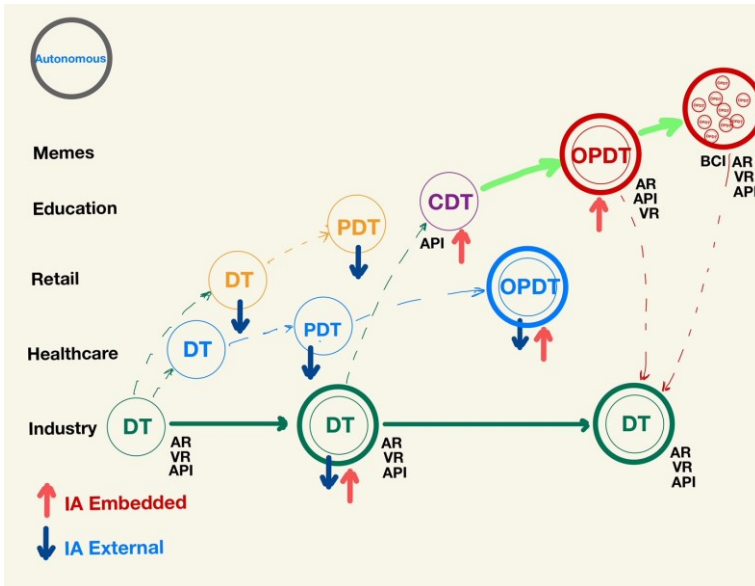
19

Figure 12. The evolution towards memes 2.0

This points to the "when" question—What is the likely roadmap that will steer this evolution?

Digital twins have been adopted, and finely tuned, by industry (over 15 years ago). Industry has set up a new environment to host and leverage digital twins (the digital transformation has been the conceptual landscape). Other areas have, more recently, started to adopt (with every little modification) the digital twin paradigm. Most crucial for our discussion on personal digital twins is the healthcare and retail domains where digital twins have rapidly expanded to include the modelling of people. This is shown by the dashed lines in figure 12. Through the evolution, we saw Digital Twins first, and Personal Digital Twins (more recently) being used as data repositories onto which apply artificial intelligence. This has provided Intelligence Augmentation (IA) to digital twins, making them smarter in terms of awareness of the physical twin situation and context. Most of this IA is derived from external applications (platforms play an important role in this). The access to the data spaces and the interaction with the digital twin modelling capability takes place via API, whilst interaction with human users takes place via AR and VR.

We have also seen the emergence of the Cognitive Digital Twin, a digital twin that embeds knowledge and can transform that knowledge into an executable one. First proposed by IBM to model smart robots with machine learning capabilities, and now extended for application in the education domain, relating to people.

The next step is towards Digital Twins autonomy in the industry which has already started with autonomous robots, autonomous warehouses, autonomous factories, etc. This autonomy is represented in figure 12 by the solid circle surrounding the digital twin. The reason for this representation is that autonomy is provided through a shell (software and hardware) to make sense of the environment, analyze the goal, and find the appropriate way or actions to reach it. Intelligence is now becoming embedded in the Digital Twin.

A similar, although for the time being only at a research stage, evolution can be expected in healthcare (Covid-19 accelerated this evolution). We will see a shift from the PDT owned by the medical institution, to an Owned PDT (OPDT-owned by that person) that becomes autonomous in delivering services on-demand (including continuous monitoring of the persons health status). This autonomy rquires the capability to connect through the cyberspace to data centers, sharing with them the required data, often in a completely anonymous way, and when required, booking services, becoming the interface to its physical twin (the person).

Even further down the line is the evolution of CDTs into OPDTs (red circles). These will act as knowledge centers flanking the person to deliver executable knowledge on-demand, in a seamless way. In the long, long term, this OPDT will become an integral part of the

knowledge space of the person, similarly to how the smartphone has become an integral part of our communication space.

In perspective, and in line with a global evolution towards a distributed knowledge that is accessible here and now in an executable form, we will see these OPDT as virtual clusters of knowledge entities, each one attached to a specific physical entity be it a human, an institution, or a machine. This OPDT will grow exponentially—at the rate of world wide knowledge increase—becoming both the landscape in which we live, and the tool delivering the needed knowledge (memes 2.0).

6. Personal Digital Twin: European effort

The European Joint Research Centre (JRC) has become interested in Personal Digital Twins, and they founded *MyDigital Twin*—trusted Personal Digital Twins in a Transformed Society.

A few days ago, the JRC held a (virtual) meeting to connect industry experts (invitation only) in attempt to address the:
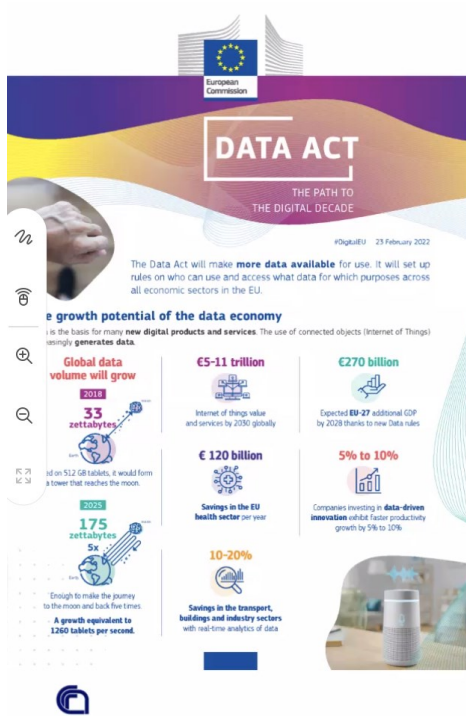

Figure 13. The expected growth of the data economy in Europe and the role of the EU Data ACT in making data available to society and businesses. Image credit: EU

• Issue of using, supervising, and controlling the rapid generation of personal data via PDTs;

• Role of PDTs in understanding complex/fast societal dynamics;

• Challenges relating to ethics and privacy;

• Opportunities to extend and advance public administration services;

• PDTs role for EU strategies (AI ACT, DATA ACT, European e-Identity).

Personal Digital Twins are seen as providing a new technological edge by creating a digital alter-ego to all citizens...the other "digital me."

The Covid 19 pandemic has catalyzed the demonstration of gaps in Governments knowledge on the behavior of society actors. This gap strongly contrasts the huge amount of personal data stored in many places on the web that could be used to create PDTs. These PDTs could be used to create a framework—supporting tools to provide an understanding of societal dynamics.

Additionally, the PDTs promise to play a key role in the creation of advanced e-Gov and private platforms and services (with an earlier focus on personalized digital and Health services).

PDTs are also considered instrumental in the fostering of Smart Cities and Local Digital Twins (usually clusters of DTs mirroring a specific, local reality).

What I find interesting is that an institution like the EU considers PDTs mature enough to be considered for regulation at the EU level. In their statement on "Why" EU should be looking into PDTs now, they list:

- Digital Twins are a Pillar of the Digital Transformation of Society: a basic component in the virtualization, datafication, and connectivity;
- The Rise of the Cyber-physical domain: the pandemic has further increased the volume of personal data (in the exabyte range today) creating the fabric that unifies the digital and the physical space through the "digital twin" pattern;
- The EU Next Generation and the Green/Digital Transition Strategies: data and its capacity to be leveraged by both by institutions and single citizens is a must.

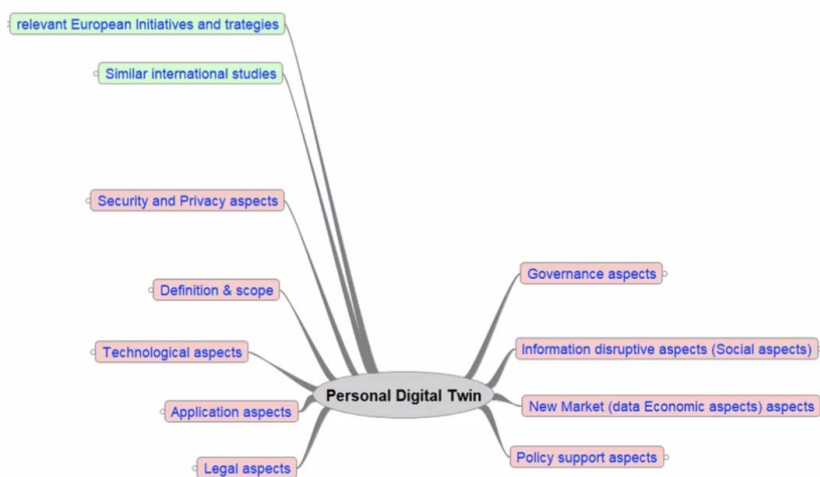## 6.1 The EU regulatory framework for PDTs

Figure 14. The approach to the regulation of Personal Digital Twins takes into account all the related aspects—from tech considerations to the application domains. It also takes into consideration the ongoing initiatives in Europe and abroad (pale green areas). Image credit: JRC EU

As shown in Figure 14, the task assigned to the experts' group on Personal Digital Twins was to look at the various aspects related to PDTs with a broad perspective/open mind.

There is obviously a technology aspect (the one I focused on) considered by the group, more in terms of reference framework (like technology platforms, AI-ready data, sharing of private/company data spaces, etc.), but that is just one among many:

- *Governance*: Including Citizen control of personal data and personal info spaces, people's acceptance (like the acceptance/refusal of Covid tracking apps), personal vs. platform driven control, and governance.
- *Information disruptive aspects:* Including behavior information, web service personalization, personalized extended reality.
- *New markets, data economics*: Including behavioral data and behavioral change recognition.
- *Policy support:* Including post-Covid society, information to monitor medical emergency situations, information to monitor economic and social crises, how to move from personal information to statistically meaningful ones generating a "synthetic" population.
- *Legal:* Including legal limitations, moral limitations, legal best practices, personal info spaces.

- *Application :* Including personal environments with ambient intelligence like home and office, smart cities, and smart places supported by urban data platforms at different levels of granularity (from neighborhood, to city, to region)
- *Security and privacy:* Including e-Identity and GDPR.

The first focus, as could be expected, was on the "definition and scope." Consensus (it might seem obvious, but it is not) is that a PDT is first of all a Digital Twin, and as such, it inherits a DT structure:

- Mirroring of (certain aspects of) the physical entity - *digital model*
- In synch with the status of its physical entity (with respect to the mirrored aspects) – *digital shadow*
- Keeps record of the evolution of the physical entity (with respect to the mirrored aspects) – *digital thread*

The PDT can act as a butler (assisting the physical entity), as an avatar (impersonating the physical entity), as an agent (like harvesting information on behalf of its physical entity. In certain situations, it can act as the digital placeholder (like in storing the person's health record).

The EU interest is au pair with endeavors run in sseveral places around the world, and it is connected to several activity areas in the EU framework like the:

- data free flow[23]
- European database[24]
- AI regulatory framework[25]
- e-Identity[26]
- BARDA[27]
- GDPR[28]
- Green deal strategy[29]

6.2 Internet of People

One of the first points the group agreed on was that a Personal Digital Twin is tied to a very specific "person"—its physical entity. To what extent does the PDT mirror the physical person? Well, that basically depends on what is the expected "use" of the PDT. For example, a PDT designed to operate in the healthcare domain will mirror physiological aspects of the person with a great range of details (weight, height, sex, heartbeats, respiratory rate, metabolism, …) and include a variety of health related data (EHR, genome sequence,

---

[23] https://digital-strategy.ec.europa.eu/en/policies/non-personal-data
[24] https://ec.europa.eu/eurostat/data/database
[25] https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
[26] https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-digital-identity_en
[27] https://sciencebusiness.net/news/eu-create-new-biomedical-research-agency-modelled-barda
[28] https://gdpr-info.eu/
[29] https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en

"omics" info, occupational hazard, living environment, parents pathologies, etc.). On the contrary, a PDT used to mimic the competence of a person won't include health-related data (although it might include health constraints, like allergies). This type of PDT focuses on the skill and knowledge of the person, with data representing the previous experiences. It might also contain information on work-related behaviors (inter-personal communication skills, leadership, stress resistance, problem solving capabilities, etc.).

Any person, of course, is all of the above (and much more, just think about the social relations that are an integral part of our life, and the impact they have on our behaviour …). Hence we can have a PDT that is the aggregate of several PDTs of that person. In effect, this aggregate of PDTs would mimic the overall person much better than any of its components alone. In perspective, we can assume that the many PDTs that will be created over time to represent specific aspects of a person will merge into the aggregated PDT.

As a matter of fact, looking into the future (next decade and beyond), we can assume that each person will have a digital counterpart, a personal digital twin, that will be "born" during the gestation period (accumulating data from exams taken by the mother), and will keep growing through the subjects lifetime. This PDT will be the full digital persona, and it can be used to share data at a specified time and to pre-authorized parties that have permission (and need) to access, under the control of the physical person.

It is clear that a cluster of PDTs can be used to represent some general characteristics of those PDTs, and that cluster can be modelled through a Digital Twin. For example, we can look at the PDTs of a community—of sport fans at a game, employees in a company, or blue collars on a factory assembly line.

This Digital Twin can be used for simulation purposes (to study the behavior of a group of people in a given situation). It will mimic the behavior of a crowd, but should not be considered a PDT.
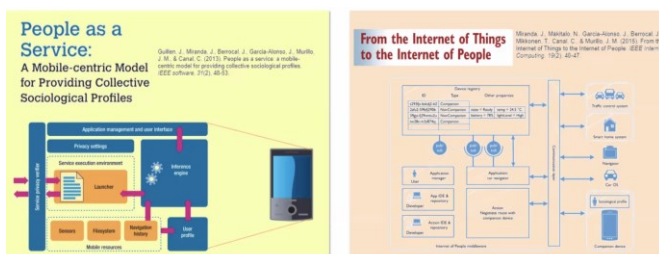

Figure 15. A PDT can be used to represent the social behaviour of a person and several PDTs in a community (an ambient) may be used to simulate and study the global behavior of a group of people. If these PDTs are shadowing their physical entities, we could say that the physical persons act as IoT and their PDTs become an Internet of people. Image credit: From the IoT to IoP, IEEE Internet Computing 19(2) and 31(2).

As shown in Figure 15, PDTs may be shadowing the status and behavior of their associated physical entity (the person). In this case, the physical entity is—functionally and for the most part—modelled by the PDT, or the equivalent of an IoT and PDTs (becoming the Internet of People, or IoP).

There are many applications foreseen by the group of experts. The consensus is that a European action can foster the uptake of PDTs and their use as IoP in the context of healthcare (including monitoring/controlling epidemics), and this area can be the one that paves the way to application in smart cities, eGovernment, etc.
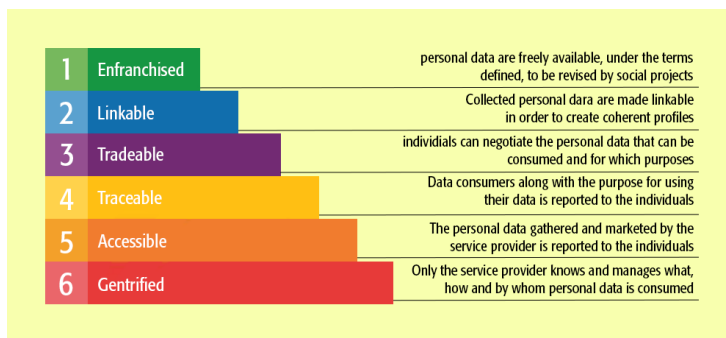
6.3 Data Sharing

Figure 16. A possible classification for the level of sharing of personal data. Image credit: Juan Luis Herrera et al, University of Extremadura, Hsiao-Yuan Chen et al, University of Texas at Austin, Niko Mäkitalo et al, University of Helsinki

The PDT is a model of a person. Part of the model consists of data, and part consists of the "interpretation" of that data to manage interactions that mirror the behavior of the person. Over time, the data sets will become broader, covering many aspects of the person, and will include the "history" of the person (like the EHR when modelling health aspects, or the evolution of the knowledge space when modelling education and working experience).

These data sets are clearly crucial for the operation of the PDT (a PDT without data does not exist). In addition, they may be useful to third parties, including other PDTs interacting with them. Sharing this data can be beneficial to the person, and in some cases, sharing can result in revenues. On the other hand, privacy and ownership issues may require restrictions and limited sharing.

The EU has a number of directives and regulations in place pertaining to the use of data—from the protection of data (GDPR) to the EU Data ACT[30] aimed at making more data available for societal and business benefits (in a way, you can see the GDPR and the Data ACT as looking at opposite needs to protect/hide on one side, and to leverage/share data on the other).

The experts' group on PDTs discussed the need for a specific regulation that could provide a framework steering the development and interoperability of PDTs among themselves, and with third parties applications.

The classification shown in Figure 16 was proposed:

- Gentrification: no data sharing at all. The owner of the data is managing them and uses them only for the reason they have been collected for in the first place. The owner of the PDT (the person) will, as a matter of fact, delegate the management of data to a service provider with the agreement that the data cannot be used for anything else other than the personal benefit of the owner (this includes their use to provide the required services),

- Accessible: the data manager (service provider) must be transparent on where the data is harvested, and the person referred to in the data must be provided instructions as to how to review their data. For example, my PDT can be "enriched" by data, resulting from the tracking of my activity on the web (search, e-commerce, entertainment), and I should be made aware of the existence of this data.

- Traceable: (some of my) data is shared with third parties for specific reasons, and I am notified both of what data is shared, for what reason, when, and to whom.

---

[30] EU Data ACT: https://digital-strategy.ec.europa.eu/en/policies/data-act

- Tradable: I, as owner, have the pootential to negotiate what data can be shared, and at what "price."

- Linkable: data can be analyzed in context and linked together to generate further data (meta-data), and results from applying data analytics to various sets/streams of data. The resulting metadata provides a higher level of information about myself (for example, linking the data of my physiology—fever, cough, heartbeat, shallow rapid breathing— with the ones of my whereabouts, and highlights potential exposure to contagious people, and can lead to an assessment of myself as high-risk of contracting Covid).

- Enfranchized: my data (or most likely a part of it) is freely available for societal benefit (for example, I might decide to provide free access to my genome sequence to researchers).

Presented above is just one possible way of structuring data sharing. It is obvious that some data will be shared more freely, and other data will be kept private depending on its meaning and the value to the person. However, a European (more generally, an institutional) framework may impose a certain degree of sharing (access) to specific data by specific parties. In any case, the person who is the owner of those data (in the sense that those data are about them) needs to be notified of the sharing/access and intended use.

The issue is not easy at all. For example, my image is picked up by security cameras many times a day. it is also picked up serendipitously by other people who are taking a picture of a monument when I happen to be in the frame… How do we protect this data? How can we be notified that this data "exists"? Notice that the issue is becoming more and more important as image recognition (face recognition) software becomes a commodity (you can search on Google using an image, and it is more and more likely that you will get results showing that same face in other photos. From there it might be quite easy to identify a person…).

If I appear in a photo and that photo becomes part of the data of another person PDT, what kind of right do I have on that? It might be close to impossible to be aware of my presence (presence of my data) in other PDTs. Of course this is a much more general problem, the emergence of PDTs and embedded AI will just make it more sensitive and impactful.
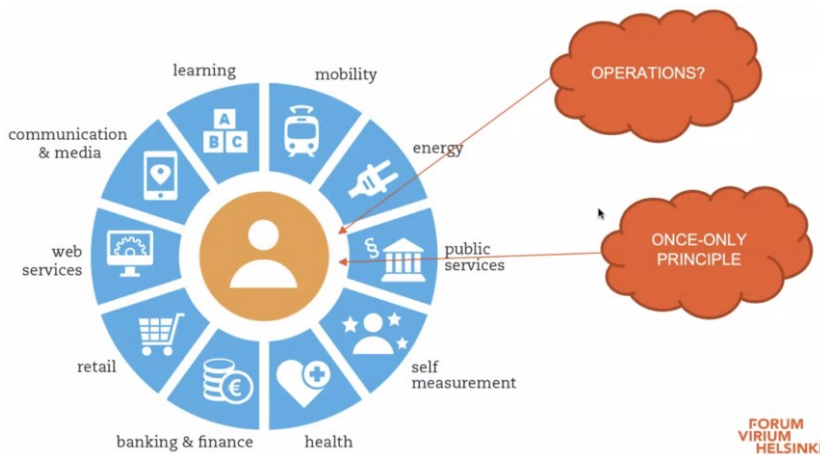
6.4 What Data?

Figure 17. The landscape of data that can, potentially, be part of Personal Digital Twins. Notice that only a small portion of them are data generated AND captured by the person (self-measurement). Most data results from interaction of the person with their environment, and is likely to be captured by third parties. Image credit: Forum Virus Helsinki

Having discussed a possible classification of data with respect to their sharing (access, usage, exploitation rights), the question now is on what data can/should be part of a PDT.

This discussion was steered by a presentation of a data "landscape" as shown in Figure 17.

This representation is interesting because it clearly shows that most data is captured through interactions that a person has with the environment, with third parties, and what is most likely to be captured by those parties. Only a fraction of personal data is captured by the person (self-measurement data).

As a matter of fact, most of our interactions with third parties today generate a digital trail that some of the parties involved will store in their client records. For example, shopping at a supermarket and online, engaging in banking transactions, watching a movie on Netflix, booking a hotel, buying a train/flight ticket, accessing a public service, seeing our doctor, etc. The resulting digital trail is not necessarily a personal digital twin—actually, and by far, it is not. It can be used by the third party as part of their accounting procedures, and in a few (growing) cases, by profiling our person as their client. More sophisticated applications are becoming available to make sense of those records, some aiming at developing an executable profile of our person that can be used to predict future demands and/or stimulate/steer those demands. This growing sophistication in the analysis and management of personal data records brings us closer to the concept of a digital twin. It would be a "personal" digital twin in the sense that it reflects our person (habits, history, future behavior), but it is not "personal" in the sense that it is mine—I have basically no control over that pseudo-digital twin.

There are signs of evolution indicating that some companies may be willing to share data with us so that we can use them (of course it must be mutually beneficial). For example, some stores/brands are starting to empower their fidelity card through apps that can reside on our smartphone. These apps will accrue the digital record of our transactions (capturing both what we bought and what we "saw"), and will provide ways to contextualize this data to our context (the here and now). For example, we can use a 3D model of our body, created by a 3D scanner in an apparel store, to try on (virtual) clothes presented on the website of a store. We can also see how a given sweater matches with a skirt purchased four months prior. Augmented reality (AR) and virtual reality (VR) are important tools to transform digital models into concrete perceptions.

In Healthcare, we are just a few steps away from seeing personal digital twins being created by medical insurance companies that are constantly updated based on visits to our doctor(s), exams, and through data generated by wearable and ambient sensors. Applications interacting with the PDT, or embedded in the PDT, can analyze the data and issue warnings, red flags, and even generate signals to intelligent agents that can make in depth analyses and interact with us (chatbots). These interactions could also become part of our PDT record.

These are just examples. We can see a general trend where profiling evolves into the creation of a PDT, and further, in making this PDT available to the person, in the form of an app, that can be downloaded on that persons smartphone.

This PDT will be, by far, controlled by the provider (with its vested interest), but it is nevertheless important because it will:

· Create a culture of using the smartphone as an extension of the person;
· Make services available through the PDT;
· Stimulate other companies to offer similar services via a PDT.

On one hand, this evolution is steering towards the co-existence of several (many) PDTs, and this may be confusing, but on the other hand, it can prompt companies to provide an integrated PDT as a cluster of existing, sectorial, PDTs of that person.

6.5 Synthetic Data

The data that is associated to a personal digital twin is data derived from that person, as explained in the previous sections. Part of this data, however, could have been created by software algorithms that used existing data derived from that person to generate additional ones—for example, through simulation scenarios.

Via software, it is possible to create data that, in a way, mimics certain characteristics of a person, but that at the same time is not "connected" to that person specifically. In this case, their data is used as a blueprint to create a generic model that can be applied in other contexts. Hence we have:

· Data modelling the person (e.g. the model of the circulatory system);
· Data shadowing the person, or representing the current status (e.g the current heartbeat and the current context while walking up a flight of stairs);
· Metadata derived from analyzing several streams of data (e.g. a fitness grade through data analytics considering the persons age, physical model, current physiological data, environmental data);
· Synthetic data representing an abstract model inspired by that persons data, along with the data of several other people (how a person having a certain degree of fitness would react in a specific circumstance).

Synthetic data is becoming more and more important in training artificial intelligence software.

Figure 18. All the people shown in this image have one thing in common: they don't exist. The "photos" have been created by a software that took inspiration from photos of real people to create artificial ones. This is synthetic data. Image credit: NVIDIA

At the same time, synthetic data can be used to create models of entities, including people. Figure 18 provides an example—this photo collage was not derived by placing photos of real people on a single canvas. Each face was created through AI trained on recognizing people's faces, which led to the creation of a synthetic model that has been used to create faces. By creating the faces based on human face models, the results cause us to believe that those are faces of real people.

Similarly, we can now train an AI software tB to learn the style of a given composer (Bach[31], Mozart, the Beatles…), and have the software create a musical composition that would be in line with their style. Such high accuracy has been achieved using this tool that the composed musical pieces can fool most people, including some musical experts.

The importance of data modelling and data shadowing for a PDT is obvious, and so is the relevance of metadata. These are the ones that usually provide a "meaning" and, therefore, are the ones used by the physical person in the interaction with the PDT (you are not interested in knowing your heartbeat, rather in knowing if your heartbeat is ok, and that is information that has to be derived through the analysis of different data, personal as well as ambient—like how steep the trail on which you are currently jogging is…).

What is the importance of synthetic data? Well, synthetic data might capture some of your characteristics in a generalized way that decouples your identity from the characteristics. This synthetic data might be used by a Digital Twin that can act as an avatar that can be studied. For example, the behavior of a person with characteristics similar to yours as part of a crowd in a given situation. This decoupling preserves the privacy of your data, and your "persona" (of your behavior).
.
Notice that if you were to use your digital persona to provide a service, like a consultancy, you would not use synthetic data—rather, a subset of your data and metadata since, in this case, you want to maintain the relation with yourself.

Synthetic data is likely to become widely used throughout enterprises to train artificial intelligence in products. For example, when providing a smart product that can interact with a variety of users—detecting each unique user within a certain range.

Marketeers are also going to leverage synthetic data to assess the characteristics of a given market in terms of reaction to specific stimuli.

---

[31] AI generated music mirroring Bach style: https://www.youtube.com/watch?v=lv9W7grYhbk

Government and institutions, such as those managing smart cities, are very interested in creating synthetic populations since they are a powerful tool that can provide accurate information without compromising the privacy of citizens, while simultaneously reflecting the complexity of the structure of a population and the characteristics of the individuals that influence their behavioral response[32].

A personal digital twin may, among its numerous capabilities, have the potential to create a digital twin to represent a virtual person that mimics characteristics of its physical entity.

6.6 Web of Digital Twins

The discussion in the group of experts highlighted the need for a general framework to support the access from one digital twin to another—For example, how we can access a document on the web from another document (hyperlink).
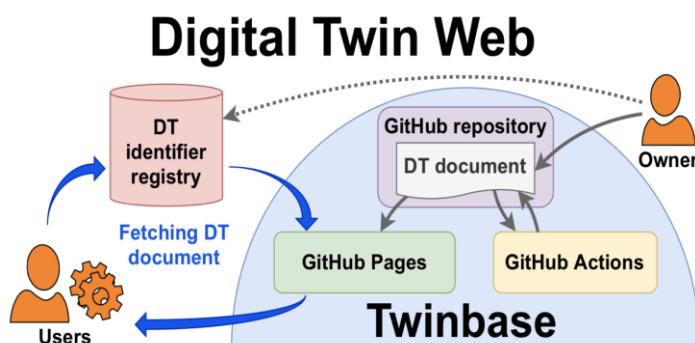

Figure 19. As digital twins become ubiquitous, they will also represent the interactions taking place among physical entities by mirroring these interactions in the cyberspace. It would be very useful to create a mechanism, similar to the world wide web, to link all these digital twins. Image credit: Juuso Autiosalo

As a matter of fact, it is easy to predict that digital twins will increasingly be created to represent physical entities to the point that, by the end of this decade, every manufactured product will have a corresponding digital twin. This will also apply to entities that are not manufactured but that exist in the physical world (including ourselves). A growing number of people will have a personal digital twin, and as we have discussed in the previous sections, we might end up with several PDTs— each one representing an "aspect" of the person, and, each one likely to have been created by a different player (the healthcare system, the company I am working for, Amazon, and the like … myself…).

The variety of digital twins, and PDTs, would surely benefit from a framework that can support their identification and the establishment of a connection amongst one another. This is what happens with the web in which there are differnt standards for identifying various entities (such as documents), and a protocol to connect one entity to another.

The group agreed that one of the goals of the EU should be to promote the creation of a Web of Digital Twins (and PDTs). This web of Digital Twins can serve as a place interchanging data in the digital space, as well as a place to access and retrieve data (or related services).

The owner will set the access/sharing rules, which users will be required to accept and comply with.

Interestingly, a sub-web (subset of a given DT web) could be seen as a digital twin with complex behavior and specific (dynamic) knowledge. In other words, for example, imagine several of our PDTs clustered into a web of which I can claim ownership and set the sharing rules.

---

[32] Modelling of a population using synthetic data: https://www.matsim.org/gallery/paris

MyData roles

From: https://mydata.org/declaration
[Blue content added by Juuso]

Data sources

activity, temperature, heart rate

location, acceleration, digital profiling
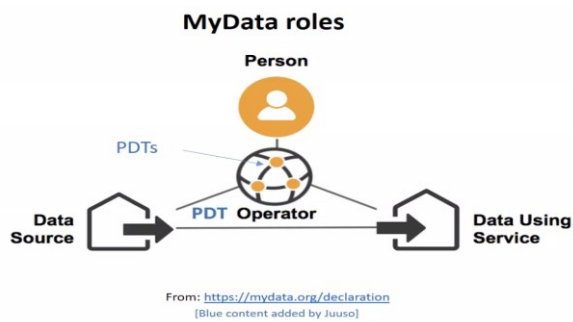
Data using services

APP

Figure 20. There are multiple personal data sources that can feed data to a PDT. Likewise, there can be several (interested) users for this data. Rather than having the data directly accessed by the users, the PDT can intermediate, thus maintaining control of sharing personal data. Hence the PDT becomes a crucial element in the data governance. Image credit: Juuso Autiosalo

At the same time, the PDT formed by clustering my elemental PDT can create an emerging representation of myself. The same, of course, will apply to other clusters similar to those of people, such as clusters of a city …

These clusters represent the overall knowledge and behavior of the component DTs once they inhabit that cluster. This goes both ways. For example a smart city DT cluster can:

• Represent the collective behaviors of people in a city as conditions change, for example, when the city has insufficient resources, or as new services become available, in the event of an epidemic…

• Be used by a single PDT to evaluate how a persons behavior changes when present in a given cluster.

A key aspect for fostering this web is the governance.

6.7 Personal Digital Twins Governance

One of the basic components of Digital Transformation is that data, from an economic standpoint, are nonrival[33].

If you eat a loaf of bread, that loaf is no longer available to anybody else. There is competition on that resource, and governance policies must establish how resources shall be allocated (given to a specific person, split in half and each half given to a person, set aside for future consumption, ...).

On the contrary, data can be used simultaneously by any number of users at different times of the day. The use of data is not limited if in use by another user. As a result, allocating data resources as an attempt to govern data will fail since these data and resources, in practice, are infinite and can be leveraged by any number of users. The European Commission has already issued (and is working on issuing) several documents that set the stage for data policies—from the basic General Data Protection Regulation (GDPR[34]), specifically looking at Personal Data, to the Data Governance Act (DGA[35]), to the definition of European Data Spaces (Gaia-X[36]), and the European Digital Identity Wallets (eIDAS 2.0[37]).

The EU is now starting to work on the governance of Personal Digital Twins as a way to protect individual (privacy), and society as a whole, in accordance with western values of societal wellbeing. Governance of PDTs requires the establishment of acceptable frameworks for their creation and exploitation, thus involving PDT operators (these should be/become trusted parties, such as your

---

[33] Data as nontrivial entities: https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.20191330

[34] GDPR: https://gdpr-info.eu

[35] DGA: https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-data-governance-act

[36] Gaia-X: https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html

[37] eIDAS 2.0: https://digital-strategy.ec.europa.eu/en/policies/eidas-regulation

doctor) and the sources feeding data to the PDT, the services provided by the PDT, and those that are using PDTs. We should pay extra attention to the ways people may be using their PDTs (and who is responsible for what).

One of the points under discussion is how such a governance shall take place. Should it be an indirect governance in the form of laws and standards, or shall it be a direct governance through the provision of services and technologies (in other words, should the PDT be similar to a passport that is "issued" by the EU?).

These are not easy questions, and, in part, the answers will depend on what is happening (or will happen) outside of Europe. Will there be international standards for PDTs? Will companies provide services based on a PDT and/or will PDTs actually result from services people are provided? For example, if you are subscribing to a service resulting in the creation of your PDT, will you be able to use and control how that PDT integrates with others and how it is tied to you?

I do not have a crystal ball, but if I had to make a prediction, I think that the evolution will be bottom up rather than top down. It is unlikely that an institution (like the EU) will dictate and steer the evolution of PDTs. It is much more likely that several companies will offer services that result in the creation/adoption of PDTs of people. Other companies will provide tools to integrate fragmented PDTs into a single entity that can be compliant with a general framework (this one could well be the result of work at an international standardization level) that can have some characteristics regulated by a body, such as the EU.
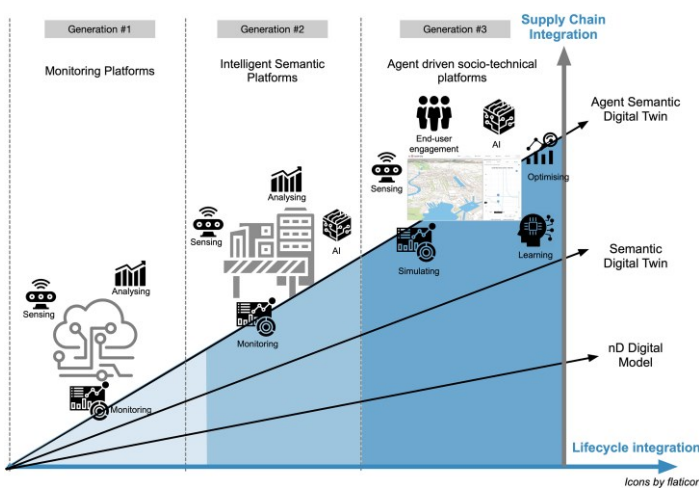
7. Conclusions

Figure 21. A representation of Digital Twin evolution towards semantic Digital Twins. The lower part of the graphic demonstrates the evolution of a multidimensional DT (representing different characteristics of its related physical entity), the middle line represents the evolution of a semantic DT, and the upper line the evolution towards a Semantic DT acting as an Agent for its physical entity. Image credit: Calin Boje et al, Towards a semantic Construction Digital Twin: Directions for future research, Elsevier

Wrapping up this long discussion on Personal Digital Twins, it might be appropriate to remember that PDTs are "Digital Twins" and … Digital Twins are evolving!

As shown in figure 21, DTs have evolved from a monitoring platform (mirroring the status of their physical entity in quasi real-time) to an intelligent platform with a growing understanding and awareness of the environment. Continuing to evolve towards becoming agent driven, context aware, and part of a socio-technical platform.

It is also worth noting that the graphic shows three evolution paths corresponding to different types of internal capability. On the lower line, the evolution of a digital twin can be said to mirror the syntactic aspect of a physical entity, like its temperature, acceleration, speed, location, etc. The "nD" associated to the Digital Model is to highlight that there may be n dimensions. In other words, n different parameters captured to mirror the physical entity (in the given example, the temperature, acceleration, speed, and location—hence 4 independent characteristics). The middle line, the Semantic Digital Twin, includes the evolution of a Digital Twin mirroring semantic characteristics such as overheating, normal attitude (speed/pitch), route position, etc. Notice that the semantics could be associated by

software related to the data present in the previous type of digital twin. However, in that case, the DT does not "know" the meaning of the data received since that knowledge is external to the Digital Twin.

The upper line represents the evolution of an Agent semantic digital twin. This DT knows the meaning of the data received and can contextualize this meaning with respect to a goal. In other words, it understands the meaning in a given context (as a result, it can take appropriate actions when needed). For example, it can detect if a vehicle is overheating due to current acceleration while determing if the increased heat is being caused by a part malfunction or from the vehicle taking action to aoid an obstacle.

Notice that this mapping of evolution is not in contrast with the now established stages of DT evolution (from pure mirroring to autonomy), it just provides additional perspectives.

Personal Digital Twins can be applied in several ways:

· Mirroring the physiological data of a person (pressure, heartbeat, metabolic rate …)
· Knowing the expected mix of physiological data and be able to raise flags when these differ from the baseline;
· To understand the physiological data in a specific context (for example, a racing heart beat and shallow breathing during a game of cards—emotion driven vs. Resulting from strenuous exercises that should be accompanied with motion indicators and deep breathing), and initiate recovery actions when needed (none would be needed in the card game example, however, a rapid heart beat and shallow breathing during a strenuous exercise may raise a red flag).

Personal Digital Twins are bound to evolve first as shown in the lower line, likely becoming better and better mirrors of the person by increasing the number of dimensions (nD), and will be used by external applications to derive meaning (cloud services in healthcare will monitor the PDT and provide feedback, both directly to the physical entity—the person—and via the PDT). However, I consider the g5rowth of the PDT to inevitably become autonomous in assessing the meaning in the data (For example, the PDT in my smartphone will go beyond harvesting my physiological data—it will soon have the capability to analyze the data and provide feedback). Then from this, become capable of understanding the data in any given situation (and/or with respect to a pre-assigned goal). This is just a small step.

Artificial intelligence is crucial technology and, over time, will become partly embedded—countinously growing in capabilities. Then, it will reach a point in which it becomes part of a global intelligence, thus becoming aware of the context and able to determine, at any given time, what information should be shared, and what must be harvested from existing global knowledge.

Regulating PDTs will most likely boil down to regulating AI, something I am skeptical about in terms of practical feasibility. I am more confident that we can create trusted frameworks for technology to provide accountability and enforce compliance.