observatoryNano:

European observatory for science-based and economic expert analysis of nanotechnologies

Work package 4: Ethical and societal impacts

TOOLKIT FOR ETHICAL REFLECTION AND COMMUNICATION

(DELIVERABLES D4.4.1 AND D4.4.2)

CEA-LARSIM

Marc Pavlopoulos,

Alexei Grinbaum*,

Vincent Bontems

June 2010

* Address: CEA-Saclay/SPEC/LARSIM, 91191 Gif-sur-Yvette Cedex, France. Email: alexei.grinbaum@cea.fr

Copyright notice

Foreword

The ethical debate on nanotechnology is large and tangled. It is often unclear what the right questions in this debate are, nor whether these questions are specific to nanotechnology in comparison with other emerging technologies. This Toolkit for ethical reflection and communication does not claim to provide a definitive picture of all options in the ethical debate on nanotechnology. Its aim is more modest: we wish to provide the reader with means to frame her own vision of the debate and to sharpen ethical awareness of the parties involved in the development of nanosciences and nanotechnologies. We hope that this will foster the dialogue between philosophy, science, industry, and society. The toolkit does not replace academic research on the subject[†]. It is our firm conviction that those who think about nanosciences and nanotechnologies are better equipped to do so with a notion of philosophical ethics. This is because the views elaborated over centuries can enable the construction of an argument to respond to new problems and because philosophical reflection itself will suggest new lines of questioning. The approach behind this Toolkit rests on two assumptions:

- That nanotechnology as a new wave of cutting edge technology is likely to lead us, over time, to deeply modify both our norms for action and our worldview. This was the case with nuclear technology, space travel, information technology and genetic engineering, and it will likely be the case with nanotechnology. The role of philosophy is to study this change and to accompany it with a continual reflection.
- Certain fundamental choices must be made now, which concern the future. Yet uncertainty prevails with regard to our own future, for reasons connected with unforeseeable technological development as well as the arbitrariness of human decision and policy-making. Rather than guessing at what the future might be, it is crucially important to install a practice of ongoing normative reflection about the problems that appear on the horizon.

From the methodological point of view, this Toolkit must not be characterized as a set of rules, moral or otherwise, imposed on nanotechnology. There can be no ready-to-use ethical code for any new technology. Principles, norms and values all evolve in time. Rather, the approach should be seen as being in the form of what the American philosopher John Rawls calls a "reflective equilibrium". This consists of a double-way exchange between ethical theory and practice in a process of deliberative mutual adjustment. The Toolkit must not be seen as a recipe book offering plug-and-play solutions. It helps, and perhaps sometimes guides, ethical thinking by offering concepts, notions, and methods for an application to practical cases.

We welcome comments from all readers.

[†] E.g., see the journal *Nanoethics* published by Springer.

TABLE OF CONTENTS

Part 1. Introduction

1.1. What is ethics?

Since Antiquity, ethics has been dealing with the question of good life: what is it to live well? The conceptions of good life vary across societies, historical times and individuals. Philosophical ethics does not seek to establish absolute rightness of a particular view but its robustness. For any given answer, the ethicist asks: What arguments count in its favour? Is it consistent? What objections does it face? How can one reply to these objections? What are the weaknesses and strengths of this view?

The job of the ethicist is not to issue ready-made prescriptions. He won't give you answers to the eternal questions of mankind. The ethicist can only help you to structure, nurture, strengthen and deepen your own position.

1.2. What is special about nanotechnology?

Answer 1: The scale.
Chemistry has been working at 10^{-9} meter for centuries. Biology studies processes at this scale too.

Answer 2: Nanotechnology is not just about working at the nanoscale, but about designing new devices, objects and methods. It employs instruments that reach out directly to nanoscale objects.
True, this is really what is new on the technical side of things. But this just means that nanotechnology is an engineering technology working at the nanoscale! Every new technology opens a new field and employs its unique instruments. What we want to know is the *sense* to be made of this new development in engineering.

Answer 3: The nano revolution will radically modify our ways of life, opening up a posthuman future.
We would already look posthuman to a Neolithic man or even an Ancient Greek. Technology has dramatically transformed humans and the whole world. That nanotechnology will do so too is no surprise. That the change it will bring about will be more radical that what humanity has already seen is mere speculation.

Nanosciences do not carry with them any radically new scientific theory. They deepen and widen the knowledge of the nanoscale, but this is not a scientific revolution.

Nanotechnology is a label for a vast and heterogeneous array of new ways to engineer matter. It is not an exception nor even a rupture in the long history of science and technology.

So far the nanotechnology revolution has been more of a social one: nanotech has brought about important transformations in the way society thinks about science, questions the very notion of technological progress, changes its values and develops imagination. Economy and industry are transformed by nanotechnology as they have been transformed by many other generations of technology. Nanotechnology's aura of unheard-of novelty does not arise from nanoscience, but from the impact it makes on society.

1.3. What can ethics do for nanotechnology?

Nanoscience and nanotechnology transform lives as they influence some aspects of how people get cured, die, give birth, eat, communicate, etc. They also transform the society: the circulation of information, mobility, energy resources and so forth. Society develops new ways to use technology. In the 1990s, when computers and internet already existed, who would have guessed that blogs or social networks would flourish globally and change our lives?

Sometimes we say that technology has made our lives better, but sometimes we aren't so enthusiastic. Technological changes can cause an abrupt change in social history but can also go unnoticed. It is difficult to predict with certainty what a given technology will do. So what should our attitude be with regard to this uncertainty? What shall we do? How to make a judgment? This is where ethics enters.

Scientists no longer live in an Ivory Tower. Society wants its share of influence on what comes out of the laboratory, because it knows that the uses of technology will have consequences for its own well-being. Whether the nanotechnologist wants it or not, he is not alone: his fellow citizens are interested in his work and want to control future applications. The responsibility of the scientist *de facto* exceeds his own sphere of action. How to deal with this situation? How to make sense of the scientist's new role? This is where ethics enters.

Dreams are not innocent. Fiction and imagination are immaterial, but they help to tell stories and to make sense of actually existing material objects. What today's teenagers will do tomorrow as grown-up scientists and engineers depends on the books they read and films they see. But fiction usually parts radically with the realities of the laboratory. The nanotechnologist is often appalled at how unscientific stories attract more attention than his own hard work. What attitude should the nanotechnologist have towards fiction? Dismiss it or embrace it? This is where ethics enters.

Part 2. Classifying ethical and societal issues

Any classification of ethical and societal issues concerning nanotechnology is conventional, however having such a classification helps to structure both one's own thinking and the public debate. Numerous issues cannot be strictly assigned to one category only. We propose a conceptual map.

Detailed material will be found in separate boxes, each time linked to a specific group of questions. One can skip these boxes on a quick reading, but they are instrumental for further structuring one's thinking on ethical questions.

2.1 Nanobiotechnology

- Does the heuristics of “mimicking nature”, common in top-down as well as bottom-up nanosciences, pose in itself an ethical question? In other words, is this imitation of nature itself a moral problem?
- Does nanobiotechnology blur, or lead to the transgression of, the border between nature and artifice, human body and machine?
- Is there a clear rational distinction between the living and the inanimate?
- Biological molecules such as the DNA are associated with the “sacred essence of life”. When they are put to use in nanomechanical devices or as templates for the arrangement of other molecules, does their moral status change?

Living and Inanimate

It has long been thought that living beings were made of a special vital matter that would set them apart from inanimate things: this theory was vitalism. The discovery of the DNA by Watson and Crick in 1953 and the subsequent rise of molecular biology gave vitalism its final blow: the DNA, the “essence” of life, is built from an “alphabet” of molecules.

Since then, many biologists impressed by the achievements of cybernetics and information theory have adopted the view that the DNA is an algorithm that controls the development of a cell. Life as a whole can be reduced to the DNA, much like all we need to know about the computer is contained in its processing instructions. But other biologists insist that various epigenetic and environmental factors should also be taken into account for explaining living phenomena. And the informational point of view still leaves open the question of what is life.

Two main concepts associated with life are (1) self-reproduction, (2) an equilibrium of the living object’s interaction with its environment across a membrane that separates them. The notion of life as an algorithm encompasses many factors that do not satisfy these two conditions: computer software and inanimate objects such as crystals also behave according to a “program”. This either calls for a deep revision of what we consider as living beings or pushes one to admit that biology alone is not sufficient to define life.

In both cases, **a scientific notion of life will not be identical with the common-sense notion of life,** which lies at the foundation of our moral judgment. Hence we have no scientific basis for our moral judgement about life, such as what exactly must be respected or protected. We don’t even have a clear idea at which level life begins: at the level of the DNA or, say, proteins or perhaps at the level of the entire cell? For instance, what about viruses with a metabolism but no self-reproduction? Biology traditionally considers cells to be the simplest systems where the two conditions above can be satisfied. But again, there is no scientific proof that these conditions cannot hold for some other sort of systems, perhaps having nothing to do with cells.

Natural vs. Artificial

“Nature” is a tricky word. Here are its three main meanings:

1) **Scientific notion of nature.** The point of view of natural science since Galileo and Newton has been that nature is a set of universal laws. Everything in the physical universe is natural, even man-created artefacts, for they obey natural laws. This notion of nature inspired by science leaves no room for the opposition of natural and artificial. So the natural *vs.* artificial is based on a different notion of nature.

2) **Ancient opposition between nature and technique.** When people talk of what is natural and what is artificial, they make an implicit reference to a philosophical concept of nature which, historically, preceded the advent of natural science. According to Aristotle, a natural object contains *in itself* the cause of its movement, whereas artefacts or technical objects have their behaviour determined from the outside, i.e. through the action of their creator or craftsman. This definition carries a basis for normative judgment: a behaviour springing from the object’s inner constitution is authentic and free, whereas behaviour imposed on it from the outside, e.g. through engineering, is constrained. Hence the modern idea that what is natural is more authentic, more autonomous and free, in short *better* than the artificial.

3) **Nature and culture**: This distinction is a modern elaboration on the Ancient opposition #2. Here “nature” means a global order of all things on Earth, including Man. Man holds substantial relationships with other things and beings, so when he transforms his environment, he induces a substantial change in himself. Beginning with Romanticism in the early 19th century, it has become commonplace to believe that this indirect transformation of Man by himself is a denaturation, for it perverts the “essence” or nature of Man in the Ancient sense. In other words, when Man transforms his environment, he loses his own true nature.

Does Man have a fixed, unchangeable nature? According to a long philosophical tradition going from Rousseau to Marx and Habermas, human nature consists in having no fixed nature: Man indefinitely perfects himself through arts, sciences and technology. These lead to a change in the conditions of human life, which in turn influence Man himself. Thus, what the Romantics described as a denaturation of Man is in fact an essential characteristic of his nature. Indeed, Man alone among animals has a history, which is nothing other than a continuous transformation of the human condition by Man himself.

2.2 Nanomedicine

- Think of infra-red vision for the blind or a memory prosthesis implanted in the brain. Is it possible to draw a clear boundary between a therapeutic and an enhancing technique?
- Can regulation suffice to prevent wild use of nanomedical methods and products for enhancing bodily and cognitive capacities?
- What becomes of informed consent when diagnosis points to a probability of developing a disease within several years? Given that the way doctors speak to their patients influences the choices made by the patients, how should they present such a situation?
- If permanently implanted biosensors monitor health and transmit their output to ICT devices for analysis, how can privacy of medical information be secured? Is it feasible to prevent insurance companies from accessing these data?

Transhumanism

The term *transhumanism* was first used in 1927 by Julian Huxley, brother of the author of *Brave New World*. As an intellectual movement it originated in California in the 1980s. In 1998, Nick Bostrom founded the *Word Transhumanist Association*, which has since changed its name to *Humanity+*. It is an influential international association holding chapters in several European countries including France, Germany, the Netherlands, Italy, and the UK. For *Humanity+*, "Humanity stands to be profoundly affected by science and technology in the future. **[Transhumanism] envisions the possibility of broadening human potential by overcoming aging, cognitive shortcomings, involuntary suffering, and our confinement to planet Earth**". Transhumanists also target "involuntary" death: they support technologies such as cryonics (freezing the body after death in hope of a future "resurrection" when technology will permit) and fictions such as uploading an individual human mind on a computer. For the human species as a whole, transhumanists support **human enhancement** to progressively generate a new species.

Transhumanism has roots in the humanism of the Enlightenment. Like Humanists, Transhumanists hold that scientific progress goes hand in hand with moral progress and that Man's distinctive nature consists in constantly improving himself through knowledge, science and technology. **Unlike Humanists, though, Transhumanists believe that Homo Sapiens has a fundamentally limited nature** and that present science and technology will progressively give rise to a superior species. Some Transhumanists even think that Homo Sapiens is essentially defective both cognitively and morally. For these **Posthumanists, the evolution of Man towards a superior being will be disruptive** and Homo Sapiens will eventually be replaced by a fundamentally different being integrating technology and biology.

Tranhumanism was strongly opposed by the bioconservative movement in the US. Their advocacy in favour of the NBIC convergence was influential in the early years of the US National Nanotechnology Initiative.

- Lab-on-a-chip and biosensors make self-diagnosis likely in the near future. How should patients learn about the probability of developing severe conditions such as cancer? Should they get first-hand access to the results of the tests themselves, should the treatment decision be delegated to a computer or should the results only be disclosed in conversation with a doctor?
- Evolution of the notion of disease: is the identification of one molecule or harmful agent in the patient's body enough to characterize him as ill? What does the notion of disease become when a patient can be "cured" before she even develops symptoms?
- On what basis should the probability to develop a disease be calculated? By reference to a country's population? To a targeted population? To several subclasses of population?
- Could personalized medicine lead to an individualization of norms of health and disease, each individual deciding for himself on what counts as a healthy state and what level of actual or probable malfunctioning justifies treatment?
- As a rule, public health insurance only covers diagnosis and cure of a *disease*. How will public health system consider treatment decided after early diagnosis, when there is only a probability of developing a disease?
- How will public health insurances consider regenerative medicine, given that it may blur the border between curative and enhancing treatments?

Health and Disease

Traditionally, health is defined as absence of disease. While this definition is at the basis of many European public health systems, **the World Healh Organization has a more ambitious approach: "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity".** Since different societies and even different individuals have different ideas on what qualifies as well-being or as disease, the WHO definition opens the way to dramatically different interpretations.

In medicine proper, there are three different conceptions of health and disease:

(1) Disease can be identified with the presence of a **pathogen**, such as a virus, bacterium, or tumour. The main problem here is that the presence of a pathogen is neither a necessary nor a sufficient condition for disease. Influenza virus, for example, need not always cause flu, and some diseases have no identified cause.

(2) Disease can be defined physiologically as **malfunctioning of one or several organs.** Though prevalent in medicine, this functional view presupposes a definition of "normal functioning", while in fact this concept varies a lot from one culture to another. Hypercholosterolemia is considered disease in most European countries, but not in Northern Africa; and until the 19th century mental disorders were not considered diseases. It is very difficult to find a purely biological criterion for "normal functioning" of the human body.

(3) Health may be positively viewed as a **global state of good functioning, i.e., a condition enabling individuals to reach their "vital aims".** This view is clearly a normative one and close to the WHO definition. It endorses variability and possible individualization of the norms of good health. It may prove to be the most nanomedicine-friendly definition of health, but for the time being it is only marginal within the medical community and completely at odds with the practice of public health systems.

2.3 Food and cosmetics nanotechnology

In the Western society, certain spheres of human life can be called "sacred": these concern all kinds of intervention on the human body through nutrition (food), cosmetics, or medicine. Their "sacredness", connected with the inviolable status of bodily integrity, implies that every individual in particular and the society as a whole tend to adapt a radically precautionary approach on such interventions. This is true even to a higher degree with respect to children and old people, whose body is especially fragile.

Technological progress remains almost indisputably desirable insofar as it does not venture on a "sacred" terrain, whichever meaning of sacredness one could have in mind. The clash of technological innovation with a culturally defined concept of "sacredness" is the reason why we observe a strong precautionary attitude with regard to nanotechnological applications in foods and cosmetics.

- Should the public be informed of the presence of nanoparticules in food, food packaging and cosmetics? How to balance the benefits (full information, autonomy of the consumer) but also the risks (spreading fears about safety just because the question has been raised) of such information?

- How to regulate the presence of nanoparticules in food? Should industries provide proof of safety, when, and under what form [question of governance]?

- Given that *in vivo* analyses are yet scarce and that we will only see the long-term effects of some nanoparticules in years, what amounts to a proof of safety [question of epistemology]?

- Should the application of the precautionary principle be stricter on food and cosmetics than in other domains (see *The Precautionary Principle*)?

2.4 Information and Communication Technology

- Invisibility: it is now possible to gather, treat and transfer information through devices that are unnoticeable for the human eye. Does this novelty radically change the social perception of ICT? Will invisibility allow new social uses for ICT? Will it render problematic some of the existing uses of ICT?

- Miniaturized artificial objects and the problem of privacy. Will further miniaturization of information-gathering artefacts lead to extended surveillance and control of the State over citizens? Or, on the contrary, to a society where everyone can view and be viewed by anyone else?

- Who controls the data gathered by miniaturized devices? How does one make sure that the mechanisms of control are safe from manipulation?

- Miniaturization of electronic devices allows ordinary citizens to record and upload information quickly and without the need for mediation, thus avoiding third-party control. How does this technology change the social bonds? How does it influence politics?

- What is the impact of ICT on the partition between public and private spheres? Will the spread of ICT blur the frontier between private and public to the point where this distinction will disappear? Will it create a hierarchy of levels between the strictly private and the strictly public?

Private and public

Historically the notion of a private sphere in the individual's life has developed around the question of the choice of one's own religion. Before the Reformation, to follow a particular religion was not a matter of individual choice. The extremely violent Wars of Religion forced the Europeans to invent the notion of freedom of consciousness: each individual could choose his faith for himself insofar as public expression of this choice didn't interfere with the interest of the Prince. This notion of free choice was later extended beyond religion to encompass profession, political opinion and, to a certain extent, moral values. Following a general secularization of European societies, the focal point of the private sphere has progressively shifted from the quest of Salvation to a quest of individual happiness. The notion of a private sphere itself does not come for free nor is it given to *Homo sapiens* biologically. Human history has invented the private vs. public distinction, but human history could also dissolve it.

The division between public and private spheres was theorized in the 19th century by political liberalism. The State should only impose such constraints on the individual that are necessary to ensure their peaceful coexistence. Exact appreciation of such constraints can vary from one European culture to another: for instance, in Britain security and justice dominate, while Germany and France tend to focus on social welfare and education. But the general idea is the same: the public sphere is restricted to the conditions of living together, and all the rest is private, i.e. all particular actions of an individual citizen belong to his private sphere, where he alone defines his well-being.

According to political liberalism, personal choice must be autonomous and must not be influenced by the State or other citizens. Hence, the State should guarantee that the private does not become known in public. This is the notion of privacy.

The 21st century may witness an important blurring of the private-public dichotomy. Social networks such as Facebook render public huge amounts of data on the private life of individuals. For example, even one's own stream of visual experience can now be shared through EyeTap, a miniaturized camera placed above the eye to record what one sees and upload it on the Web. On a social network, the value of a person is the amount of attention she attracts: you are worth the number of friends or visits you get. Private data has thus become a tool for enhancing one's value in the public space. Contrary to the 19th century, individuals care much less about keeping private things private. The formerly private has now become an instrument to attract the others' attention. On the other hand, the sphere of public debate, where previously matters of common good and living together have been discussed, is now invaded with huge amounts of the individuals' private information. The private vs. public distinction thus suffers a double blow: both at the level of individual data which becomes an instrument of public life and at the level of the content of the society's public space. The frontier between the public and the private gets blurred or indeed disappears.

One can also argue that the frontiers of the private sphere get redefined rather than blurred. Private information becomes an easily available resource used for social interaction. Even if this interaction is potentially accessible to all, it is normally only used by a selected group of `friends', which effectively restricts its publicity. We may witness the creation of an entire hierarchy of levels between the "private" and the "public" spheres in the classical sense, which will render inapplicable our current norms of moral judgment and legal dispositions based on the strict public-private dichotomy. The new borderline of what the individual keeps strictly for himself may vary from one person to another so that the private vs. public frontier will not be the same for everyone. This threatens the political function of the private vs. public distinction, which would require the existence of a frontier common to all.

- Problem of **human autonomy** in a world where information-processing is ubiquitous in all spheres of human life. Can constant observation prompt individuals to conform themselves to prevailing norms in anticipation of the fact that they will be observed? Or is one always free to obey or disobey incentives?

- **Ambient intelligence** should alleviate us from routine but time-consuming tasks, such as shopping for food. Intelligent refrigerators will be able to compute our profile and needs and to order food on the internet. A gain of time gives more freedom to attend to more rewarding tasks. Can it also imply a lack of autonomy, for example because it will be harder to change one's consumption habits?

- **Persuasive technologies** have already been developed to influence people's behaviour. For example, the Foodphone shows obese people what they will look like in 10 or 20 years if they don't put on a diet. Accenture's Persuasive Mirror deforms one's likeness according to the health risks ensuing from one's way of life. Do persuasive technologies diminish human autonomy?

- **Brain/Machine interfaces (BMI).** Technologies spanning from cochlear implant to neuroprosthetics and neurofeedback already allow direct interface between the central nervous system and computers or robots. While they remain restricted to test use or medical necessity, these technologies have the potential to transform human communication on a large scale. What limits and controls should be created to frame the development of BMI?

- BMI blurs the frontier between human body and machine. What influence will BMI have on the meaning of human dignity, which we currently associate with inviolability of the human body?

- Principles of bodily integrity forbid to remove body parts even if they are artificial. For example, should a patient unable to pay for his artificial limb or a brain implant be allowed to keep it?

- BMI interfaces can significantly alter moods and tempers. Some patients react to deep brain stimulation by adopting behaviour that sharply contrasts with their personality. When our emotions, moods or temper are influenced by technological intervention, are we still responsible for the behaviour that ensues?

Human Dignity

Human dignity is a fundamental principle of European ethics and law. The first article of the 1949 German Constitution states: “Die Würde des Menschen ist unantastbar” (The dignity of Men is inviolable). This phrase is taken up in the first article of the 2000 Charter of Fundamental Rights of the European Union.

In Europe, dignity is a foundation for other human rights, such as mutual respect, education, freedom of speech, etc., and for bodily integrity. **In European Law, dignity prevails over all other rights, including freedom of choice**; hence the European ban on the commercialization of human organs even in the case of donor’s informed consent. In the 2005 opinion on ICT implants, the European Group on Ethics (EGE) advocates a similar ban on making profit on ICT implants (e.g., by reselling them). In so doing, the EGE implicitly (a) considers that ICT implants are fully part of the human body just like biological organs; and (b) confirms that bodily integrity is integral to human dignity.

What gives Dignity this absolute supremacy? Like many other European ethical or law institutions, the EGE considers human dignity to be “intrinsic to human substance”. But what is it about human beings that makes them so special among all animals?

It is hard to find a clear answer to this question. Most often, the implicit assumption is that Man is endowed with reason. This line can be grounded either in religion via the notion of soul or in moral philosophy such as Kant’s. Still, we know from biology that there is no natural clear-cut break between Man and other animals. A chimpanzee, for example, can learn a language or imagine complex goals better than a human baby. Why shall we refuse a monkey the dignity that we bestow on human babies or embryos? Is a sheer potentiality of becoming a member of human society in some remote future sufficient to separate such biological organisms from all others?

Another difficulty arises because of cultural diversity. If dignity were inseparable from human nature, it should be the same everywhere. But what seems contrary to human dignity for a European in the 21st century may be evaluated differently in a different cultural context or may have had a dramatically different meaning in the past. For example, robots are used in Japan to take care of the elderly and are considered almost as members of the society in their own right; or a public execution of a criminal is unacceptable to us today, while it used to be a major public attraction several centuries ago.

It is very likely that the idea of a sacred substance of Man, whether it is rooted in soul or in reason, will come into conflict with scientific knowledge. Neuroscience progressively reveals and manipulates brain mechanisms that underlie human “mind”. If dignity cannot be grounded in any scientific fact, how can we continue to justify it? An evolution of the meaning of this notion seems inevitable, although the term *dignity* itself will continue to play a key role in our morality.

The Cyborg

A term coined in the 1960s, *cyborg* is an abbreviation for "cybernetic organism" and has many synonyms such as "bionic man". Science fiction, especially films (*Terminator* or the *Six Million Dollar Man),* popularized the cyborg, who looks like a human being while his body is in fact stuffed with artificial devices based on information technology.

Some view the cyborg as a doomsday prospect, while others consider it a chance for the emancipation of humanity from biological constraints. Libertarian thinkers, especially in the US, claim that men are already on the way to become cyborgs. Things as different as tennis rackets, computers or toothbrush already serve as artificial prostheses to the human body, because they enhance a function of the biological body or add a new one. Still, even if they considerably augment our cognitive capacities such as memory and change our relationship with other human beings, they remain external to the body itself. Transhumanists call for an intensification of this trend: if our bodies become cyborg by internalizing prostheses, this would enable us to radically change not only the human condition in the world, but the human nature itself. The questions are: what will "we" mean in this case? Can we recognize "us" in these modified bodies? Is our body a fixed given that cannot be modified or do we possess the power to choose a different material support for our consciousness? Also, will this future cyborg still identify himself as continuing the history of humanity?

Replicants and robots

Richard Sennett pointed in *The Craftsman* that some technical objects can be viewed as "specular" tools, i.e. machines that invite us to reflect upon ourselves and the human condition. According to Sennett, there are two types of such instruments: "replicants" (the term comes from the movie *Blade Runner*, which is an adaptation of Philip K. Dick's book *Do androids dream of electric sheep?*) and robots. Replicants are simple substitutes for humans but robots are better than humans in performing a given task. Because they excel in the formerly exclusively human role of *operator*, robots present a threat to skilled laborers and an increasing number of less qualified workers. Hence, the very idea of "nano-robots" can be viewed as a threat by many professionals, when they imagine that such robots will replace the need for human labour in their field. Such is the fear expressed in Bill Joy's *Why the future does not need us*, where he imagines machines that can reproduce without any human intervention. Replicants also contribute to humanity's anxiety, because by mimicking us they provoke animosity and position themselves as our adversaries. Consequently, most science-fiction stories involving androids and replicants speak about the conflict between "real" humans and cyborgs. Nanotechnologies do not create pure replicants, but the perspective of growing biomimetic technology shows that this problem may emerge in the future. Although these fictional issues are somewhat far from reality, they raise relevant ethical questions about technology: how nanotechnologies are going to change the perception that we have of ourselves? Seen from this somewhat unusual perspective, are advanced technological objects bound to be our new enemies?

2.5 Nanotechnology in the military: questions of dual use

- Does the European society "automatically" condemn any development of nanotechnology oriented toward military use?
- Will the development of nanotechnologies in the military initiate a new arms race, comparable to the nuclear arms race from the 1950s to the 1980s? If yes, how to make this arms race compatible with world peace?
- Does nanotechnology protect from terrorism or does it provide it with new weapons?
- Will nanotechnology truly result in killing fewer soldiers and civilians at war by reinforcing the protection of the former and targeting attacks more precisely to spare the latter?
- It is often impossible to distinguish between nanosciences dedicated to civil and to military use. How is it possible to restrict or control only the military applications?
- How does a scientist involved in a big research project funded by military agencies or having potential military uses decide if he should continue working or quit? Which argument is more convincing: "It will be done anyway, so I better participate and try to do good from within the project" or "It's military, so I should have nothing to do with it"?

2.6 Questions relative to risk and uncertainty

- Nanotoxicology and ecotoxicology are new and complex disciplines. Materials whose behaviour is known at the macroscopic scale may exhibit a specific toxicity at the nanoscale. Most results showing toxicity of nanomaterials involve high exposure rates and almost exclusively *in vitro* tests. Is there reasonable ground to believe that nanoparticles and nanomaterials *could* be harmful? What action should be taken, i.e. on what grounds should we better protect ourselves and the environment even before risks can be fully evaluated?

- Is there a reason to be more suspicious about nanomaterials than about other potentially harmful substances, e.g. in oil industry?

- 4 to 6% of the EU's and USA's public funds in nanotechnology research are devoted to nanotoxicology. Is this numeric indicator a good way of assessing progress in nanotoxicology? What is the role played by this indicator in policy debates?

- What is the best reaction to those who issue dramatic warnings about nanotoxicity? Shall we refuse to admit that risks are radical? Shall we oppose the risk aversion of developed societies as a symptom of weakness? Or shall we leave it to toxicologists to speak about risks in public?

- Governance of nanotechnology in the face of uncertainty is a major challenge. The EU has adopted the Precautionary Principle. What real influence does it have on research? What influence should it have?

- European citizens live in a "risk society" with its high awareness of the risks inherent to any technological change. What are the consequences of a "zero risk" society? Are we going in this direction?

The Precautionary Principle

There are several definitions of the Precautionary Principle (PP). The definition prevailing in Europe is: *The precautionary principle applies where scientific evidence is insufficient, inconclusive or uncertain and preliminary scientific evaluation indicates that there are reasonable grounds for concern that the potentially dangerous effects on the environment, human, animal or plant health may be inconsistent with the high level of protection chosen by the EU* (Communication from the Commission, 2000).

Thus the PP is supposed to apply in the following cases:

(1) Scientific evidence exists but is inconclusive; probability of harm is a subjective bet rather than objectively established value.

(2) Probability of harm may not exist but future harm itself is envisaged and it would lead to "unacceptable", "irreversible" or simply "serious" consequences.

The PP then requires one to take precautionary action proportional to the harm. The concept of proportionality remains unspecified and subject to interpretation by the decision-maker.

The justification of the PP often makes an appeal to **responsibility towards future generations.** This idea was developed by the German philosopher Hans Jonas (*The Imperative of Responsibility*, 1979) and is now recognized as the main ethical rationale for the PP. However the very notion of being responsible before a being who does not exist is problematic.

The European promotion of the PP sharply contrasts with the US reliance on cost-benefit analysis. The PP does not appear in the US Federal law and only exceptionally in state laws. The US chamber of Commerce considers the PP to be a "relatively new theory" pushed by "radical environmentalist groups" and rejects the PP in the name of "scientifically sound and technically rigorous standards". The European Union defends the PP via a reference to "the high level of public health, safety, consumers and workers protection, and environmental protection chosen by the Community" (Communication on the regulatory aspects of nanomaterials, p. 3). This typically means that if hypothesized harm is significant, the risk should not be taken even if scientific proof of this hypothesis has not been carried out. International competition pushes the EU toward adopting a cost-benefit analysis too. Both the EU and the USA have made research on the risks of nanomaterials a priority. When consensual quantified risks have been objectively established, the PP loses its relevance.

2.7 Questions relative to public communication on nanotechnology

- What are the consequences of nanotechnology being proclaimed revolutionary, catastrophic, or radically novel?

- Social enquiries show that the public still knows little about nanotechnology. Is there a risk of a widening social divide between those who know and those who don't, as nanotechnology becomes more widely used? How to counter this divide through education or outreach?

- Nanotechnology will likely influence our lives. In a modern democracy, this implies a need for transparency and information for citizens and for the involvement of the public in the decision-making. Should the scientist remain neutral in such debates in his position of expert or should he act as a scientist *and* as a citizen?

- How should governments and research institutions take into account the outcome of such debates? Should they be advisory or mandatory?

- Ethical education is usually absent from science and engineering curricula. What are the arguments for and against its introduction?

- Is the scientist trusted or feared by the public? If both, how can the scientist understand and deal with this paradox?

- What is the best reaction to hype: support it, ignore or refute?

- Researchers are tempted to use positive and negative hype strategies (e.g. by overemphasizing toxicological risks) to get attention or to attract funding to their research. What are the consequences of this use of hype for the society and for the research community?

Hype

Nanotechnology became a major direction in research policy with the launch of the National Nanotechnology Initiative (NNI) by President Clinton in 1999. In 2002 a report from the US National Science Foundation listed the benefits to be expected from nanotechnology, predicting a new golden age:

> Technological convergence could become the framework for human convergence. The twenty-first century could end in world peace, universal prosperity, and evolution to a higher level of compassion and accomplishment. It is hard to find the right metaphor to see a century into the future, but it may be that humanity would become like a single, distributed and interconnected "brain" based in new core pathways of society. This will be an enhancement to the productivity and independence of individuals, giving them greater opportunities to achieve personal goals (M.C. Roco and W.S. Bainbridge, *Converging Technologies for Human Performance*, NSF, 2002, p. 6).

Such statements seem to be pure dream but they were instrumental in convincing both the US government and Congress to further fund the NNI. As a consequence, between 2002 and 2005 its budget more than doubled. The use of visionary language to attract public attention is called **hype.** Hype is fictional but has very real effects.

2.8 Questions relative to visions and fictions

- From the very beginning nanotechnology was promoted by fiction (Drexler, *Engines of Creation*, 1986). Nanoscience fiction can be utopian as well as dystopian, i.e. prophesising a new golden age or a new apocalypses (Crichton, *The Prey*, 2003). Imagine a scientist is working on a new development in nanotechnology. Can the scientist predict which of the new features that he's developing will be most fascinating for the public? Is it possible to anticipate the fictional scenarios that will be built around his work?

- Artists take part in nano communication. Nanoart produces images with vivid colours and shapes typical of the human scale and adapted for human perception. This adds a uniquely human dimension to the perception of nanometric phenomena, which are perfectly invisible to the eye and do not possess intrinsic colours nor definite shapes. What is the effect of this "anthropomorphisation of nano" on the perception of nanotechnology?

- The prefix "nano-" has been widely used in a metaphoric way (for naming portable music players, automobiles and even cigarettes). What impact does this metaphoric use have on the image of nanotechnology? Do scientists need to worry about it?

Magic Nano

Magic Nano was a protective glass and ceramic sealant on sale in many German supermarkets. In March 2006 it was called back from the German market after causing 110 severe breath problems. Media presented it as the first major accident with nanotechnology. End-of-the-line distributors did not know exactly what Magic Nano contained and were unable to justify “nano” in its name. The ingredients were then analyzed by the German Federal Institute of Risk Assessment. It turned out that Magic Nano contained no nanosize material, although when sprayed as an aerosol it covered a surface with a film less than 100 nm thin. The hazard was due to droplets emitted through the aerosol which could penetrate into the lungs. The product itself contained no nanomaterials but was instrumental in creating the conditions for hazard production at the nanoscale. This is an example of both a positive use of the “nano” label for marketing and a negative use for creating suspicion regardless of whether nanomaterials are actually present in the product. In 2006 no one had studied independently whether the sealant was “nano” until the breath problems occurred.

2.9 Questions of social justice

- Nanotechnology-enabled products are often costly. What effort should be made to guarantee equal access to these products?

- Labelling a patent as “nanotechnology-related” can enhance its visibility on the market but also hide the concrete disciplinary applications of this patent. Should “nano” be a part of patent classification or should it remain domain-specific? Should patent rights in nanotechnology be reformed to allow better circulation of scientific knowledge and know-how?

- Will nanotechnology result in a deepening of the gap between North and South? Nanotechnological devices consume less energy than macro devices. Is this sufficient to claim that they provide new and more effective resources for economic development?

- Will nanotechnology’s complex know-how lead to the concentration of decision-making and economic power in the hands of a scientific and technocratic elite? Will it reconfigure the existing situation (its structure, the steakholders or their relations)?

Distributive justice

Distributive justice is concerned with dividing a common good between all those who have shares in a group or in a society. Typical questions include the division of wealth between citizens or between employees in a firm. Distributive justice also concerns social goods and rights such as health, education, land, freedom, knowledge, and power.

The Aristotelian principle has it that a just distribution must be based on merit. How exactly merit is defined and measured varies according to the nature of the good being distributed. The matter is often open to debate. With regard to income, a plausible criterion for merit is the amount of work people invest in creating common wealth. With regard to education, all people equally deserve a good education. With regard to technology, who deserves to have a privileged access? Is technology like education, i.e. should all people have equal rights to access and use technology? Or is it like specialized knowledge, where a hierarchy (i.e., inequality) exists based on cognitive merit and the time that goes into learning a complex subject?

Aristotelian principle of justice has competitors. In his *Theory of Justice* (1971), John Rawls proposed a "Difference Principle", which allows for unequal distribution of a good if inequality benefits the least advantaged: "All social primary goods – liberty and opportunity, income and wealth, and the base of self-respect – are to be distributed equally unless an unequal distribution of any or all of these goods is to the advantage of the less favoured." Like with the case of the Aristotelian principle of justice, there exist objections to the Rawlsian principle too. When the degree of inequality becomes significant, individuals tend to lose their sense of belonging to the same group or community. Rawls assumes the existence of some common-knowledge method that allows one to compare the well-being of different individuals in the situation when they possess, or do not possess, a certain good. Strong inequality means that no such principle is available within the group and the concept of distributive justice according to Rawls loses its meaning.

Can the Rawlsian principle be applied to technology? Can the concentration of technological know-how in the hands of a few lead, at the end of the day, to a benefit for the least advantaged? Or is an equal distribution a necessity? Is technological inequality too strong, leading to a breakdown of the Rawlsian idea of justice?

2.10 Questions of responsibility

- Technology is about introducing new instruments into the world but it does not dictate how these instruments should be used. It doesn't carry in itself a moral value. On the other hand, technology brings about changes in the way of life. In the long term, these changes can lead to a redefinition of societal values and norms. So how can we still say that technology is value-neutral?

- Is the scientist only responsible for the uses of his work that he develops intentionally or is he also responsible for all social uses that will occur in the future? Or, again, only for those that he can reasonably anticipate?

- Individual vs. collective responsibility. Nanoscience and nanotechnology involve large-scale institutional science which produces collective research results. At the same time, responsibility remains associated with individual human beings. Should we say that no particular individual can be held responsible for the consequences of nanotechnological innovation? What would it mean to say that only a collective is responsible? Is it the same as putting the responsibility on every member of this collective?

- It is impossible to predict all future uses of scientific discoveries. Harmful uses of our technology will put the blame on us as its creators. What are the implications *hic et nunc* of this unpredictability? What is the impact on research funding and the choice of priorities in public and private R&D?

- Nanotechnology will influence the way in which we live. Many think it will even irreversibly transform the human condition. Are we responsible towards future generations? What entitles us to judge on their behalf? How do we know that they will have the same values and norms of judgment as we do?

Responsibility: collective and individual

The public often points to individual scientists being responsible for technology when it leads to harm. The scientist may have formed no intention to use the results of his work for a particular harmful purpose, but if someone else did, then the scientist is held responsible by the society because he has *brought about* the technology needed to produce that harm. This raises the question of responsibility without intention, which is a paradox from the point of view of law (liability without intention cannot exist), but a well-known albeit problematic case from the point of view of moral philosophy. Max Weber was among the first to distinguish the notions of ethics of responsibility (*Verantwortungsethik*) and ethics of conviction (*Gesinnungsethik*). The first one applies to those who are in charge of a group and who must foresee the consequences of their actions on the whole collective, while the second one applies only to a particular person who acts according to his own beliefs. Weber himself used this distinction in the philosophical analysis of the politician (*Politik als Beruf, 1919*). Today – and probably unexpectedly for Weber himself, for whom science was neutral and devoid of values (*Wissenschaft als Beruf, 1917*) –, it applies to the scientist too.

Individual researchers are held responsible for the collective action of institutions, firms or groups they belong to. Collective responsibility in the legal sense is unacceptable, because liability is always based on individual will. Still this phenomenon exists *de facto*. How to make sense of it? This paradox was famously studied by Hannah Arendt. Collective responsibility is neither legal nor does it refer to individual morality. Rather, it is political and it is used as an argument for discrediting one's opponent in a debate on future policy: you are guilty, because you belong to a certain collective.

- **Legal responsibility** is liability: you are responsible for what you have willingly done or intended to do. Liability leads to legal normative judgment.
- **Moral responsibility** leads to moral judgment. It is not codified by law but encoded in the morality of the society or moral values of the individual, both of these being vague and subject to debate rather than a univocal interpretation.
- **Political responsibility**: Many factors bear on the judgment pronounced with regard to our actions, such as for example our position in a social or scientific hierarchy. *Even though we haven't really chosen them, these factors do not exculpate.* Political well-being of the institution to which we belong is decided among policy-making individuals who personify the institution, although they may have not participated in the elaboration of a particular technology. On the other hand, the scientist who discovers new technology does not decide what development it will receive. Whether they want it or not, all members of the institution, from a working scientist to the director, have a share of responsibility simply by being members and contributing to the collective action of the institution. This calls for an ethics of responsibility in Weber's sense. Individual ethics of conviction is not enough.

2.11 Questions of epistemology

- Do nanosciences carry radically new scientific conceptions and theories about the behaviour of matter at the nanoscale?
- Is there a sharp difference between nanosciences and nanotechnologies?
- Has technoscience become a new paradigm of knowledge, i.e. is it possible to maintain that pure curiosity has been replaced by the know-how science, where knowing and making are one?
- How are scientific institutions modified by the growing proximity of science with technological, industrial and economic outputs? What is the relationship between the end of the Ivory Tower science and the spread of grant-funded research?
- Is nanotechnology a unified research program? Does it stimulate the emergence of a new, interdisciplinary scientific community with its own culture and body of knowledge?
- Is the ethics of science compatible with that of the entrepreneur?
- The scientist is often pictured as “morally neutral”. How can this be compatible with the fact that science has its own ethics, around values such as truth, honesty, progress, and advancement of knowledge?
- Is disinterestedness a core component of scientific ethics?

“Plenty of Room at the bottom”: myth or reality?

Richard Feynman’s famous “Plenty of room at the bottom” speech at the American Physics Association in 1959 is often quoted as expressing the fundamental visionary idea of nanotechnology. This is by far the most quoted paper in nanosciences today, although it has been scarcely noticed when it originally appeared (see Ch. Toumey, *Nature Nanotechnology*, 4, 2009, pp. 783–784).

In the “Plenty of room” speech, Feynman talked of the possibility to write all of *Encyclopaedia Britannica* on the head of a pin. Miniaturization of transistors has since made this vision a part of reality. But back in 1959, when computers were nowhere to be found except huge machines in military and cybernetics research centres, Feynman had also made it clear in his speech that he had absolutely no idea of the technology which could realize his dream. Indeed, the word “nanotechnology” did not even appear before 1974. Feynman did not *name* nanotechnology in his speech nor indeed was he himself a nanoscientist.

Our current abundance of references to the “Plenty of Room” speech is a typical case of retrospective reinterpretation which elevates an otherwise ordinary event to the status of a mythical Original Act. Since the 1980s, Feynman’s vision has helped to create and to keep together the nanoscientific community. Indeed, nanoscientists coming from various disciplines do not spontaneously share any substantial scientific body of knowledge, for their expertise lies in heterogeneous fields (from computer science to medicine, electronics or material science). For the same reason, they have very different work cultures and research practices. The great name of Feynman, known and venerated by all, is a symbol that unites this community around a common vision of the goals of their new field. But in truth the authorship of the visionary promises of nanotechnology belongs to this community and not to Feynman.

Part 3.Thinking with the help of ethical concepts

We present in this section a selection of philosophical concepts, which are at the same time useful ethical tools. These concepts are rather simple. However the difficulty for the scientist usually lies in the identification of a moral question and the application of these tools to such concrete questions. Hence the composition of this section: we begin with a clear question, we then present a case study where it is exemplified, and finally we introduce a relevant philosophical concept. Some of these cases are imaginative while others are connected with the reality of today's laboratory science. However, both sorts of issues are part and parcel of the public debate about nanotechnology.

A. What is an ethical question in nanotechnology?

Case study: Take toxicology of nanoparticles. Toxicological studies are mandatory both for societal acceptance of technology and for its harmless use. As a discipline, toxicology strives for a correct description of the toxicity of a given material. It is descriptive, i.e., not normative. Norms are set with regard to the result of the toxicological study, but they also involve a human appreciation of how much damage is acceptable.

Think of this case using the notion of Ethics. Ethics is normative. It is about how the world should be – as opposed to how it is. If we treat 'ethical' and 'moral' as synonyms, then ethics is concerned with what is right and wrong, good and bad. Toxicology in itself is a scientific, not an ethical question.

B. Do moral values and norms evolve in time?

Case study: Take the example of robot rights. Imagine that a future technology will create ubiquitous robots for whom we'll have affection and who will respond to us emotionally. Will a robot need love? Should a man be allowed to love a robot? Can a robot propose to a girl? Can a robot retire and claim a pension?

Think of this case using the notion of Metaethics. Metaethics is the study of ethical systems. It is chiefly concerned with the meaning of moral terms ("good", "right") and with the evolution of this meaning in time. Technology modifies the norms that exist in the society, but does so on a much slower time scale than the one at which operates technological innovation. Values change even slower. Values and norms that belong to different human generations are often incompatible: e.g., slavery was normal for ancient Greeks and Romans and still for a lot of people in the 19th century. Animal experimentation did not pose any problem to our grandparents. So it is not inconceivable that the concept of dignity in the future be extended to robots.

C. Does nanotechnology pose specific ethical questions?

Case study: Some people claim that “nanoethics” is a distinctive field of ethics, like bioethics for instance, because it is induced by a new generation of cutting-edge technology. Systems in applied ethics are typically characterized by the specific technology that enables a new set of moral questions. Contrary to robotics or animal experimentation, nanotech involves a variety of highly heterogeneous techniques in different areas of science. If the scientific field can be formally unified by a scale (that of nanometer), it is increasingly difficult to argue that nanoethics could be conceptually unified.

Think of this case using the right definition of Applied ethics. Applied ethics is a set of professional applications of ethics to a specific scientific field. These include bioethics, medical ethics, environmental ethics, animal ethics, neuroethics, roboethics, etc. Whether nanoethics belongs here is often a matter of terminology. Yet there is more to this question than a choice of words: one must face the blurring of frontiers between particular domains of applied science and engineering, so the frontiers between respective systems of applied ethics become more difficult to define.

D. Is technology in itself good or bad?

Case study: One can hear a full spectrum of opinions. Luddites deem technology undeniably bad. Technophiles think it is always good. Sometimes it is said that technology is neither good nor bad, hence has no moral value in itself: technology merely provides new means of action but not goals. This leads to the laissez-faire attitude of the scientist’s “inborn ethics”: as Feynman said, “Shut up and calculate”, i.e., do your best as a researcher or an engineer who creates new things. Only when a discovery has been made or a new technology has been proven feasible shall one start thinking about ethical questions. Society tells the scientist to abandon this Ivory Tower attitude, yet the scientist resists and argues that it is impossible to do science otherwise. Is nanotechnology a challenge to the laissez-faire attitude or can it claim ethical irrelevance, delegating all moral issues to communication and policy-making?

Think of this case using the notion of Ethical theory. Ethical theory is a set of moral doctrines, each having distinctly different rules by which one determines what is right. Three major ethical theories are consequentialism, deontology and virtue ethics.

Three ethical theories

Consequentialism: Whatever the intention, the value of an action is determined by its consequences. Ethical assessment amounts to a cost-benefit analysis. John Stuart Mill (1806–1873) was a staunch consequentialist. His sort of consequentialism goes by the name of utilitarianism: the guiding principle is maximization of global utility for all. Weaknesses:

- Is there a universally agreed norm for what counts as a benefit and what counts as a cost?
- If a wicked intention results in an unintended overall benefit, consequentialism still says that the action is good. This seems to contradict our moral intuitions.

Deontologism: Whatever the consequences, the value of an action is given by the intention of the agent. A good intention is an intention that abides by an unconditional, universal moral principle. According to Kant (1724–1804), one should never treat a human person as a means of action, for she deserves unconditional respect. Modern deontologists value human dignity above all other principles including progress and knowledge. This view holds dignity as an absolute. Weaknesses:

- How do we deal with conflicting principles, e.g., unconditional respect for human person and the advancement of medicine in the case of stem cells research?
- A good intention may sometimes lead to disastrous consequences, which a deontologist cannot condemn according to his moral theory.
- Principles which we believe to be universal may vary over time and across cultures. Deontologism cannot account for this variation, so the deontologist will typically hold that the meaning of principles does not change, e.g. “human dignity” has always meant exactly the same thing throughout history.

Virtue Ethics: There is no general rule to assess the value of a particular action. It all depends on the circumstances. Only virtues present in the agent’s character, such as courage, loyalty, justice, honesty, etc., can guarantee that the action is good. Different life situations call for the manifestation of different virtues. Aristotle (384–322 BC) invented and defended virtue ethics. Principles may say what is good in general but are helpless for making concrete choices in complex situations. Acting well requires situated judgment and the capacity to discern between good and bad even in unpredictable circumstances. These qualities require experience and, above all, a good character. Weaknesses:

- Virtue ethics concerns individual virtues and does not show how to promote them to abstract principles of action applicable to all; one is left to merely imitate the individual who possesses certain virtues.
- If general rules for action cannot help, how can we rationally justify our acts? Thus virtue ethics seems to open the way for arbitrariness.

E. Would you like to live in a world where everyone is judged by his or her body's propensity to develop a future disease?

Case study: Early diagnostics of diseases with the help of nanomedical tools is a typical case of consequentialist reasoning, If a future disease leads to the consequences judged as morally bad, such as pain or death, it is preferable to avoid the disease before any such consequences can occur. Pushed further, early diagnostics deals with probabilistic predispositions to diseases rather than the factors that cause them deterministically. Predictive medicine enables early diagnostics, as well as – in certain cases – predictive treatment of future ailments, and is instrumental in turning the human body, with all its formerly random biology, into a utilitarian machine.

Think of this case using the notion of Consequentialism (see bow 13). In order to reason consequentially, one must specify which consequences are morally relevant and how much weight we should give them. Can a simple algorithm assess the goodness or badness of a probable disease? Can there be a universal such algorithm applicable to all individuals?

F. How far do we go with remote control technology? What if the brain of one person commands the body of another?

Case Study: Remote control of the human body seems within reach. Imagine a tiny communicating implant forcing another human being to raise her hand. Her own awareness or desire to act have no bearing on the performance. Are good or bad: a) this technology, b) the development of this technology by the scientist, c) the use of this technology by another human being?

Think of this case using the notion of Deontology. Independently of the consequences of the action (raising one's hand may either do good or harm somebody), a negative judgment can be justified solely by the principle of human dignity. Neither consequentialism nor virtue ethics provide an immediate judgment that the use of such technology is always *wrong*.

G. Should the scientist be trusted?

Case Study: Once upon a time scientists were bestowed with trust. To be a scientist meant to be wise, reliable, and rational. All these virtues contributed to justifying the scientist’s opinion or action whatever content they had. The end of this Ivory Tower science and the loss of trust have altered this situation dramatically. The scientist is now in the center of public arena, equal to other mortals, and has no more the moral standing of a virtuous man.

Think of this case using the notion of Virtue ethics. What are the virtues of the scientist? Of the expert, i.e. a scientist speaking out in public? Do these virtues suffice to justify the scientific stance and the action of science in the eyes of the public? Do scientists need to acquire other virtues, e.g. eloquence or ecological militancy, in order to restore society’s trust?

H. Am I sure to be right if I follow all the rules?

Case study: If one follows in his work a set of rules of professional conduct, such as one of the codes of conduct in nanotechnology, it may give the impression of a perfectly moral scientific practice in full agreement with ethical principles. Nuclear research in the 1930s or the development of asbestos technology in material science were examples of such conduct, yet later they were morally condemned due to harm far beyond control of the researchers who developed the original technology.

Think of this case using the notion of Moral luck. Moral luck is a special situation showing the limits of any form of rule-based ethics. It is a situation when a future outcome, which is accidental and not willfully intended, will have retroactive impact on the way the agent's conduct will be judged. This *post factum* judgment is clearly in opposition to the Kantian, markedly deontologist insistence that morality is immune from luck and to the consequentialist’s attempt to eliminate luck by predicting future benefits and costs. The case of moral luck underlines the presence of chance in life. It suggests that one’s positive intention does not suffice for a positive result; rather, one should stay alert during the development of a given project or action and not claim the benefit of success.

Moral Luck

There seems to be a contradiction between two moral intuitions:

(1) We think that responsibility extends only to what depends on us.
(2) We often blame people for unforeseeable consequences of their actions.

Examples of (2) are:

- John Smith is driving back from a wedding party. He has drunk two glasses of champagne but feels fine, although his reflexes may be a little slower than usual. Suddenly a child runs in front of the car. On the first scenario, John Smith pushes the brakes and stops just in front the child, who is a bit frightened but otherwise safe. On the second scenario, John's car runs over the child, who is killed. John is then arrested and judged, and the court's verdict is very severe, because John was drunk while driving. Now, the moment at which John took his car after having drunk some champagne (call it t_1) is strictly in the past of the time when a child appeared before his car (call it t_2). Normally in moral philosophy John's conduct at t_1 must be judged on the basis of John's intentions at t_1 as well as prior to t_1. The future cannot influence moral judgment, yet it does! On the first scenario, what occurred at t_2 has no repercussions on the judgment with regard to the event that occurred at t_1. However, on the second scenario it is perfectly clear that killing a child at t_2 has influenced the verdict. John Smith could not have known what would occur in the future: it's a question of pure *luck*. The influence of this luck factor on the moral judgment with respect to past events is called *moral luck*.
- Gauguin is now a famous painter, often acclaimed as a genius. He painted his most famous works in Tahiti. When he left for Tahiti, Gauguin abandoned his family, leaving them without money and in dire straits. He did so because he wanted to become a great painter, and believed he needed a more "primitive" environment to develop his talent. He was lucky enough to succeed. Had he not succeeded, we would judge his decision to abandon his family morally wrong.

Philosophers Thomas Nagel and Bernard Williams (*Moral Luck*, 1981) observe that the notion of moral luck doesn't seem rational yet cannot be dismissed. Our lives are influenced by "luck" (=random) factors such as education, environment, personality, or genes. If one wanted to eliminate luck altogether from the assessment of responsibility, we would at most be responsible for the firing of certain neurons in our motor cortex, or for purely mental acts of decision. Since our actions depend on factors which we don't control and cannot even predict, responsibility is always in part for things that do not depend on us.

I. What if a device designed to do good is used to do evil?

Case study: Imagine a medical scientist who collects loads of personal biological data for developing a new treatment of cancer. For scientific purposes, it is unavoidable to store this data for long periods of time and to grant access to it to numerous members of the laboratory. It may happen that, once collected, such private biological data be used for other purposes, e.g., surveillance or insurance pricing.

Think of this case using the notion of Double effect. Double effect is a special case of deontological ethics, when one makes a difference between causing harm or evil as an unintended side-effect or intending it directly. Justification of action from double-effect requires that: a) the intended action is good; b) the agent intends the good effect and not the evil one, even as a means for his goal, and c) the achieved good is bigger than the achieved evil.

In practice, double-effect reasoning faces the difficulty of distinguishing between known side-effects and intentional outcomes. In the imaginary case above, it will be argued that even if the scientist knows or suspects that the stored data will be used by a health insurance company, he only gathers these data for research purposes. A proof of his intention is provided by the claim that, even if the data were not ultimately to be used for health insurance purposes, they would retain their value as means for scientific inquiry. Generally, the chosen means is intentional just in case it is necessary to achieving one's aim, whereas a side-effect is a consequence that could have been avoided without compromising success.

But what if the insurance company also funds the research?

Part 4. Responsible communication

4.1. Nanotech halo

Public debate on nanotechnology was initially dominated by extreme, totally unrealistic visions. Until a few years ago, social utopias and other futuristic optimistic scenarios were opposed to no less radical dystopic visions, focused on a major catastrophe to which nano would doom us. Since then, the debate has shifted toward the topic of "responsible research", even if nanohype still occupies an important place in the literature. Various social groups that take part in the public debate on nanotechnology are easy to identify. Science fiction writers use “nano” to spice up their novels, and artists, to add graphic and bruising colours to their works. Laboratory scientists hardly express themselves at all. Perhaps only toxicologists attract public attention when they warn of the risks associated with nanomaterials. Computer scientists are audible too, for they recycle in nano a futuristic discourse previously developed with regard to artificial intelligence. Policymakers use the promise of “nanoworld” to justify reorienting research policy towards essentially economic ends. They often refer to visionaries, who dream that new technology will radically change our lives. This configuration tends to reduce the examination of nano-related ethical and social issues to the assessment of toxicological risks and opportunities for investment. If they do not want to be spectators in this debate, scientists must realize that society does not evaluate nanotechnology products solely on the basis of technical criteria and values acknowledged within the scientific domain. If scientists only tell others about their science, they leave the field open for public anxiety with regard to what happens when the results of scientific research are taken out of the laboratory into a bigger life.

Other factors come into play: the ubiquity of technology, its impact on environment and economy, and also the overall spirit that it gives to our life. Most often, these are the factors which define, more or less explicitly, our opinion about technical objects. The way in which these objects fit into our culture puts their scientific assessment into a larger perspective. More than by anything else, social career of technical objects is determined by their “psycho-social halos”. Every such object circulating in the society is surrounded by an ensemble of perceptions and opinions that redefine its identity based on its cultural resonance. Following the French philosopher Gilbert Simondon, we call it a halo. Like the halo of a source of light, the psycho-social halo of a technical object becomes larger and more significant than the technical virtues of the object as such. The halo *impresses*, so it is of great importance. Inside as well as outside the scientific field, nanotechnological halos are a seed for symbolic crystallization: cultural echo is amplified by the use of the prefix "nano".

Today, any object which is called “nano” is perceived as modern and innovative. Its halo promotes desirability; its attraction goes far beyond the mere technical utility of the object. The halo of nano makes us believe that nanotechnology will improve our life, although behind the shining light of this halo we often remain blind to the actual technical identity of each object, and to the realistic degree of improvement it could possibly bring into our lives. With too much shining, one is unable to distinguish the details. Thus, adding a prefix “nano” to the name of a technical object intensifies public attention and is conducive to both strengthening and polarization of emotions. There is a clear need for ethical insight to control these collective sentiments.

4.2. The virtues of principled communication

Aristotle said that virtuous act is a middle way between excessive and insufficient action. Courage, for example, is to be opposed to cowardice but also to temerity. Not only does courage lie between these passions, for indecision or indifference do too; it also corresponds to a “happy balance” of the extremities. The courageous man acts only at the right time, i.e. when he can dominate fear rather than ignore it. Courage does not blind: one remain lucid throughout the action. It is then considered virtue and the actor is endowed with trust.

Responsible scientific communication proceeds similarly. Communicating the technical content of scientific research is not enough: the public wishes to know what kind of person the scientist is. Here, fortunately, the scientist has an advantage: his own ethic (see *Glossary*). Conveying the fundamental principles of the scientist’s ethic is part and parcel of scientific communication.

There is no Ivory Tower science in today’s world, and a life of research can no more be isolated from the wider societal concerns. Emerged from the fresh ruins of the Ivory Tower, the scientist is put in a debate with the citizen, the industrialist, the expert, the engineer and the politician. Whatever the scientist says, he will be heard as either setting free the demons of fear or the angels of hope. Both of these will come back and chase the scientist: his future depends on the attitude of society. When this happens, as a scientist, assume your responsibility. If you deflect it, your argument will become weaker than your opponent’s. You may lose your future. Avoid reckless chattiness as well as diffident soliloquy. Choose a middle way: even if a scientific argument does not tell society what ought to be done, measure your public speaking by the standards of modesty, lucidity and openness as one does among peers. Courageous Greek men did not always win battles, but they were nevertheless proclaimed virtuous. Keepings to the values associated with fair research will at least not betray the trust with which science is endowed by society.

	lucidity	openness	modesty
Reflexive application: to oneself	Be aware of the principles you violate: admit honestly to yourself that communication may involve hype and exaggeration.	Say what you internally believe to be true, do not conceal what is important for the argument.	Understand uncertainties and weigh your words against them.
Intersubjective application: to others	Be aware of the perception of your message: public reaction is often other than rational.	Let others see the effort you're making. Listen to what they have to say in response.	Accept that there can be a grain of truth in what the other says, even if you disagree completely.
Global application: to society	Be aware of the myths, fictions, symbols and archetypes: your public message often uses and always evokes many non-scientific narratives.	If you are challenged but decide to maintain your view, do not appeal to gloomy authority or hermetic knowledge, better promise your opponent to continue the conversation.	Doubt your competence beyond its limit.

4.3. Perception is framed by narratives

Communicating on nanotechnology is both necessary and utterly difficult.

It is necessary, because science is no more in an Ivory Tower position: when technoscientific research leads to a large scale production of ubiquitous devices that change people's daily lives, citizens hold scientists accountable for at least a part of this change. Technological progress is no more consensually accepted as being *good*.

At the same time, scientific communication is very difficult because of the public's lack of knowledge and methodology and, even more importantly, because of the public's resistance to consider nanotechnology from a realistic vantage point reflecting actual science as it exists today. A recent European report on the public perception of nanotech[3] shows that even after being given information about nanotechnology, laypersons tend to react by invoking ancient narratives and images that certainly do not take in any scientific information. Many of these narratives were already prominent in the public's perception of previous generations of advanced technology. Like genetic engineering, nanotechnology is feared to "open Pandora's box", nanoscientists are seen as "playing God" or acting as "the sorcerer's apprentice", they are believed to be "messing with nature" and the public tends to think that nanotechnology will lead to "the rich getting richer and the poor poorer".

Such narratives have been around for centuries and sometimes millennia. Pandora's myth first appears in Hesiod's *Works and Days*, 7th century BC. The written version of the Sorcerer's Apprentice is a tale by Grimm brothers, 19th century. Why is the technology of the 21st century understood through the prism of ancient narratives?

The answer may lie in our most general relationship with technology. Unless they fall within their professional area of expertise, most people don't know how technological devices work. We approach them as if they were black boxes: we know the input we feed into them (say, pressing a particular button) and we expect a promised output (taking a photograph or making a phone call); but we don't have a faintest clue as to what happens inside the device between the input and the output. We feel annoyed, powerless and even anxious when our computers, cameras, TV sets, cars, etc., break down, for we cannot say what went wrong inside. In such cases, many of us tend to deal with machines as if they were intelligent or sentient beings: we pray or curse them, we endow them with will ("My computer *just won't* work"), emotions and moods ("My cell phone is slow, it's a bit grumpy today"), or cognitive capacities ("Stupid machine!").

This attitude is a typical case of a magical relationship. Like rain dancers invoking a god to make water pour from the sky, we are ignorant of the functioning of our proximate environment. We appeal to the *soul* of the machine as if it could make a phone call go through or a film finally appear on the black screen. We know that saying that a machine is

---

[3] S. Davies, Ph. Macnaghten and M. Kearnes (eds.), *Reconfiguring Responsibility: Lessons for Public Policy (Part 1 of the report on Deepening Debate on Nanotechnology).* Durham: Durham University, 2009. <http://www.geography.dur.ac.uk/projects/deepen/NewsandEvents/tabid/2903/Default.aspx>
See also: Ph. Macnaghten, S. Davis and M. Kearnes, "Narrative and Public Engagement: Some findings from the DEEPEN proaject", in: R. von Schomberg and S. Davies (eds.), *Understanding Public Debates on Nanotechnologies*, Office of the European Union, 2010 <http://www.nanocap.eu/Flex/Site/Page.aspx?PageID=15408>

stupid won't make it work any better – and still we say it. What is going on here? Obviously, we project human, anthropomorphic features on machines. On the one hand, such a projection appears purely irrational, because we are fully aware that machines aren't like human beings. On the other hand, this projection is not new, for humanity has used such practices before. These machines are magical objects as long as we try to please them or to curse them as if we were dealing with gods. Through this relationship technology takes the place of a new sort of divinity whose power transcends the sphere of human action: this is a chief reason why, in our societies, nanotechnology is loaded with symbolic meanings and imaginary powers.

There are two further reasons for this that are specific to nanotechnology. First, nanoscale devices are absolutely invisible and imperceptible. We cannot in principle establish a relation with these objects based on the direct perception by human senses. They are therefore all the more mysterious to laypeople: not only can we not comprehend nano-objects by our minds, but we cannot even perceive them by our senses. Second, many promoters of nanotechnology fuel a "halo" of hype around nanotechnology by spreading the idea that nanotechnology will allow people to go beyond both their human condition and their biological nature. Transhumanists and the proponents of NBIC convergence claim that nanotechnology will bring mankind closer to redemption; unsurprisingly, the opponents of nanotechnology respond that it will put us in front of the gate of a modern version of Hell. Despite this apparent opposition, both "technophiles" and "technophobes" share a common premise that nanotechnology is an unprecedented technological development and a formidable, almost divine tool. The debate between these two radical positions has an air of religious controversy.

The only way to dispel the magical vision of nanotechnology, if it can be dispelled at all, is education. But education has always to start with what students know and think. Although they typically know nothing about the science behind nanotechnological devices, lay people tend to have a lot of things to say about it and to form spontaneous opinions. The magical character of the non-scientific relationship to technology explains why it comes without surprise that such opinions frequently refer to the wonderful world of myths and tales rather than the realities of a science laboratory.

What is surprising, though, is the resistance of non-scientific and emotional appraisals of technology even after scientific explanation has been provided. Why, for instance, do people tend to still cling to myths and tales after they have been confronted with objective scientific information?

Collective imagination is structured by narratives that express deep emotional and normative reactions to important moments of human life: birth and death, sex, blame, violence, pleasure, good and evil, health and disease. Technology transforms our lives and reaches out to such special moments. People may resort to narratives to oppose this influence of technology on the fundamental instants of their being or, on the contrary, to make room for technology and to make sense of their new human condition. This is why narratives, however irrational and unscientific they may seem, should be taken into account when communicating on nanotechnology: they are the only way to understand and modify the reasons of fear instead of simply waving it off as irrational. A critical analysis of important myths and narratives about technology reveals that the "magic" and "blackbox" attitude toward technology is not blindly irrational, but framed by a subtle and powerful reasoning.

Narratives are dangerous to play with, for they are loaded with intense emotional responses. The scientist is normally not trained to use such tools. But he must at least be aware of the presence and of the power of narratives in public perception. The words used by the scientist evoke fictions, myths and tales that often have nothing to do with science or nanotechnology. The scientist needs to understand this mechanism himself. This is the goal of next section of this Toolkit.

Part 5. Narratives of nanotech

Introduction: What is a myth? Why does it matter for science?

A myth is not simply a story taking place in some supposedly remote past. Insofar as it speaks of the origin of nature and of society whose existence continues to this day, myth is a *living* story. But myth is not situated in historic time and its heroes do not have birthdates according to calendar. It is atemporal or, as said Lévy-Strauss, “myth is a machine to destroy time”. Norms and values expressed in myths appear as ahistorical and are never questioned within the narrative. Myths dealing with the question of foundations connect these foundations with the action of a power that transcends human capabilities. Myths dealing with birth or death are often arenas to the action of magical forces, whose meaning transcends the usual capacity of comprehension. In all cases myth is intimately linked with the sacred (e.g., Gods) and projects moral values on the actions undertaken by humans as well as superior beings (e.g., centaurs can be good and titans very bad).

Why do human societies resort to the sacred to legitimize their norms, values and social organization? Man is not simply a social animal: he is an animal who needs to produce and reproduce society in order to live. Social organization and hierarchies vary a great deal: human societies are much less constrained by the natural givens than animal societies. More than any other species, humans are endowed with imagination and symbolic capacities. They are able to transform reality and to present it through the prism of imagination. Ongoing imaginary patterns inherited from the previous generations are encoded in myths; they provide a foundation for the society. Without these tales, social order may appear for what it seems to us from history: somewhat arbitrary, contingent, but assertive in its use of power and violence to perpetuate itself. Social cohesion rests not so much on the memory of past contingencies as on the shared representations of meaning provided in age-worn narratives, myths including.

Our modern societies are complex: some criticize or even want to destroy what others adore. Hence the ambivalence of modern myths and modern readings of old myths about science and technology. Still, myths are useful because they express how society perceives science: they are a part of the social “halo” of technology. Prometheus is both a figure of redemption and of possible damnation. Pandora’s myth expresses the danger of conceiving technology simply as a means to satisfy desire. Golem and Frankenstein express the consequences of a science devoid of moral values and question the place of intelligent artefacts in human society. Whether old or new, myths shed a precious light on the reasons why our societies appreciate or depreciate science and technology.

For a scientist, it does not help to ask if he or she is Prometheus or Daedalus. The answer is that he is not, because he lives and works in a highly complex world where such labels are only good for marketing campaigns. But it helps to compare oneself with, say, Daedalus, and see if questions arise that are common to the situation of moral choice with respect to contemporary technology and the situation in which Daedalus found himself according to myth. Why do we accept new gadgets without asking questions? It helps to ask this question about Icarus who quietly accepted the invention of wings by his father. Is the sin of hubris inevitable or can we indeed restrain technology to good use only? It helps to ask if Icarus could avoid flying toward the sea and the sun. What stands in the place of these images? The sea and the sun in the myth of Daedalus surely have their analogues for every technology that we develop for the benefit of the world. What happens after the catastrophe? It helps to understand that it is not Icarus's death but Daedalus's emotional response to this death that stopped him from making myriads of new wings for other people. So do feelings still govern the technological policy-making?

There is no point in trying to identify oneself as *being* the mythical hero, because the scientist is simply not the same person as these characters. The interest of comparing the scientist with the mythical hero lies elsewhere. Human imagination has always been shaped by the world in which it lived, first a natural world of wild beasts and thunderstorms, then a technological world of airplanes, telephones and medical scanners. The main ethical question with respect to technology: what sense shall we make of the new human condition that our technology will help to bring about, has always made sense, not just in the world of iPods, but also in the world of teapots. The interest of the myth is that it is a unique reservoir of human memory that enables us to compare our current answers to the main question of ethics with the answers given by the previous generations, whose world did not resemble ours and whose thinking has been shaped by a different set of concepts. If our moral thinking makes room for the idea that that we have been born into a world in which human thought lasted for centuries, then the danger will be a little smaller, of us operating a discontinuous change that will terminate the history of mankind and erase the records of human memory.

I. Prometheus

Story

Hesiod's original myth was written in the 7th century BC. At the dawn of humanity, Men used to be provided for all their needs by the Gods. The Gods, though, wanted to keep their powers for themselves, especially the mastery of fire. Prometheus, the wisest of all Titans[4], pitied men, stole fire from the Gods of Mount Olympus and gave it to men. Through this act, he is variously credited with bringing to men craftsmanship, knowledge, enlightenment, and labour. But Zeus, the King of Olympian Gods, ordered Prometheus to be chained to a rock and have his liver eaten out every day by an eagle. Then every night his liver would grow back. This cycle would continue for eternity.

In the 5th century, Plato in his dialogue *Protagoras* writes another version of the myth, in which he emphasizes the role of Prometheus' brother, Epimetheus. The name *Epimetheus* means "with afterthought", whereas *Prometheus* means "with forethought". Epimetheus was as clumsy and unprepared as Prometheus was wise and provident.

[4] The Titans were the primordial deities who reigned before the Olympians, according to Greek mythology. The Titans were overthrown by Zeus and imprisoned in the Tartar.

The Gods were about to create animals. They asked both brothers to equip creatures with proper qualities, so that each can find his food, protect himself from heat, cold, and predators:

> Epimetheus said to Prometheus: "Let me distribute, and you inspect." This was agreed, and Epimetheus made the distribution. There were some to whom he gave strength without swiftness, while he equipped the weaker with swiftness; some he armed, and others he left unarmed; and devised for the latter some other means of preservation, making some large, and having their size as a protection, and others small, whose nature was to fly in the air or burrow in the ground; this was to be their way of escape.[5]

Thus Epimetheus distributed all the qualities he had received from the Gods among animal species with the exception of Man. Prometheus came for inspection and found out that among all animals...

> ... Man alone was naked and shoeless, and had neither bed nor arms of defence. The appointed hour was approaching when man in his turn was to go forth into the light of day; and Prometheus, not knowing how he could devise his salvation, stole the mechanical arts of Hephaestus and Athena, and fire with them (they could neither have been acquired nor used without fire), and gave them to man. Thus man had the wisdom necessary to the support of life, but political wisdom he had not; for that was in the keeping of Zeus, and the power of Prometheus did not extend to entering into the citadel of heaven, where Zeus dwelt, who moreover had terrible sentinels. [...] And in this way man was supplied with the means of life.

Analysis

5.1. The Ambivalence of Technology

In Plato's *Protagoras*, the original state of Man before civilization is purely negative: man is almost less than an animal, because he has no natural means of survival. For Hesiod, on the contrary, before Prometheus's theft of fire, all man's needs were directly satisfied by nature; the burst of technology into man's life was followed by the appearance of labour, pain and suffering.

Since Antiquity, then, technology has been perceived either as a way to save mankind from its original destitution or, on the contrary, as putting an end to the original Golden Age, when men used to live much longer than today, were provided for all their needs by Nature and lived in harmony with the Gods. These two conflicting images of technology are still with us: think of the reactions to climate change on the one hand (human technology used on an industrial scale is responsible for the devastation of our planet); and of the

[5] Plato, *Protagoras*, transl. B. Jowett, 2005.

promises of nanotechnology on the other (cure cancer, live longer, etc.) Both images are equally far-fetched: one may argue that the existence of technology *per se* is not responsible for climate change, nor can we expect that nanomedicine will put an end to "involuntary death", as transhumanists would have it. Still, it is important to notice that in the foundational narratives of Western world, technology is both a promise of a better, fully human life and a vision of catastrophe. What comes as a surprise is that as long ago as the Antiquity ethical judgment on technology was already split: technology in itself is neither undoubtedly good nor doomed to be bad; in and by itself it may be amoral rather than immoral, i.e., it may be neither good nor bad.

5.2. Technology and politics

If technology in itself is neither good nor bad, why then is it perceived as being good or, at other times, bad? What is the true meaning of this ambivalence? Where does it come from? Plato's text offers one answer: that it stems not from technology itself, but from its relationship to politics. Here politics does not reduce to the clever tricks and cunning speeches of the political leaders of today; it is an art of organizing the *polis*, i.e. men's *living together* in accordance with the principles of justice[6]. When Protagoras tells his version of the myth of Prometheus, he insists on the fact that technological intelligence is separate from political wisdom. Thanks to his technical knowledge, man has learned to produce food and weapons and even invented language. But men don't know how to organize themselves so as to benefit collectively from the instruments they have invented. They lack the "ordering principles of cities and the bonds of friendship and conciliation": these are to be found via political activity.

The upshot is simple: technology creates new instruments, but does not say how they should be used, by whom, for what purpose, and to what extent. Technology gives the means, but not the right ways to employ them.

Who shall decide on how to use technology? Who should be allowed to give an opinion on the right way to distribute and use technical inventions? For Protagoras, who supports democracy, the answer is that everybody should, i.e. all citizens should be allowed to express their view on the uses of technology. This is because all people were endowed by the Gods with a share of "political virtue", which makes their opinion valuable. Plato, who was a supporter of aristocracy, had a different view. For him, political wisdom is the privilege of rulers, for only they possess political virtue. Protagoras's and Plato's positions illustrate two opposing interpretations that can be given to the myth of Prometheus. One can hold that Men are able to identify the good uses of technology and act on these; then technology appears as a bless and the vision of technology is a radiant one. Or one can say that men cannot do it, because they do not have enough political wisdom, i.e., they don't see the consequences of the uses of technology for their common well-being. In that case, the decision should belong to an elite of enlightened and virtuous rulers, for otherwise technology may lead to a catastrophe.

[6] E.g. in Plato's *Statesman*, politics is defined as the art of "weaving" men together into an harmonious whole, much like a weaver entangles threads together to produce fabric. This "weaving" starts by distinguishing men into different groups by the division of labour (peasants, artisans, warriors, etc.).

Today these two visions correspond to the alternative between a technocratic regime, where experts and politicians alone decide on the proper uses of technology, and a "technological democracy", whereby all citizens have their direct say on the right use of technology and its political consequences. Plato's myth demonstrates that technological democracy rests not so much on the assumption that all citizens could get access to the knowledge of all technologies (which is obviously impossible, as Protagoras himself admits), but more importantly on the idea that all men have an equal share of political virtue. A contemporary interpretation of the alternative to this view is not necessarily political aristocracy, but a modern democracy in which the sphere of technology is ruled by the technocrats, i.e., a specially trained and selected elite.

5.3. Technology and *hubris*

In Greek mythology and culture, *hubris* is an attempt – always in vain – to resemble divinity. In the European languages this term is often translated as excess, immoderation, démesure. Prometheus is a major figure of *hubris*: by giving men fire, which symbolizes technology in the myth, he offers them a privilege of the Gods. Hence the Promethean idea that technology offers Man a way to escape his finite, mortal condition, and to divinize himself. This idea inspires, sometimes very literally, many of today's technophile movements.

In the Greek world, *hubris* is always punished: the Gods will not allow a mortal to be like them and Man must stay within his human limits. Prometheus is punished for his *hubris* by being chained to a rock, having his liver eaten by an eagle, symbol of Zeus, for eternity; and Icarus, who flew to the Sun, drowns at sea.

Although we have ceased to believe in Olympian Gods a long time ago, the fear of technological *hubris* is still with us. Scientists are often accused of "playing God" or "going too far". What is "too far"? Clearly, this means going beyond a certain limit that humans must not cross, which is a typical case of hubris. The role that hubris still plays in our ethical judgment should be surprising, because it is a religious notion and our society does not rest on religion anymore. Since the Enlightenment, the Christian God has disappeared as a mandatory reference for justifying the social order. Constitutions decided by men replaced the "divine right" as the only source of political authority. But, in a certain way, we have only replaced the figure of a transcendent, unique God by something else: by Man himself as the ruler of his proximate environment. A multitude of limits have stayed, whose origin can now vary. Human finitude, being one such source, leads to hubris for that who wishes to cross this finite border. One may say that we again live in a polytheistic world, where gods are no more transcendent, but the corresponding ethical notions, hubris including, are present nevertheless.

There is a reading of history that allows one to say that science has become "sacred" in our societies. It explains why our attitude to science, whether fear or veneration, is always passionate. Only a few remain unmoved when presented with cutting-edge scientific observations or new inventions. Descartes said in the 17th century that through science Man becomes "master and possessor of nature". In the 19th century, Science appeared with trains and electricity as a perfect tool for Man's mastery of his own destiny. Positivists believed that

Science will help man to transform and even to transcend nature. Cosmists like Tsiolkovsky spoke very seriously of immortality. Our attitude to science has become a faith, almost a religious one, with scientists and engineers in the role of priests. But every faith has its iconoclasts, i.e. those who rebel against sacred symbols in the name of a "true" faith. Attacks on science and technology by technophobes and luddites are the attacks on the symbols of this modern faith. People who take part in such attacks are generally those who wish to restore a "purer" human condition: often because they view Man as living in harmony with his environment, not as transforming it; or because they believe Man is a fundamentally equalitarian being, while they view technology as creating inequalities. Science has become a stake in political battles because, *prima facie*, it has become an issue in the controversy between two visions of the sacred.

Clearly the scientist alone cannot influence this battle. He is not a virtuous hero like Prometheus, nor does he openly assume his society-imposed position of the priest. But he participates in the controversy whether he wants it or not. Political activists on all sides use the scientist in their arguments, so he better be clear about the real stakes expressed in the passionate declarations on science and technologies. When people accuse scientists of "playing God" or "messing with Nature", this is not only a reaction to the novelty of scientific research or an expression of moral concerns. It is also a political claim, which the scientist may agree or disagree with. In any event, he must not remain silent and let others exploit himself politically.

II. The Golem of Jeremiah

Story

Legends of artificial beings, which in the Jewish tradition are called golems, also belong to the class of narratives that enrich our contemporary thinking about technology. We are not interested here in the actual methods that the rabbis used – according to various Jewish books – in order to create animals or men. Only a few aspects of the large mystical and philosophical literature are of direct relevance to the ethics of technology[7].

Golem legends start with some rabbis making calves for the Saturday meal and other rabbis trying to make a real man but failing to achieve it, as we see, for example, when we learn that their creations could not talk. Then comes one legend that is of interest to us here. It dates back to the late Middle Ages and was recorded among the Jewish population of the Rhine region. For these people, the prophet Jeremiah was the standard of a wise man. Indeed, according to the legend, Jeremiah was so knowledgeable and wise that he could do many things that other people couldn't. And he also felt an urge to demonstrate his high level of knowledge practically: because he could do it, he would do it just as a proof of his skill. Jeremiah decided to perform the ultimate task: to make an artificial man, and so high was his wisdom that he succeeded. His creature – the golem of Jeremiah – was perfect: he could walk, talk, etc., and no one could tell the difference between the golem and an ordinary man. Immediately after he rose from the dust from which he was made, the golem

[7] See M. Idel, *Golem*, Albany: State University of New York Press, 1990; H. Atlan, *Les Etincelles de hasard*, Paris, Seuil, 1999.

spoke to Jeremiah. "What have you done?" he asked. "If now a man meets another man in the street, he will not know if that is a man born from a mother and a father or a man made by another man." Jeremiah was puzzled, for in all his wisdom he had not thought about this question. "What shall I do?" he asked the golem, who immediately answered: "Undo me." And this is what Jeremiah did, thereby ending the story of his golem.

Analysis

5.4. Knowing is making

One can point to various similarities as well as differences between Jeremiah and the modern scientist. Like Jeremiah, the scientist only considers true the knowledge that he can transform into a know-how and apply practically. Alternatively, the scientist may have chosen to know that he possesses some knowledge without applying it in the real world; but, like for Jeremiah, this option is untenable for a real scientist. Practical implementation in the form of technological know-how has become the ultimate criterion of truth. Note that moral choice or the question of good and evil are absent from the decision to submit one's knowledge to practical test. The test is performed for epistemic reasons that have no goal other than the establishment of true knowledge. Words, such as utility, benefit of innovation, solution to the problems of society, or the well-being of the public, are completely absent from Jeremiah's thinking. Jeremiah behaves as if he has had no ethical or political position.

This is quite stunning, and the golem points it out to Jeremiah at no delay. Jeremiah, who is supposedly the wisest man, has not thought about the moral question that the golem puts to him, which has to do with a blurring of categorical distinctions (man born from another man or made by another man) that are at the foundation of the existing system of moral values. A man who is born possesses inherent dignity, while a man who has been made has none. He is a machine, an artificial being brought about with a particular goal in mind that his creator had first formed and then realized in the form of a human machine. Natural human life does not possess a specific goal, but golem's life does, and this goal depends on the intention of the man who created this golem. Jeremiah's golem rightly points at this enormous ethical shift as a *problem*, and Jeremiah, who is instantly terrified because he hadn't thought about this evident problem before, appears to be lost. Why hadn't he thought about it before? And what shall he do now?

There can be many answers, and myths usually admit a variety of different interpretations. Perhaps, contrary to the Jewish tradition, Jeremiah's omitting such an obvious ethical question is a sign that he had sinned indeed, as Greek or Christian morality would have it. For a Greek, Jeremiah committed an act of hubris as he tried to put himself on a par with gods, to whom the right to make new living creatures belongs. For a Christian, Jeremiah is to be condemned for he tried to "play God", an accusation often used today in public debates and not so different from the Greek notion of hubris. For the original Jewish thought, though, such considerations were totally alien. Perhaps Jeremiah did not think about the ethical problem simply because he was carried away by the very process of creation, so that no other thought could enter his mind. The scientist knows this feeling of being enticed away from the entire world, when a long-prepared experiment finally starts working as he had hoped for. This is akin to the Greek notion of enthusiasm or, to put it simply, this is what we all experience when we say, "I've totally given myself to this work. I couldn't think about anything else before it was done." For a Christian, this phenomenon is a manifestation of the limited nature of human being, whose difference from God lies precisely in not being able to contemplate too many things at once. Jeremiah's bounded apprehension introduced a dose of evil in his work, and the golem, who is Jeremiah's and not God's creation, cannot be free of evil because Jeremiah as a creator is only a limited one.

Jeremiah has been very fortunate in that he had available to himself an "easy" solution of the ethical problem that he had created. The golem could be undone, which is what Jeremiah did. Without asking how he proceeded to undo the golem, let us note that this solution is unlikely to be available to whatever technological innovation we introduce into the world today. This is because the agent who does so (the "we") is a complex network of research organizations, technological firms and policy makers, and simple reversibility is not an option for "us" at all. To take just one example, the introduction of the mobile phone surely cannot be played back or undone. So, in a sense, we are doomed to fare better than Jeremiah and to find other ways to solve ethical problems created by our technology.

5.5. Intermediate status of technical objects

Why does the legend use the verb "to undo" rather than "to kill"? Because in the Jewish tradition a golem is not a living being. This distinction allows a comparison with an ethical question that is often raised in nanobiotechnology and synthetic biology. Jeremiah's golem is perfect, which means that there is no difference that one would be able to perceive between him and a man who was born naturally. On the basis of what, then, can we judge if the golem is a man like any other or not? There is a mystical answer given by the exegetic tradition: that a "God's sparkle" is absent, which is necessary for life. This mystical answer should be interpreted in one or another way in the language which is familiar to us. Some say that the conception and the birth of a living being must involve randomness or chance in an essential way and that it is this chance that makes a man alive. The function of the chancy component in man's birth is to remove the intentionality that could have been put into the creation of a new man by his human designer. In a way, the element of randomness is what endows a new being with autonomy. For example, whatever the intention of the parents, their child is never a mere machine to fulfill that intention. Being intentional is equal to being artificial, so removing the goal-orientedness from one's life through an intervention of chance makes the person truly alive.

Another interpretation of the undo/kill distinction takes the side of the golem. Like the products of synthetic biology, the golem is fully undistinguishable from a living being, so he is – in a sense – a living being. This means that he is neither inert nor living, but occupies a third, intermediary stage, which does not exist in the natural order of things. Our language and our experience have no word or room for such an intermediary status, but it can still exist if we decide to introduce it in our society. For example, certain representatives of the Jewish tradition argued about the legal status of the golem and concluded that he can take part in a prayer very much like a man who was born, although he does not possess some other rights and obligations of a born man. Such an introduction of the intermediary legal status: not quite a living being but not a mere machine either, is an attractive option for solving the problem of social and juridical place of genetically modified foods, genetically synthesized bacteria and other engineered organisms, including robots. The society will inevitably organize a debate on the place in the social structure that must be given to technological artifacts, and the more we learn from similar debates in myths and legends, the better we'll do for our own generation of emerging technologies.

III. Frankenstein, or the Modern Prometheus

Story

Mary Shelley's 1831 novel *Frankenstein, or the Modern Prometheus*, is an adaptation of the myth of Prometheus to the context of industrial revolution.

Victor Frankenstein is a bright and passionate scientist fascinated with human nature and the source of life. After years of fanatic studying, he is able to "bestow animation upon lifeless matter", so he creates a monster of gigantic proportion from the assembled body parts taken from graveyards, slaughterhouses and dissecting rooms. As soon as the creature opens his eyes, however, the beauty of Frankenstein's dream vanishes: the creature is horrible. Frightened by his own creation, Frankenstein flees his laboratory.

When he returns the next day the monster has escaped. He killed Frankenstein's younger brother, whose body is found in the woods. Frankenstein knows who is the murderer but says nothing, for he thinks that no one would believe him. A family friend, Justine, is accused of the murder, found guilty and hanged. Frankenstein leaves his parents' house and goes wandering in the Alps. There he meets his creation who tells him his story.

After leaving Frankenstein's laboratory, the monster went to the village where he was insulted and attacked. He found refuge in a hovel next to small house inhabited by an old blind man and his two children. By observing the family and by reading their books, the monster learned to speak and to read. He sympathized with the family who struggled to survive and helped them in their work without making himself noticed. Longing for kindness and protection, one day he decided to meet his hosts. He got into a pleasant conversation with the blind man but the reaction of his children was quite different. They beat him and forced him to flee the house. Completely disillusioned, the monster was filled with rage and decided to find his creator. By chance he met Frankenstein's younger brother in the forest and choked him. Then he placed his portrait in the lap of a sleeping young girl, Justine.

The monster asks Frankenstein to create a female to accompany him. If Frankenstein would agree, the monster and his bride would stay away from other people and live in the wild. At first Frankenstein says yes. Then, when he gets quite far in the work on his second creation, he becomes afraid that they may hate each other or that they may produce a whole race of monsters. When the monster visits to check on the progress, Frankenstein destroys his work. The monster swears revenge and promises to be with Frankenstein on his wedding night. The following day, the monster kills Henry Clerval, Frankenstein's best friend. Frankenstein is in despair but, after some time, he marries Elizabeth. Remembering the monster's threat, Frankenstein is convinced that he would be killed that night. The monster, however, kills Elizabeth instead. Frankenstein's father dies after hearing the news. Frankenstein can now only think of revenge. He follows the monster everywhere, ending in the Arctic region where he is taken aboard the ship of the novel's narrator, Walton.

Frankenstein's health deteriorates and he dies. Just after his death, Walton finds the monster hanging over his body. The monster now hates himself because of his crimes. Since his creator is dead, he decides it is time that he too should leave the world. After stating that he will build a funeral pile for himself, he disappears on his ice-raft into the darkness.

Analysis

5.6. Making life from death

Frankenstein's story is a pessimistic version of Plato's Prometheus. It illustrates in a tragic fashion how technology might escape its own creator and turn against himself. The tragedy has its origin in *hubris,* which is now incorporated in the Christian morality of the European society: by creating a man, Victor Frankenstein claims for himself the privilege of the Christian God. His vain attempt to equal God leads to his punishment: there is no place on Earth for his creature, who rebels against his creator and finally kills him.

Shelley's own point of view is not a strictly religious one. If golems were made of dust, Frankenstein creates life out of human body parts. Here Shelley vividly imposes a repulsion for Frankenstein's transgression of the barrier between the dead and the living. For the modern reader, hubris here takes a dimension of transgressing a purportedly natural rule: life can only come from life.

Of course, nature does not really impose this rule: all forms of life feed themselves from dead or non-living matter. Still there is a powerful common belief in a sharp distinction between what is alive and what is not. Human society enacts this distinction by burying corpses far away from the territory of the living, as many animals also do. This suggests a common anthropological root of the live vs. dead distinction. But what about body transplants taken on deceased bodies? Organs transplanted from dead bodies to save lives as well as the use of human corpses for scientific purposes already pose ethical questions: e.g., for how long after death is the human body still to be considered inviolable? Surely synthetic biology, if it succeeds, will pose more similar questions: assembling parts to create a living organism, be it a cell, will likely lead to the question of origin of these parts. Do they originally come from the matter that has been a part of another living being? Do they come from earth, e.g. oil or soil, or from chemical synthesis starting with atoms? Can we assemble a living organism from recycled waste? And so forth.

5.7. Blindness to ethical consequences

It is important that, like Jeremiah, Frankenstein is *thoughtless*: the need to consider the consequences of his actions never occurs to him in his first act of creation and only later during the second process, when he was making a partner for the first creature. Frankenstein, the modern Prometheus, rather has the temper of Epimetheus: he is completely unable to foresee the ethical consequences of his actions. Frankenstein is a caricature of the passionate scientist forgetting the entire world for the sake of an important discovery he's working on. Once the ethical consequences suddenly appear to him, it is too late and he is totally powerless: he deserts his laboratory. Even after he had posed the ethical question and seen catastrophic moral consequences, Frankenstein continued to believe that a remedy could be made by yet more technology: when the monster asks him to create a second being, i.e. his partner, Frankenstein first complies, before realizing that by doing so he will only multiply catastrophes.

Unlike in the Golem legend, Frankenstein cannot undo what he has done: the monster escapes his power altogether. This is modern reality as opposed to the medieval vision of Jeremiah's powers. Today scientific inventions live a life of their own as soon as they have been unleashed. Technological feats are so appealing to other scientists, research institutions and industry that there is no stepping back once they have been achieved and no "un-inventing" them.

The only way to avoid possible catastrophic consequences would be to foresee them before the invention goes public. The question, though, is what the word "possible" means and whether this meaning provides a sufficient ground for action under uncertainty. This relates to the debate on the Precautionary Principle. Still, creative work itself is only complete at the end of the process and it is usually impossible to foretell where the scientist will end up. Furthermore, the scientist's own creative work is not the end of the chain: enter the technologist, the designer, the industrialist and the salesman. Their action, too, participates in the production of the final result of technological innovation, whether revolutionary or catastrophic. Hence some form of moral responsibility must be put on the original inventor as well as the entire body of institutions that took part in the production and mass distribution of technological devices. The inventor brings innovation to the world, but what this innovation does to us is not due exclusively to his intention. It is a complex product of the entire industrial chain of modern society.

5.8. The social status of artefacts

Frankenstein's artificial creature has no home and no place in the human society. At the same time, he is very much like a human: he can speak, be benevolent when he helps the family of the blind man who unknowingly shelters him for some time; he is curious and wants to learn, as shown by the fact that he teaches himself to read; he needs company and has desires, as shown by his request to Frankenstein to create a female monster; and he also has affects such as revenge, jealously and even self-consciousness, when he finally feels guilty for his crimes. The monster wants to be human and to be included in human society. But what really sets him apart from human society is his inhuman face that induces fear.

This leads to the question of how we recognize a human being and what should be demanded of an man-made artefact in order to consider it a part of human society. There already exist robots able to perform intelligent tasks. Are they human because their inventors or their users enter into emotional relations with them, feeling sympathy, intimacy, friendship and perhaps even love for a robot? If a robot is caught in the web of human emotions, does her own behaviour become emotional or interpreted as emotional? Does she become human?

Shelley's answer seems paradoxical, for her monster shares many typically human features (emotions, feeling of guilt, desire for a social life) but is refused any status in the society. Yet facial features are fundamental in our brain recognition systems and it is not surprising that a creature with inhuman looks, however humanly he behaves, be rejected by society. However, it is clear that as intelligent, autonomous artefacts like robots grow in number and become ubiquitous, the need will emerge to give these artefacts an intermediary social and legal status, maybe similar to that of pets: not completely human, but not quite like material goods either. Robots could perhaps be sold but not disposed of as a piece of furniture. They may even have dignity and rights.

IV. A positive Prometheus?

Paul Manship, then one of the most famous American artists, received a commission for a sculpture of Prometheus in January 1933 from John Rockefeller Jr. It was a part of the original design for an ambitious Art Deco urban centre to be built. Manship's Prometheus is seen as he descends from the Mount Olympus with fire he has just stolen from the Gods, hovering over the ring of the zodiac. The zodiac symbolizes the entire Universe. The sculpture gives a sense of strength and seems to defy gravity, just like John Rockefeller Jr. wished to defy the collapse of American economy by building his gigantic Center. Carved in the red granite wall behind the figure is a quote from Aeschylus: "Prometheus, Teacher in Every Art, Brought the Fire That Hath Proved to Mortals a Means to Mighty Ends".

Analysis

This sculpture is typical of the positive interpretation of the myth of Prometheus by modern engineering. More than the scientist, here Prometheus personifies the American engineer and the Industrialist (Rockefeller), coming as heroes to the rescue of the country in the midst of an economic crisis. The engineer, whose profession was clearly established at the end of the 19th century, was then a quite recent figure: even more than the lab scientist he was apt to make use of technological progress for creating tools for the public. His contribution is shown as unambiguously beneficial to mankind: by imagining, designing and making goods and services, to be exposed in Rockefeller's center according to the original plan, the engineer is the shown as the main vector of civilization and of well-being. Note that the political question of how these "means" are to be used is not even asked: the means necessarily lead to "mighty ends", beneficial to all.

V. Pandora's Box

> "The story of Pandora's box is *a familiar one*: a temptingly closed box, once opened, releases the whole range of human evils. People spoke in these terms about nanotechnology's uncertainty, their sense of the

hubris of meddling with things that should be left alone, and of danger and, ultimately, disaster."[8]

Story

The myth of Pandora is an appendix to that of Prometheus. After the theft of fire from Hephaestus, Zeus ordered Hephaestus to create the first woman, Pandora. She was made of clay and given many seductive gifts by Athena, Aphrodite and Hermes (hence her name, meaning "gifted from all"). Zeus gave Pandora as a gift to Epimetheus. Prometheus warned his brother not to accept any gifts from Zeus, but Epimetheus did not comply and married Pandora. Pandora was sent to Epimetheus with a large jar, which she had been instructed to keep closed in any circumstances. Unable to resist curiosity, she opened the jar. At once all the woes and illnesses that had been unknown to men escaped from the jar. Thus Zeus punished mankind for Prometheus's theft of fire. Still, at the very bottom of her jar one item remained, kept there because Pandora put the lid back on the jar. It was hope.

Analysis

5.9. Technology and desire

When one wants to condemn a new technology, one often says that it "opens Pandora's box". Indeed, the lesson of the myth is that when we invent a new technology, we don't know what we wish for[9]. Pandora and her jar symbolize magical gifts that promise to free us from suffering and need[10]. Such promises provoke desire. Today, new technologies are often celebrated as putting an end to suffering or need: nanomedicine will cure cancer, radio-frequency chips will make the world safer. Such language too provokes desire. But the desire for what? With new, yet unseen technology we do not really know what it is that we wish for, until we gain some experience of using this technology: we only know what benefits to expect from it, but until we put the technology to use we do not know the side effects, which may end up being greater than the benefit.

For the German philosopher Günther Anders as well as for his one-time partner Hannah Arendt, technology has reached a point where our capacity to *make* bypasses our capacity to *imagine* and *rationally conceive* what will come out of our own inventions: we literally do not know what we make. Not only is the image of "black box" adequate to depict laypeople's attitude towards technology. Black box also symbolizes our inability to predict the effects of technology. Still, one may ask: shall we blame technology for the evil that occurred? Or shall we blame our own desire?

First, it is not clear that human desire – the public's desire to have new tools – would be the motor of technological innovation. Rather, science seems to follow its own logic, where most of new inventions grow as developments of the already existing ones. Graphene, for example, would not have been discovered without the work on nanotubes; or much of

[8] Matthew Kearnes & Sarah Davies, *Nanotechnology, Life and Lay-Ethics: A case study on 'upstream' public engagement*, op. cit.
[9] Jean-Pierre Dupuy draws this lesson.
[10] Pandora's "box" really was a large jar which the Greeks called a *pithos*. Such jars were used to store wine, oil or grain, in a society where food was scarce.

nanotechnology would not have come about without the unintended discovery of the STM's capacity to move atoms one by one. New technologies open new vistas and allow the public to dream of new tools; when turned into real objects by the scientists, these tools may in turn spur new desires. But the scientist working on a technological innovation is usually motivated by a purely technological or scientific question: to him, research appears as a problem-solving process not necessarily oriented toward a final goal of increasing well-being or obtaining pleasure. It is only from the point of the view of the public that technological objects respond to desires. So the desire for technological artefacts is not a motor of scientific research, but research and innovation produce artefacts that may, upon leaving the sphere of technological invention, spur desire in the public.

Second, Pandora's myth alerts us not so much to the ambivalence of technology as to the boundlessness of desire. We may desire things we do not possess for different reasons. Sometimes we suffer from the absence of material goods and wish they materialized in our world. Sometimes we desire what another person possesses just because of the mere fact that she possesses it: this is mimetic desire. Desire perpetually renews itself and is never complete. In a sense, it is a part of the human condition never to be fully satisfied with this very condition. Finite human beings possess boundless desire, and for Pandora's myth this is the source of evil. We are all similar to Pandora, male or female alike, when we yield to our desire without knowing the moral consequences of this act.

For the very reason that its applications are still mostly unknown, every new technology is prone to provide a realistic or a fictional basis for a great number of desires. In this sense technology falls prey to infinitely many desires. But is technology merely the jar, not its contents? On one view, merely bringing the jar along was for Pandora an unavoidable condemnation, for if Prometheus had alerted his brother not to accept *any* gift from Zeus, it means that the gift would prove fatal whatever its precise content had been. So Epimetheus's desire was the source of evil and technology only served to help this evil come about. On another view, the jar serves as a receptacle for multiple desires, which may later prove to be evil, but in itself it is not evil. It all depends on the *use* that is made of it[11]. The question posed by the myth of Pandora's jar therefore is not "how to limit technology?", but rather: how to limit the desires spurred by technology? This is an ethical problem of first importance.

VI. Daedalus

Story

"Daedalus" means "cunning worker". He is a mythical patron of craftsmen, technicians and engineers.

According to Plato, Daedalus invented sculpture and made moving images of men. First he made a wooden cow in which Pasiphae, the wife of king Minos of Crete, could hide herself and mate with a bull. Then he became famous as the builder of the Labyrinth, which

[11] One can imagine many examples here. Ritalin, a drug developed to treat Attention Deficit Disorder, was subsequently widely used by healthy students who wanted to concentrate better when preparing for exams, without caring much for possible secondary and long-term effects.

Minos ordered as a place to keep the Minautor, a monster part man part bull who was born from the use of Daedalus's first invention. The Labyrinth was such an astute construction that Daedalus himself could hardly find the way out.

After the completion of the Labyrinth Minos was taken by the fear that he might spread the word about the Minotaur. The king imprisoned his architect in a tower on a peninsula far out in the sea. Daedalus set to escape by air: he has made wings for himself and his young son Icarus. In doing so, he tied the feathers together, from smallest to largest, so as to form a shape with increasing surface. The larger ones he secured with thread and the smaller with wax. Daedalus equipped his son with another pair of such wings and taught him how to fly.

Daedalus warned Icarus not to fly too high, because the heat of the sun would melt the wax, nor too low, because the sea foam would soak the feathers. While flying, Icarus forgot the warning and began to rise toward the sun. The wax softened and the feathers disentangled, so Icarus fell into the sea and drowned. Upon the death of Icarus, Daedalus bitterly lamented his own arts and called the land near the place where Icarus fell into the ocean Icaria in memory of his child.

Eventually Daedalus arrived safely in Sicily. Meanwhile Minos searched for Daedalus. He travelled from city to city asking the same riddle: he presented a spiral seashell and asked for a string to be run through it. Minos knew that Daedalus alone would find a solution.

Sicilian king Cocalus took Minos' seashell to Daedalus. Daedalus tied the string to an ant which, lured by a drop of honey at one end, walked through the seashell and spread the string all the way through. Then Cocalus brought the seashell back to Minos, who asked to hand Daedalus over to him. Cocalus managed to convince Minos to take a bath first, where Cocalus's daughters killed Minos. In some versions, Daedalus himself poured boiling water on Minos and killed him.

Analysis

5.10. Science, tinkering and rationality

Like Homer's Odysseus, Daedalus is full of *metis*, a Greek word approximately equivalent to cunning, but deprived of any moral judgement. Daedalus finds technical solutions to all sorts of intricate problems, but he does not apply a fixed and premeditated method. He really has *no method*: his talent is for finding *tricks*. When he is imprisoned on a tower with no access to a road, he escapes through the air, helping himself with the only raw material he can access: feathers. When his finger is too thick to go through the shell, he uses what can pass through it: an ant. Daedalus uses imagination more than reason. He regards every new problem as an opportunity for a surprising action, not a set of passive *data* that have to fit in together. The figure of Daedalus embodies an approach to technology that is more akin to tinkering than to abstract thinking or computer simulation.

Is this figure so remote from the contemporary technological and scientific invention? Take the discovery of fullerenes by Smalley, Curl and Kroto in 1985. Kroto was doing microwave spectroscopy of long molecules of carbon and nitrogen present in stellar atmospheres. Smalley had devised a laser apparatus to vaporise atoms of metals and study the clusters they produced; his interest was in semi-conductors, and he conducted his experiments with Curl, an infrared spectroscopist. Kroto discussed Smalley's laser with Curl and thought of applying it on carbon atoms for his research on carbonated stars. The vaporisations of carbon showed that C_{60} was a particularly stable and symmetrical carbon cluster: fullerene had been discovered, in a completely unexpected and unpredictable way. It took a lot of serendipity and the encounter of three imaginative scientists coming from disciplines as different as radioastronomy, spectroscopy and semiconductors.

This is not how science and technology are usually presented. In manuals and science classes, scientific truths seem to unfold smoothly and almost naturally, deriving from each other in a rational and linear order. In fact, the history of science and the actual logic of discovery are full of surprises, discontinuities and U-turns. Order and unity only appear *post factum*, when a scientific discipline has completed its main discoveries and solved its core problems. As much as a finished scientific theory is rational and linear, scientific practice is akin to the "creative disorder"[12] exemplified by Daedalus's cunning rationality. Of course, one should not deny that science also progresses by structuring and uniting the existing body of knowledge; but this is not how radical expansions of knowledge occur. Both rationalities, the "cunning" and the "straight", are complementary and necessary for science.

5.11. Technology and values

Daedalus has many craftsman's qualities but no moral sense. His technical feats do not aim at moral goodness, e.g. making human life better. He is a pure fabricator, an ingenious assembler of material elements into technical devices. In a way, Daedalus is an ultimate machine not unlike the conveyor on which our complex devices such as airplanes are assembled. He serves first Pasiphae's lust for a bull, then Minos's wish to protect Crete from the monster generated by this very lust. He seems motivated uniquely by technical curiosity: Minos knew he would not resist the temptation to solve his sea-shell riddle. Suddenly something human appears in Daedalus when his son Icarus dies from using the technology he has made for him. He seems to have emotions only for his family; he has no political allegiance and belongs to no society, so he can serve all cities and kings equally well – and without any moral sentiment or sense of guilt.

Of course, the modern scientist in the Western world is no Daedalus: like any citizen, he makes moral judgments, has a moral life and is attached to certain ethical and political values. However, exceptions exist, like Abdul Quadeer Khan, the father of Pakistan's atomic bomb, who leaked the technology to other countries. This nuclear scientist might be closer to Daedalus's status *outside* morality.

More than men who happen to invent and develop technology, Daedalus really personifies technology itself. Is technology itself amoral, like Daedalus? Is technology foreign to the world of human values and emotions? It seems that a moral argument can never stop a new technology. But this does not mean that technology is valueless. Technology

[12] Henri Atlan, *Enlightenment to enlightenment, Intercritique of Science and Myth*, New York: SUNY Press, 1993, p. 126-127.

legitimizes itself and asserts its autonomy; its core value seems to be the pursuit of technological progress.

Jacques Ellul writes: "Technology demands a certain number of virtues from man (precision, exactness, seriousness, a realistic attitude, and, over everything else, the virtue of work) and a certain outlook on life (modesty, devotion, cooperation). Technology permits very clear value judgments (what is serious and what is not, what is effective, efficient, useful, etc.). This ethics is built up on these concrete givens [i.e., technological artefacts]; for it is primarily an experienced ethics of the behaviour required for the technological system to function well."[13] Ellul thinks that technology has its specific ethics, foreign to the usual ethics of humans and of social life. This is in line with Daedalus's myth. The pursuit of technology's own logic, "make it whenever possible", was the driving force behind Daedalus's achievement. Similarly, scientists adopt the values of the technological system they work in, even when these values are not in harmony with moral values of their fellow citizens.

The clash between human desire and the virtues proper to technology, between Pandora and Daedalus, may lead to the collapse of technology. Icarus's desire to come closer to the sun conflicted with the modesty of Daedalus's intention when he sets limits for the use of his technology, such as the need for a moderate temperature in which the wings could function; an ensuing technological collapse resulted in moral catastrophe, i.e. the death of Icarus.

5.12. The scientist and other humans

The "technological system" is not a machine producing artefacts: it employs men and women, the scientists, who are not mere conveyors creating new technologies, like Daedalus. Still, since they belong in the technological system, scientists cannot but share its values. Is the scientist different from other humans? Does he have different emotions and a different morality? The question seems naïve, but it is part and parcel of the image of Daedalus as the patron of engineers; and it has some appeal to the public.

Ellul's list of virtues is to some extent a caricature: it omits all passions and strong emotions. However, important discoveries, whether scientific or technological, involve and produce passion, sometimes even exaltation. Emotional reaction is a part of the process of scientific discovery and technological design and innovation, even if the layman may believe to the contrary.

There exist several conflicting *cliché* images of the scientist. Sometimes he is pictured as a solitary genius, almost a hero or a wise man, a little eccentric but certainly not cold and mechanical (e.g., the iconic view of Einstein). This image rarely corresponds to the reality since at least the 1950s, when large scientific institutes and corporations took the place of small labs and solitary desks, but it still has some popularity. At the opposite of the spectrum, a more recent figure of "scientific entrepreneur", maybe a myth as powerful as the first one, pictures a scientist who likes to have fun, show off and make money[14]. There is also the figure of the expert, expected to tell the public "what is to be known" on a subject

[13] Jacques Ellul, *The Technological System*, "Autonomy", New York: Continuum Publishing Corp., 1980, p. 338.
[14] Craig Venter, one of the founders of synthetic biology, is often cited as the icon of scientific entrepreneurship.

when, in fact, science really is about investigating the unknown. All of these images are somewhat extreme and none of them resembles the actual scientist in the lab. Nevertheless, they show that the "technological system" is not fully autonomous: it connects with disinterested knowledge (the Ivory Tower scientist), industry (the scientific entrepreneur), and politics (the expert). Each of these spheres have specific virtues and values, with which the technological system must compromise.

Science and technology still possess a special ethics, an honest scientist's set of principles of professional conduct, which is strongly determined by the nature of knowledge that is pursued and of the objects that are produced. This ethics does not prohibit emotions or desires, but at the same time it is structured by specific values and it accounts for the special position of the scientist in the society (see Part IV and Glossary).

VII. The Matrix

Film

The Matrix is a 1999 American science-fiction film written and directed by Larry and Andy Wachowsky. It received five Academy Awards and has been watched by millions of people in the world.

The film is a fable on computer-generated simulation. In what seems to be an American buzzing city of 1999, computer programmer Thomas Anderson leads a secret life as a hacker under the alias "Neo" and wishes to learn the answer to the question "What is the Matrix?" Cryptic messages appearing on his computer monitor and his encounters with three sinister agents lead him to a group at the head of which stands a mysterious underground hacker Morpheus. Morpheus gives Neo a choice between two pills: red pill to learn the truth on the Matrix, blue pill to return to the world as he knows it. Neo chooses the red pill, and subsequently finds himself in a liquid-filled pod, his body connected by wires and tubes to a vast mechanical tower covered with identical pods. The connections are severed, he is rescued by Morpheus and taken aboard his hovercraft, the Nebuchadnezzar.

Morpheus informs Neo that the year is not 1999, but estimated to be closer to 2199. Humanity is fighting a war against intelligent machines created in the early 21st century. The sky is covered by thick black clouds created by humans in an attempt to cut off the machines' supply of solar energy. The machines responded by using human beings as their energy source, growing countless people in pods and harvesting their bioelectrical energy and body heat. The world which Neo has inhabited since birth is the Matrix, an illusory simulated reality construct of the world as it was in 1999. All humans grown in the human nursery created by the machines are connected to the Matrix, where they live a fake life that keeps them away from attending their true condition of prisoners. Morpheus and his crew belong to a group of free humans who unplug others from the Matrix and recruit them to their resistance. Morpheus believes that Neo is "the One", a man prophesized to end the war through his limitless control over the Matrix.

Neo is trained to become a member of the group. He learns that injuries suffered in the Matrix, while not real, can cause enough trauma that they are fatal: if he is killed in the Matrix, his physical body will also die. He is warned of the presence of Agents, fast and powerful sentient computer programs with the ability to take over the virtual body of anyone still directly connected to the Matrix. Their purpose is to seek out and eliminate any threats to the simulation.

The group enters the Matrix and takes Neo to meet the Oracle, a program who has predicted the eventual emergence of the One. The Oracle tells Neo that he has the gift of manipulating the Matrix, but that he is waiting for something, possibly his next life. From her comments, Neo deduces that he is not the One. She adds that Morpheus believes in Neo so blindly that he will sacrifice his life to save him.

Returning to the hacked telephone line which serves as a safe exit from the Matrix, the group is ambushed by Agents. Morpheus allows himself to be captured so that Neo and the others can escape. They later learn that they were betrayed by the crew-member Cypher; as a reward for his treason, the Agents allowed Cypher to return forever to the fake but pleasurable world of the Matrix and to be plugged back in the human nursery. Neo and Trinity rescue their leader. Neo becomes more confident and familiar with manipulating the Matrix, ultimately dodging bullets fired at him by Agents. Morpheus and Trinity use a subway station telephone to exit the Matrix, but before Neo can leave, he is ambushed by Agent Smith. After a long chase, Neo is finally able to escape, having in the meantime died and revived in the Matrix.

A short epilogue shows Neo back in the Matrix, making a telephone call promising that he will demonstrate to the people imprisoned in the human nursery that "anything is possible". He hangs up the phone and flies into the sky.

Analysis

5.13. The reality of simulation

What is the difference between a computer-simulated experience and a physically real one? The Matrix really is a computer-generated dream, but a dream shared by all. How and why should an experience in the Matrix be less real than an experience in the physical world?

This question is not new. At a time when computer-generated simulation was completely unconceivable, René Descartes asked the following question in his *Metaphysical Meditations* (1640). Suppose that I am fooled by a wicked demon or a misleading almighty God. I believe that I sit in a room writing, when in fact I have no body and am simply dreaming. How should I know it? What criterion do I have for distinguishing my experiences from a well-formed, consistent dream? *The Matrix* replaces Descartes's demon with a simulation program, but the argument is the same.

In the 1980s the American philosopher Hillary Putnam rephrased Descartes's question in modern terms: suppose I am a brain in a vat and that all my sensory inputs actually are generated by a powerful computer wired to my nervous system. Could I notice anything in my experience that would allow me to know that it is not real?

When *Matrix* came out in 1999, "virtual reality" was on everyone's lips. Internet was *the* new technology. It has since become common to talk of one's avatar on the web, to play on-line simulation games with other internet users, and to virtually interact in various ways.

Simulation is everywhere in our everyday life: the letters we type when writing an email simulate hand- or typewritten letters we would have produced with previous technology. Further, simulated reality is moving beyond sight and hearing to include the feeling of movement and spatial position, perhaps even smell and touch.

We believe that we are not living in a scenario like *The Matrix*'s. We can easily tell when we interact with a computer-generated simulation and when we act in the material world. What *The Matrix* or Descartes's and Putnam's philosophical thought experiments ask is: what is the status of simulated experience? Is it real? If so, what sort of reality does it have? Since there is no general criterion to distinguish simulation from reality and a perfect simulation could *in principle* entirely fool us, one must conclude that simulation *is* indeed *real*, however less detailed and phenomenally different from the material reality it may look.

The term "virtual reality" appeared in the late 20th century as an attempt to qualify a special sort of reality produced in a computer simulation. But the very notion of "virtual reality" is not clear: something virtual is only possible and what is only possible is not real; so "virtual reality" is literally a contradiction. The expression tries to point at some third mode of existence, which would be intermediary between the actual (or real) and the merely possible. But *what* is this mode of existence? What is this new logical category, beyond the three traditional notions of existence as possibility, actuality, and necessity?

It seems that the notion of "virtual reality" does not help to capture the reality produced by computer simulation. Back in the 17th century, Bishop Berkeley answered Descartes by claiming that only what is perceived really exists and that we live in a world of pure spirits affected by sensory but immaterial impressions, which for him where directly sent by God to us. Berkeley's world was not so different from that of the humans in the Matrix, who live a life disconnected from their bodies. So *materiality* is the only difference between simulation and physical experience. In a simulation, we access a reality where our body is required at most as a tool (say, by typing letters or pressing a button, you cause something to become visible on a screen). In material reality, on the other hand, our body is not a tool but a *medium*: it is the place in which and through which perception occurs, hence it is somehow present in all our perceptual experience. One cannot perceive anything in material reality without at the same time perceiving, or feeling the presence of, one's body in that reality, whereas one doesn't have to feel one's body in order to perceive a simulated reality: on the contrary, attention to simulated reality requires to forget, to a certain extent, perception of one's own body. To push the point further, it is conceivable that we should not even *need* to use our body in order to navigate in a computer-simulated experience in the near future: research on BMI interfaces is already on the way to making this happen.

What is truly new with computer-simulated reality is not that it is "not really real", for it *is fully real*. What is new is that our body is not present with us when we move about in this sort of reality. In the near future, we may not even need our body to navigate computer-generated reality. Are we on the way to becoming pure spirits, as in Berkeley's philosophy? Transhumanists, with their dream of "uploading" people's mind on a computer, might be closer to reality than we usually think.

5.14. Reality, a matter of choice

The Matrix is science fiction. It overestimates the ontological difference between material and simulated experiences and clings to the idea that only the material world is real. This is a conservative vision that can be criticized. The use of old visions, norms and values in ethical judgment in the face of new technology is typical of people's reluctance to recognize the effects of technological innovation on their daily life and on their conceptions and norms. The evolution of norms is a slow one and the one of values even slower.

An interesting feature in *The Matrix* is that the choice between simulated and physical reality is presented as a *moral* choice: in the "real" world are the good guys, while in the simulation there are wicked machines and the meek humans who accept their condition of prisoners in the human nursery, plus traitors. This divide provides an implicit ethical judgment on simulation technology. This is not quite unexpected of a Hollywood film. The question before Neo when he chooses between the blue and the red pills, which is a crucial moment of the film, is the following: do you prefer to live in a fake world of pleasures or a painful world of the Truth? What is remarkable is that the choice here is moral.

Here *The Matrix* revisits an old philosophical allegory: Plato's cave. In *The Republic*, Socrates asks his disciple Glaucon to imagine a cave where chained prisoners are seated facing a wall. Light enters behind them. Between light and the wall, but unseen to the prisoners, objects move, projecting their shadows on the wall of the cave. Socrates further supposes that the prisoners were born in the cave and have always been chained this way. He asks: will these prisoners not believe that the shadows they see on the wall are the only reality? Will they not believe that all there is in the world is *shadows*? He adds: if one of the prisoners is unchained and taken to the entry of the cave, will he not be blinded by the light of the sun? Will he not suffer and want to get back to his cave? And if, by luck, he has a noble soul and chooses to face the outside reality, gets used to sunlight and then goes back to the cave to teach the other prisoners and to liberate them, will they not reject him, attack him and even kill him? Of course they will, answers Glaucon; for they feel more comfortable in the cave, even though they are chained. Men tend to prefer the comfortable prison of illusion to the dangerous world of truth.

Plato's allegory asks questions on the moral value of truth and illusion. Can illusion be morally good? If illusion brings pleasure, what is the value of this pleasure? The world of commercial computer-generated simulation is often guided by hedonism: simulation is used to create artificial pleasures and keep users in a state akin to addiction. If addiction is bad but we believe that any reality could be acceptable, even if it is an illusion, is being true and being good the same thing?

Computer-generated simulation is a real experience; at the same time it might be false and, consequently, bad in the moral sense. It may portray virtues where there is only an illusion or encourage the idea that action can be performed at no effort.

These questions are relevant with respect to the existing technology of computer-generated simulation. We now have the choice of reality in which we prefer to spend our time. We can choose physical reality or we can opt for the artificial reality of computer games and software. The choice needn't be as dramatic as that of Neo: most of us have no trouble alternating between these two realities. Still, we have actual life experiences in both worlds, and these are governed by different norms. Ethical concatenation of these norms is far from trivial.

The experience produced in a computer simulation is artificial: for the first time in human history, men are able to devise their own experience entirely by their minds. This gives us more freedom of choice, since there is in principle no limit to the realities that can be generated through simulation. At the same time, it puts on us a new responsibility: we are responsible for the reality we live in, which we may share with other human beings. Creating one's individual reality becomes an ethical dilemma at the moment human relations are to be brought into this hedonistic exercise.

Conclusion

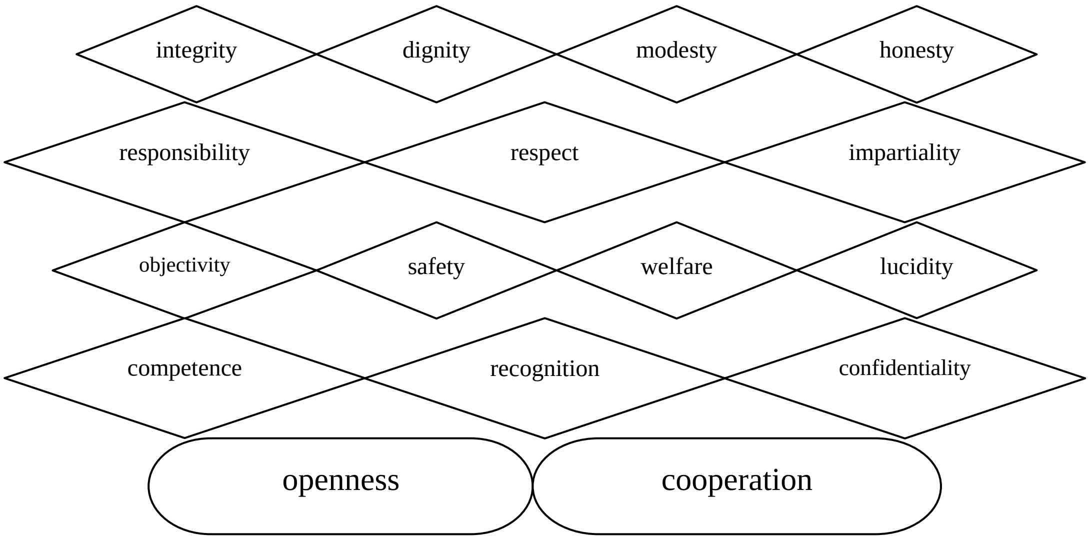
The case of computer-generated simulation is a good example of ethical problems raised by the impact of new technologies on the society. Social norms evolve over time, but ever more slowly that the technological change. When a new technology offers new possible uses, it pushes for the creation of norms for such uses. This process is rapid, not because technology is modern and the society is more conservative, but because technological innovation operates in a normative void: the novel uses of technology usually break away from the old social norms and create a new reality that needs a new normative setting. Technology creates new moral spaces and the society consequently adapts to this configuration and fills these spaces with new norms. In a way, new technology offers the society an arena of *moral experimentation*: it creates new opportunities and new practices and leaves the society the task of inventing respective norms. It also sets against each other moral values that already exist but were not necessarily seen as incompatible. Which one: the dignity of life or the pursuit of knowledge, will win?

As a consequence, it is very hard to predict how society will use new technology. This is why, far from doing abstract scenarios for a hypothetical future, what technology needs is an ongoing ethical assessment of the norms governing the practices of new technologies. What really is under scrutiny in this sort of ethics is not so much technology in its relationship with science, but rather society in its relationship with technology.

Glossary

Principles of the scientist's ethic

- **Integrity.** Integrity is the quality of being upright in character and resistant to corruption. More profoundly, integrity characterizes a person who acts in accordance with his principles. A scientist who promises to follow certain rules displays integrity if he is truly respects his word. Integrity is a necessary condition of moral behaviour.
- **Dignity/Respect.** Dignity is the quality of a human person or of another creature who earns or deserves respect. For Kant, dignity implies that our acts must treat every human being as an end in itself, and not only as a means to some end. The special dignity of the scientist means that he should himself experience respect and must provoke it in others, by acting as implied by his position of a scientist, i.e. by labouring toward progress of knowledge.
- **Impartiality.** To be impartial is to make a fair judgment without giving privilege to one side of the dispute. In the scientific field, impartiality is connected with making a scientific assessment without discrimination by sex, religion, political belief or other



factors, and with communicating in an open and fair manner about the results of one's research.

- **Honesty.** Honesty is the quality of being free from deceit or cheating. For the scientist, honesty implies disinterest and objectivity in the pursuit of knowledge. The scientist must carry out research not for a worldly recompense, financial or other, but for the intrinsic value of knowledge.
- **Competence.** Competence is being able to carry out one's work correctly. This social principle implies the existence of a community, which sets a standard of 'correctness'. In a collective, the principle of competence prescribes to entrust with the power to act the person who is competent with regard to the intended action. For an individual, it

implies that one must evaluate his own competence fairly and avoid action or claims of expertise in matters that lie beyond his area of competence.

- **Safety.** The principle of safety prescribes to act cautiously so as not to endanger one's own safety or the safety of others. It requires from the agent an evaluation of risks before undertaking the action. The concept of safety lies in the foundation of such risk-averse attitudes as prevention and precaution.
- **Objectivity.** Objectivity is the quality of a judgment grounded in independently verifiable facts, contrary to the one based on personal opinion. Objectivity is a condition of scientific argument.
- **Honour/Recognition.** Honour is one's attachment to his or her own reputation. In the scientific field, the principle of honour implies that one must include in the assessment of his reputation, a judgment of recognition or of rejection based on the evaluation of one's work by peers.
- **Cooperation**. The principle of cooperation represents an idea that an appropriate division of labour between different researchers can increase the rate of scientific progress. Knowledge, when shared, is not divided, but grows. Competition between scientists must not endanger this surplus of knowledge due to their cooperation.
- **Welfare.** Welfare is a condition of good life: today one typically associates it with being in good health and enjoying comfortable living and working environments. The modern idea of Progress is based on the assumption that scientific development leads to better technology, which in turn leads to economic and industrial development and enhanced welfare. Reality is more complex than this simplistic idea: welfare depends on many factors other than a steady scientific progress.
- **Confidentiality.** Confidentiality, with respect to human beings, is the quality of knowing how to keep a secret and of being trusted with the secrets of others. In science, particularly medicine, confidentiality implies that if information is collected for the needs of particular research, it should not be used for another purpose. The principle of confidentiality puts limitations on the conditions of storage and accessibility of personal data.
- **Lucidity**. To be lucid is to be aware of the nature and meaning of one's own actions. Lucidity implies that the scientist must not deceive or be deceived about what is real in his or other's work. Even if scientists sometimes willfully exaggerate reality, the principle of lucidity requires from every individual scientist to remain internally aware of the fact of exaggeration. Lucidity also demands that the scientist reflect in an objective manner on the perception of his outreach message by a larger audience. This perception often seems irrational, and the scientist must not extend or impose his own patterns of thinking on others.
- **Openness**. Openness is the quality of somebody who listens to others and erects no barriers at the entrance to her internal world. In a debate, openness requires that one presents her arguments as reflecting only her current state of knowledge and remains open to future amendments. Scientists speak out publicly in the spirit of openness when

they do not appeal to authority or hermetic knowledge. The scientist is also entitled to demand from his interlocutor or opponent to respect the principle of openness.

- **Modesty.** Modesty is the quality of a humble state of mind in the face of complexity and uncertainties. Scientists exercise modesty when they understand the uncertainties in their work and knowledge and when they doubt their professional competence beyond the limit of their direct expertise.
- **Responsibility.** To regard an agent (including oneself) as blameworthy or praiseworthy for his action or inaction means to ascribe him moral responsibility. This notion is different from liability (putting someone at disadvantage), legal responsibility (holding to one's obligation as prescribed by law), or accountability (disposition to render an account of actions of events). Moral responsibility requires the recognition of human freedom and autonomy that we exercise when we make choices. However, contrary to guilt or legal responsibility, moral responsibility does not extend only to the desired or intended consequences of our actions. Scientists are held morally responsible for the future uses of their discoveries even if they have not sought to empower such uses deliberately at the time of discovery. This is rooted in the fact that a human being who exercises his freedom by bringing things or creatures unto the world, is held morally (although not legally) responsible for all future consequences of such action. A pure exercise of human freedom, even without any explicit expression of the agent's free will, suffices to render him morally responsible, as argued by numerous philosophers from Spinoza to Sartre.
- **Collective responsibility.** The philosopher Hannah Arendt emphasized that collective responsibility requires the presence of two conditions: a person must be held responsible for something she has not done, and the reason for her responsibility must be her membership in a group which no voluntary act of hers can dissolve[15]. As distinguished from the legal responsibility of a group or the moral responsibility of an individual, collective responsibility owes its relevance and interest to the political predicament: it is always political, and in the center of political considerations always stands the world, while in the center of moral considerations stands the human self. Contemporary science is organized in large scientific organizations, or hierarchies, to which every individual scientist belongs but none has a command of the whole hierarchy. With respect to this world of "Big Science", collective responsibility is an argument in the debate on R&D policy-making and must not be mistaken for a moral or a legal claim.

Nanoethical terms and notions

- **Ambient Intelligence.** Nanotechnology allows miniaturization of ICT devices and their incorporation virtually everywhere in our environment. Cars are already full of ICT. Tomorrow, refrigerators could place shopping orders when empty; "intelligent dust"

[15] Hannah Arendt, *Collective Responsibility*, in H. Arendt, *Responsibility and Judgment*, Schoken books, New York, 2003.

would be able to analyze the atmosphere and detect the presence of lethal gazes, for military use, but also in case of a terrorist attack.

- **Artificial retina.** Currently under research, artificial retina would enable blind people whose visual nerve remains unimpaired to recover sight, according to a mechanism similar to cochlear implant for hearing (see « cochlear implant »).

- **Big Science.** Starting with the Manhattan Project during WWII, many sectors of science became funded by huge Government programs. Big science involves (a) huge public funding; (b) numerous research teams; (c) big and costly instruments; and (d) collective work and publications rather than individual discoveries. The LHC in Geneva is a contemporary example of Big Science. Big Science as an institutional context for the development of science contrasts with the myth of the solitary genius working with no or little instruments, and with industry or philanthropy funding. Strikingly, the disciplines covered by nanosciences today had mostly been neglected by Big Science from WWII to the 1990s.

- **Bioconservatives.** Bioconservatives oppose technology-induced changes in the natural and social order. They condemn cloning, GMOs, genetic engineering, stem cell research, etc. The US bioconservative movement gathers right-wing religious and social conservatives such as Francis Fukuyama, as well as left-wing environmentalists and technology critics. Its right wing was prominent in the President's council on Bioethics under US President Bush.

- **Cochlear implant.** Developed in the 1980s, cochlear implants are placed in the inner ear and convert sounds into electrical signals which are then transmitted to the auditory nerve. They allow certain deaf patients (those presenting a lesion in the inner ear) to recover hear.

- **Consequentialism.** Ethical theory defending the thesis that an action must be evaluated through its consequences, and not the intention guiding it. Cost-benefits analysis is typical of the consequentialist approach. Consequentialism opposes both deontologism and virtue ethics.

- **Cyborg.** Abbreviation of "cybernetic organism", designating a human augmented by ICT and electronic prostheses.

- **Cryonics.** Transhumanist-supported technology for freezing dead human bodies, to keep them sufficiently unaltered for the hypothetical day Science could become able to resurrect them.

- **Deontologism.** Ethical theory according to which the worth of an action is given by the intention guiding it, whatever its consequences. Typical of the deontological approach is the prescription of action guided by universal principles, such as respect for human dignity. Deontologism opposes both consequentialism and virtue ethics.

- **Hype**. Literally, dishonest and unfaithful advertising. Any strategy exaggerating the benefits of a future research program in order to promote and/or fund it.

- **NBIC convergence.** Convergence of Nanotechnology, Biotechnologies, Information technologies, and Cognitive neurosciences, towards a global human enhancement. NBIC convergence is supported by Transhumanism and the Roco-Branbridge Report (2002).

- **Neurofeedback.** Detecting changes in brainwaves through EEG, neurofeedback make visible an upcoming event that is normally unconscously treated by the brain. The US army develops a helmet with binoculars for soldier to detect a danger that his brain would normally subconsciously notice, but would hardly react to. Neurofeedback is also used to warn epileptic patients of an upcoming crisis.

- **Neuroprosthesis.** It that can be an artificial arm, a robot or a computer, directly controlled by the motor cortex of the subject. When the subject wants to activate the prosthesis in a certain way, she thinks of a particular movement, thus generating an impulse in her motor cortex. The impulse is detected either by an implanted electrode or by an external EEG equipment, then modified by a software and sent as a command to the prosthesis. Neuroprosthetics is used for amputees and locked-in syndrome (total paralysis).

- **President's Council on Bioethics (2001-2009).** Appointed by US President Bush, its members recommended strong ethical constraints on biotechnology and research, arguing for instance for a total ban on stem cell research. The Council has been dissolved by President Obama in 2009 and replaced by the President's Commission for the Study of Bioethical Issues.

- **Progress**. The idea of progress is as old as modern science. Science itself is progressing towards Truth, though philosophers disagree on whether this progress is cumulative, proceeding through an accumulation of facts and the continuous extension and refinement of theories, or akin to a trial-and-error process, proceeding through the refutation of erroneous theories and the elaboration of more robust ones. The idea emerged in the 17th century that the progress of knowledge would bring about a progress in the human condition. Thus Descartes held that Science would make us "as masters and possessors of Nature", and that the real goal of Science was not so much Truth as the "commodities of life". In the 18th and 19th centuries, most philosophers and scientists took it for granted that the progress of science would automatically bring about moral and civilization progress. In the 20th century, especially during the two World Wars, technology and science also served to enhance and rationalize means of destruction, and the notion of progress became more and more dubious.

- **Singularity point.** Futurologist prediction that technological progress will reach a point where its development becomes unpredictable and qualitatively different from today, leading to the emergence of a super-human intelligence. The idea is supported by Ray Kurzweil (*The Singularity is Near*, 2005), and has been famously criticized by computer scientist Bill Joy ("Why the future doesn't need us", *Wired Magazine*, 2000).

- **Technoscience**. Invented in the 1970s, the word designates the integration of science with technology. Contemporary science is characterized by the increasing importance of operating on instruments, matter, energy, and signs (in mathematics, e.g.), much like an engineer does. The word "Technoscience" is also used critically, to point an inflexion of science towards practically and industrially useful achievements, and away from the disinterested pursuit of Truth.

- **Risk and Uncertainty**. Uncertainty occurs when a harm can reasonably be predicted, though its probability can not be quantified. Risk equals the level of harm, multiplied by the probability of its occurrence. The Precautionary Principle requires action even in the face of uncertain harms.

- **Roco-Bainbridge Report**. *Converging Technologies for Improving Human Performances* (2002) launched the idea of NBIC convergence and was very instrumental in the promotion and funding of nanosciences by the US National NanoInitiative.

- **Transhumanism.** An international association and intellectual movement promoting the technology-driven evolution of *homo sapiens* towards a superior species.

- **Virtue Ethics.** An ethical theory according to which there is no general rule to evaluate an action. The circumstances of our actions are too diverse and unpredictable for principles to apply easily. Good action depends on the moral character of the agent, that is, his virtues. Example of virtues in science are honesty, justice, competence, openness, curiosity, etc. Virtue ethics opposes both consequentialism and deontologism.