Chapter 6
The Rise of the MASs

Luciano Floridi

Abstract The post-Westphalian Nation State developed by becoming more and more an Information Society. However, in so doing, it progressively made itself less and less *the* main information agent, because what made the Nation State possible and then predominant, as a historical driving force in human politics, namely ICTs, is also what is now making it less central, in the social, political and economic life of humanity across the world. ICTs fluidify the topology of politics. They do not merely enable but actually promote (through management and empowerment) the agile, temporary and timely aggregation, disaggregation and re-aggregation of distributed groups around shared interests across old, rigid boundaries represented by social classes, political parties, ethnicity, language barriers, physical barriers, and so forth. This is generating a new tension between the Nation State, still understood as a major organisational institution, yet no longer monolithic but increasingly morphing into a multiagent system itself, and a variety of equally powerful, indeed sometimes even more politically influential and powerful, non-Statal organisations. Geo-politics is now global and increasingly non-territorial, but the Nation State still defines its identity and political legitimacy in terms of a sovereign territorial unit, as a Country. Such tension calls for a serious exercise in conceptual re-engineering: how should the new informational multiagent systems (MASs) be designed in such a way as to take full advantage of the socio-political progress made so far, while being able to deal successfully with the new global challenges (from the environment to the financial markets) that are undermining the legacy of that very progress? In the lecture, I shall defend an answer to this question in terms of a design of political MAS based on principles borrowed from information ethics.

L. Floridi (✉)
Oxford Internet Institute, University of Oxford,
St Giles 1, OX1 3JS, Oxford, UK
e-mail: luciano.floridi@oii.ox.ac.uk

L. Floridi (ed.), *Protection of Information and the Right to Privacy – A New Equilibrium?*,
Law, Governance and Technology Series 17, DOI 10.1007/978-3-319-05720-0_6,

6.1 From History to Hyperhistory

More people are alive today than ever before in the evolution of humanity. And more of us live longer and better today than ever before. To a large measure, we owe this to our technologies, at least insofar as we develop and use them intelligently, peacefully, and sustainably.

Sometimes, we may forget how much we owe to flakes and wheels, to sparks and ploughs, to engines and satellites. We are reminded of such deep technological debt when we divide human life into prehistory and history. That significant threshold is there to acknowledge that it was the invention and development of information and communication technologies (ICTs) that made all the difference between who we were and who we are. It is only when the lessons learnt by past generations began to evolve in a Lamarckian rather than a Darwinian way that humanity entered into history.

History has lasted 6,000 years, since it began with the invention of writing in the fourth millennium BC. During this relatively short time, ICTs have provided the recording and transmitting infrastructure that made the escalation of other technologies possible. ICTs became mature in the few centuries between Guttenberg and Turing. Today, we are experiencing a radical transformation in our ICTs that could prove equally significant, for we have started drawing a new threshold between history and a new age, which may be aptly called *hyperhistory*. Let me explain.

Prehistory and history work like adverbs: they tell us *how* people live, not *when* or *where*. From this perspective, human societies currently stretch across three ages, as ways of living. According to reports about an unspecified number of uncontacted tribes in the Amazonian region, there are still some societies that live prehistorically, without ICTs or at least without recorded documents. If one day such tribes disappear, the end of the first chapter of our evolutionary book will have been written. The greatest majority of people today still live historically, in societies that rely on ICTs to *record* and *transmit* data of all kinds. In such historical societies, ICTs have not yet overtaken other technologies, especially energy-related ones, in terms of their vital importance. There are then some people around the world who are already living hyperhistorically, in societies or environments where ICTs and their data *processing* capabilities are the necessary condition for the maintenance and any further development of societal welfare, personal well-being, as well as intellectual flourishing. The nature of conflicts provides a sad test for the reliability of this tripartite interpretation of human evolution. Only a society that lives hyperhistorically can be vitally threatened informationally, by a cyber attack. Only those who live by the digit may die by the digit.

To summarise, human evolution may be visualised as a three-stage rocket: in prehistory, there are no ICTs; in history, there are ICTs, they record and transmit data, but human societies depend mainly on other kinds of technologies concerning primary resources and energy; in hyperhistory, there are ICTs, they record, transmit and, above all, process data, and human societies become vitally dependent on them and on information as a fundamental resource. If all this is even approximately correct, humanity's emergence from its historical age represents one of the most

significant steps it has ever taken. It certainly opens up a vast horizon of opportunities as well as challenges and difficulties, all essentially driven by the recording, transmitting, and processing powers of ICTs. From synthetic biochemistry to neuroscience, from the Internet of things to unmanned planetary explorations, from green technologies to new medical treatments, from social media to digital games, from agricultural to financial applications, from economic developments to the energy industry, our activities of discovery, invention, design, control, education, work, socialisation, entertainment, care, security, business and so forth would be not only unfeasible but unthinkable in a purely mechanical, historical context. They have all become hyperhistorical in nature.

6.1.1 Political Apoptosis

In 2011, the total world wealth[1] was calculated to be $ 231 trillion, up from $ 195 trillion in 2010.[2] Since we are almost 7 billion, that was about $ 33,000 per person, or $ 51,000 per adult, as the report indicates. The figures give a clear sense of the level of inequality. In the same year, we spent $ 498 billion on advertisements.[3] Perhaps for the first time, we also spent more on ways to entertain ourselves than on ways to kill each other. The military expenditure in 2010 was $ 1.74 trillion,[4] and that on entertainment and media was expected to be around $ 2 trillion, with digital entertainment and media share growing to 33.9% of all spending by 2015, from 26% in 2011.[5] Meanwhile, we spent $ 6.5 trillion (this is based on 2010 data) on fighting health problems and premature death, much more than the military and the entertainment and media budgets put together. All these trillions were closely linked and often overlapped with the budget for Information and Communication Technologies (ICTs), on which we spent $ 3 trillion in 2010.[6] We can no longer unplug our world from ICTs without turning it off.

Hyperhistory, and the evolution of the infosphere in which we live, are quickly detaching future generations from ours. Of course, this is not to say that there is no continuity, both backwards and forwards. *Backwards*, because it is often the case that the deeper a transformation is, the longer and more widely rooted its causes may be. It is only because many different forces have been building the pressure for a long time that radical changes may happen all of a sudden, perhaps unexpectedly. It is not the last snowflake that breaks the branch of the tree. In our case, it is

[1] This is defined as the value of financial assets plus real assets (mainly housing) owned by individuals, less their debts.

[2] Source: *The Credit Suisse Global Wealth Report 2011*, available online.

[3] Source: *Nielsen Global AdView Pulse Q4 2011*, available online.

[4] Source: Stockholm International Peace Research Institute, *Military Expenditure Database*, available online.

[5] Source: PricewaterhouseCoopers, *Global Entertainment and Media Outlook 2007–2011*, available online.

[6] Source: IDC, *Worldwide IT Spending Patterns: The Worldwide Black Book*, available online.

certainly history that begets hyperhistory. There is no ASCII without the alphabet. *Forwards*, because we should expect historical societies to survive for a long time in the future. Despite globalisation, human societies do not parade uniformly forward, in neat and synchronised steps.

We are witnessing a slow and gradual process of political *apoptosis*. Apoptosis, also known as programmed cell death, is a natural and normal form of self-destruction in which a programmed sequence of events leads to the self-elimination of cells. Apoptosis plays a crucial role in developing and maintaining the health of the body. One may see this as a natural process of renovation. Here, I am using the expression 'political apoptosis' in order to describe the gradual and natural process of renovation of sovereign states[7] as they develop into information societies. Let me explain.

Simplifying and generalising, a quick sketch of the last 400 years of political history in the Western world may look like this. The Peace of Westphalia (1648) meant the end of World War Zero, namely the Thirty Years' War, the Eighty Years' War, and a long period of other conflicts during which European powers, and the parts of the world they controlled, massacred each other for economic, political and religious reasons. Christians brought hell to each other, with staggering violence and unspeakable horrors. The new system that emerged in those years, the so-called *Westphalian order*, saw the coming of maturity of sovereign states and then national states as we still know them today; France, for example. Think of the time between the last chapter of *The Three Musketeers*—when D'Artagnan, Aramis, Porthos, and Athos take part in Cardinal Richelieu's siege of La Rochelle in 1628—and the first chapter of *Twenty Years Later*, when they come together again, under the regency of Queen Anne of Austria (1601–1666) and the ruling of Cardinal Mazarin (1602–1661).

The state did not become a monolithic, single-minded, well-coordinated entity. It was not the sort of beast that Hobbes described in his *Leviathan*, nor the sort of robot that a later, mechanical age would incline us to imagine. But it did rise to the role of the binding power, the system able to keep together and influence all the different agents comprising it, and coordinate their behaviours, as long as they were falling within the scope of its geographical borders. These acquired the metaphorical status of a state's skin. States became the independent agents that played the institutional role in a system of international relations. And the principles of sovereignty (each state has the fundamental right of political self-determination), legal equality (all states are equal), and non-intervention (no state should interfere with the internal affairs of another state) became the foundations of such a system of international relations.

Citizenship had been discussed in terms of biology (your parents, your gender, your age…) since the early city-states of ancient Greece. It became more flexi-

[7] Using standard vocabulary, by nation I refer to a socio-cultural entity comprising people united by language and culture. By state, I refer to a political entity that has a permanent population, a defined territory, a government, and the capacity to enter into relations with other states (Montevideo Convention 1933). The kurds are a typical example of a nation without a state.

ble (types of citizenship) when it was conceptualised in terms of legal status as well. This was the case under the Roman Empire, when acquiring a citizenship—a meaningless idea in purely biological contexts—meant becoming a holder of rights. With the modern state, geography started playing an equally important role, mixing citizenship with language, nationality, ethnicity, and locality. In this sense, the history of the passport is enlightening. As a means to prove the holder's identity, it is acknowledged to be an invention of King Henry V of England (1386–1422), a long time before the Westphalian order took place. However, it was the Westphalian order that transformed the passport into a document that entitles the holder not to travel (because a visa may also be required, for example) or be protected abroad, but to return to the country that issued the passport. The passport became like an elastic band that ties the holder to a geographical point, no matter how far in space and prolonged in time the journey in other lands has been. Such a document became increasingly useful the better that geographical point was defined. Travelling was still quite passport-free in Europe until the First World War. Only then did security pressure and techno-bureaucratic means catch up with the need to disentangle and manage all those elastic bands travelling around by means of a new network, the railway.

Back to the Westphalian order. Now that the physical and legal spaces overlap, they can both be governed by sovereign powers, which exercise control, impose laws, and ensure their respect by means of physical force within the state's borders. Geographical mapping is not just a matter of travelling and doing business, but also an inward-looking question of controlling one's own territory, and an outward-looking question of positioning oneself on the globe. The taxman and the general look at those geographical lines with eyes very different from those of today's users of Expedia. For sovereign states act as agents that can, for example, raise taxes within their borders and contract debts as legal entities (hence our current terminology in terms of 'sovereign bonds', for example, which are bonds issued by national governments in foreign currencies), and of course dispute borders, often violently. Part of the political struggle becomes not just a silent tension between different components of the state as a multiagent system, say the clergy vs. the aristocracy, but an explicitly codified balance between the different agents constituting it. In particular, Montesquieu (1689–1755) suggests the classic division of the state's political powers that we take for granted today: a legislature, an executive, and a judiciary. The state as a multiagent system organises itself as a network of these three 'small worlds', among which only some specific channels of information are allowed. Today, we may call that arrangement Westphalian 2.0.

With the Westphalian order, modern history becomes the age of the state. The state arises as *the* information agent, which legislates on, and at least tries to control, the technological means involved in the information life-cycle, including education, census,[8] taxes, police records, written laws, press, and intelligence. Already most

[8] The Latin word means "estimate". Already the Romans, who were well aware of the importance of information and communication in such a large empire for administrative and taxing purposes, carried out a census every 5 years.

of the adventures in which D'Artagnan is involved are caused by some secret communication.

As the information agent, the state fosters the development of ICTs as a means to exercise and maintain legal force, political power, and social control, especially at times of international conflicts, frequent unrests, and fragile peace. For example, in 1790–1795, during the French Revolution, the French government needed a system of speedy communication to receive intelligence and transmit orders in time to counterbalance the hostile manoeuvres of the allied forces that surrounded France: Britain, the Netherlands, Prussia, Austria, and Spain. To satisfy such need, Claude Chappe (1763–1805) invented the first system of telegraphy (he actually coined the word 'telegraph'). It consisted of mechanical semaphores that could transmit messages across the country in a matter of hours. It became so strategic that when Napolen begun preparations to resume war in Italy, in 1805, he ordered a new extension from Lyon to Milan. At its peak, the *Chappe* telegraph was a network of 534 stations, covering more than 5,000 km (3,106 miles). The reader may remember its crucial appearance in Alexandre Dumas' *The Count of Monte Cristo* (1844) where the Count bribes an operator to send a false message to manipulate the financial market to his own advantage. In fictional as in real life, whoever controls information controls the issuing events.

Through the centuries, the state moves from being conceived as the ultimate guarantor and defender of a *laisser-faire* society to a Bismarckian welfare system, which takes full care of its citizens. In both cases, the state remains the primary collector, producer, and controller of information. However, by fostering the development of ICTs, the state ends undermining its own future as the only, or even the main, information agent. This is the political apoptosis I mentioned above. For in the long run, ICTs contribute to transforming the state in an information society, which makes possible other, sometimes even more powerful information agents, which may determine political decisions and events. And so ICTs help shift the balance against centralised government, in favour of distributed governance and international, global coordination.

The two World Wars are also clashes of sovereign states resisting mutual coordination and inclusion as part of larger multiagent systems. The *Bretton Woods* conference may be interpreted as the event that seals the beginning of the political apoptosis of the state. The gathering in 1944 of 730 delegates from all 44 Allied nations at the Mount Washington Hotel in Bretton Woods, New Hampshire, United States, regulated the international monetary and financial order after the conclusion of Second World War. It saw the birth of the International Bank for Reconstruction and Development (this, together the International Development Association, is now known as the World Bank), of the General Agreement on Tariffs and Trade (GATT, replaced by the World Trade Organisation in 1995), and the International Monetary Fund. In short, Bretton Woods brought about a variety of multiagent systems as supranational or intergovernmental forces involved with the world's political, social, and economic problems. These and similar agents became increasingly powerful and influential, as the emergence of the *Washington Consensus* clearly indicated.

John Williamson[9] coined the expression 'Washington Consensus' in 1989. He used it in order to refer to a set of ten, specific policy recommendations, which, he argued, constituted a standard strategy adopted and promoted by institutions based in Washington, D.C.—such as the US Treasury Department, the International Monetary Fund, and the World Bank—when dealing with countries coping with economic crises. The policies concerned macroeconomic stabilization, economic opening with respect to both trade and investment, and the expansion of market forces within the domestic economy. In the past quarter of a century, the topic has been the subject of intense and lively debate, in terms of correct description and acceptable prescription. Like the theory of a Westphalian doctrine I outlined above, the theory of a Washington Consensus is not devoid of problems. Does the Washington Consensus capture a real historical phenomenon? Does the Washington consensus ever achieve its goals? Is it to be re-interpreted, despite Williamson's quite clear definition, as the imposition of neoliberal policies by Washington-based international financial institutions on troubled countries? These are important questions, but the real point of interest here is not the interpretative, economic, or normative evaluation of the Washington Consensus. Rather, it is the fact that the very idea, even if it remains only an influential idea, captures a significant aspect of our hyperhistorical, post-Westphalian time. The Washington Consensus is a coherent development of Bretton Woods. Both highlight the fact that, after the Second World War, organisations and institutions (not only those in Washington D.C.) that are not states but rather non-governmental multiagent systems, are openly acknowledged to act as major, influential forces on the political and economic scene internationally, dealing with global problems through global policies. The very fact that the Washington Consensus has been accused (no matter whether correctly or not) of disregarding local specificities and global differences reinforces the point that a variety of powerful multiagent systems are now the new sources of policies in the globalised information societies. As a final reminder, let me mention a rather controversial report, entitled *Top 200: The Rise of Corporate Global Power*. It offered some years ago an analysis of corporate agents.[10] Perhaps the most criticised part was a comparison between countries' yearly GDP and companies' yearly sales (revenues or turnover). Despite this potential shortcoming, it still makes for interesting reading. According to the report:

of the 100 largest economies in the world, 51 are (were, in 2000) corporations; only 49 are (were, in 2000) countries.

The criticism remains, but the percentage has probably moved in favour of the number of companies, and what represents a unifying unit of comparison is that both GDP and revenues buy you clout. When multiagent systems of such dimensions take decisions, their effects are deep and global.

Today, we know that global problems—from the environment to the financial crisis, from social justice to intolerant religious fundamentalisms, from peace to health conditions—cannot rely on sovereign states as the only source of a solution

[9] Williamson (1993, pp. 1329–1336).

[10] Anderson and Cavanagh (2000).

because they involve and require global agents. However, there is much uncertainty about the design of the new multiagent systems that may shape humanity's future. Hyperhistorical societies are post-Westphalian, because of the emergence of the sovereign state as the modern, political information agent. They are post-Bretton Woods, because of the emergence of non-state multiagent systems as hyperhistorical players in the global economy and politics. This helps explain why one of the main challenges faced by hyperhistorical societies is how to design the right sort of multiagent systems. These systems should take full advantage of the socio-political progress made in modern history, while dealing successfully with the new global problems, which undermine the legacy of that very progress, in hyperhistory.

6.1.2 A New Informational Order?

The shift from a historical, Westphalian order to a post-Bretton Woods, hyperhistorical predicament in search of a new equilibrium may be explained by many factors. Four are worth highlighting in the context of this book.

First, *power*. ICTs 'democratise' data and the processing/controlling power over them, in the sense that now both tend to reside and multiply in a multitude of repositories and sources. Thus, ICTs can create, enable, and empower a potentially boundless number of non-state agents, from the single individual to associations and groups, from macro-agents, like multinationals, to international, intergovernmental as well as nongovernmental, organisations and supranational institutions. The state is no longer the only, and sometimes not even the main, agent in the political arena that can exercise informational power over other informational agents, in particular over human individuals and groups. The European Commission, for example, recognised the importance of such new agents in the *Cotonou Agreement* between the European Union (EU) and the Africa, Caribbean and Pacific (ACP) countries, by acknowledging the important role exercised by a wide range of nongovernmental development actors, and formally recognising their participation in ACP-EU development cooperation. According to Article 6 of the *Cotonou Agreement*, such non-state actors comprise:

> the private sector; economic and social partners, including trade union organisations; civil society in all its forms, according to national characteristics.

The 'democratisation' brought about by ICTs is generating a new tension between power and force, where power is informational, and exercised through the elaboration and dissemination of norms, whereas force is physical, and exercised when power fails to orient the behaviour of the relevant agents and norms need to be enforced. Note that the more physical goods and even money become information-dependent, the more the informational power exercised by multiagent systems acquires a significant financial aspect.

Second, *geography*. ICTs de-territorialise human experience. They have made regional borders porous or, in some cases, entirely irrelevant. They have also created, and are exponentially expanding, regions of the infosphere where an increasing

number of agents, not necessarily only human, operate and spend more and more time, the onlife experience. Such regions are intrinsically stateless. This is generating a new tension between geo-politics, which is global and non-territorial, and the state, which still defines its identity and political legitimacy in terms of a sovereign territorial unit, as a country.

Third, *organisation*. ICTs fluidify the topology of politics. They do not merely enable but actually promote, through management and empowerment, the agile, temporary and timely aggregation, disaggregation and re-aggregation of distributed groups 'on demand', around shared interests, across old, rigid boundaries, represented by social classes, political parties, ethnicity, language barriers, physical barriers, and so forth. This is generating new tensions between the state, still understood as a major organisational institution, yet no longer rigid but increasingly morphing into a flexible multiagent system itself, and a variety of equally powerful, indeed sometimes even more powerful and politically influential (with respect to the old sovereign state) non-state organisations, the other multiagent systems on the block. Terrorism, for example, is no longer just a problem concerning internal affairs—consider forms of terrorism in the Basque Country, Germany, Italy, or Northern Ireland—but also an international confrontation with a distributed, multiagent system such as Al-Qaeda.

Finally, *democracy*. Changes in power, geography, and organisation, reshape the debate on democracy, the oldest and safest form of power crowdsourcing. We used to think that, ideally, democracy should be a direct and constant involvement of all citizens in the running of their society and its business, their *res publica*. Direct democracy, if feasible, was about how the state could re-organise itself internally, by designing rules and managing the means to promote forms of negotiation, in which citizens could propose and vote on policy initiatives directly and almost in real time. We thought of forms of direct democracy as complementary options for forms of representative democracy. It was going to be a world of 'politics always-on'. The reality is that direct democracy has turned into a mass media-ted democracy, in the ICT sense of new social media. In such digital democracies, distributed groups, temporary and timely aggregated around shared interests, have multiplied and become sources of influence external to the state. Citizens vote for their representatives but can constantly influence them via opinion polls almost in real time. Consensus-building has become a constant concern based on synchronic information.

Because of the factors just analysed—power, geography, organisation, and democracy—the unique position of the historical state as *the* information agent is being undermined from below and overridden from above. Other multiagent systems have the data, the power and sometimes even the force—as in the different cases of the UN, of groups' cyber threats, or of terrorist attacks—the space, and the organisational flexibility to erode the modern state's political clout. They can appropriate some of its authority and, in the long run, make it redundant in contexts where it was once the only or the predominant informational agent. The Greek economic crisis, which began in late 2009, offers a good example. The Greek government and the Greek state had to interact 'above' with the EU, the European Central Bank, the International Monetary Fund, the rating agencies, and so forth. They had to

interact 'below' with the Greek mass media and the people in Syntagma Square, the financial markets and international investors, German public opinion, and so forth. Because the state is less central than in the nineteenth century, countries such as Belgium and Italy may work fine even during long periods without governments or when governed by dysfunctional ones, on 'automatic pilot'.

A much more networked idea of political interactions makes possible a degree of tolerance towards, and indeed feasibility of localisms, separatisms, as well as movements and parties favouring autonomy or independence that would have been unacceptable by modern nation states, which tended to encourage aggregating forms of nationalism but not regionalism. From Padania (Italy) to Catalonia (Spain), from Scotland (Great Britain) to Bavaria (Germany), one is reminded that, in almost any European country, hyperhistorical trends may resemble pre-Westphalian equilibria among a myriad of regions. The long 'list of active separatist movements in Europe' in Wikipedia is both informative and eye opening. Unsurprisingly, the Assembly of European Regions (originally founded as the Council of the Regions of Europe in 1985), which brings together over 250 regions from 35 countries along with 16 interregional organisations, has long been a supporter of subsidiarity, the decentralising principle according to which political matters ought to be dealt with by the smallest, lowest, or least centralised authority that could address them effectively.

Of course, the historical state is not giving up its role without a fight. In many contexts, it is trying to reclaim its primacy as the information super-agent governing the political life of the society that it organises.

In some cases, the attempt is blatant. In the UK, the Labour Government introduced the first Identity Cards Bill in November 2004. After several intermediary stages, the Identity Cards Act was finally repealed by the Identity Documents Act 2010, on 21 January 2011. The failed plan to introduce compulsory ID in the UK should be read from a modern perspective of preserving a Westphalian order.

In many other cases, it is 'historical resistance' by stealth, as when an information society is largely run by the state. In this case, the state maintains its role of major informational agent no longer just legally, on the basis of its power over legislation and its implementation, but also economically, on the basis of its power over the majority of information-based jobs. The intrusive presence of so-called State Capitalism with its State Owned Enterprises all over the world, from Brazil, to France, to China, is a symptom of hyperhistorical anachronism.

Similar forms of resistance seem only able to delay the inevitable rise of political multiagent systems. Unfortunately, they may involve not only costs, but also huge risks, both locally and globally. Recall that the two World Wars may be seen as the end of the Westphalian system. Paradoxically, while humanity is moving into a hyperhistorical age, the world is witnessing the rise of China, currently a most 'historical' state, and the decline of the US, a state that more than any other superpower in the past already had a hyperhistorical and multiagent vocation in its federal organisation. We might be moving from a Washington Consensus to a *Beijing Consensus* described by Williamson as consisting of incremental reform, innovation

and experimentation, export-led growth, state capitalism, and authoritarianism.[11] This is risky, because the anachronistic historicism of some of China's policies and humanity's growing hyperhistoricism are heading towards a confrontation. It may not be a conflict, but hyperhistory is a force whose time has come, and while it seems likely that it will be the Chinese state that will emerge deeply transformed, one can only hope that the inevitable friction will be as painless and peaceful as possible. The financial and social crises that the most advanced information societies are undergoing may actually be the painful but still peaceful price we need to pay to adapt to a future post-Westphalian system.

The previous conclusion holds true for the historical state in general. In the future, political multiagent systems will acquire increasing prominence, with the problem that the visibility and transparency of such acquisition of power may be rather unclear. It is already difficult to monitor and understand politics when states are the main players. It becomes even harder when the agents in question have fuzzier features, more opaque behaviours, and are much less easily identifiable, let alone accountable. At the same time, it is to be hoped that the state itself will progressively abandon its resistance to hyperhistorical changes and evolve even more into a multiagent system. Good examples are provided by devolution, the transfer of state's sovereign rights to supranational European institutions, or the growing trend in making central banks, like the Bank of England or the European Central Bank, independent, public organisations.

The time has come to consider the nature of political multiagent system more closely and some of the questions that its emergence is already posing.

6.1.3 The Political Multiagent System

A political multiagent system is a single agent, constituted by other systems, which is

a. *teleological*: the multiagent system has a purpose, or goal, which it pursues through its actions;
b. *interactive*: the multiagent system and its environment can act upon each other;
c. *autonomous*: the multiagent system can change its configurations without direct response to interaction, by performing internal transformations to change its states. This imbues the multiagent system with some degree of complexity and independence from its environment; and finally
d. *adaptable*: the multiagent system's interactions can change the rules by which the multiagent system itself changes its states. Adaptability ensures that the multiagent system learns its own mode of operation in a way that depends critically on its experience.

[11] Williamson (2012, pp. 1–16). The expression 'Beijing Consensus' was introduced by Ramo and Foreign Policy Centre (London, England) (2004) but I am using it here in the sense discussed by Williamson and Halper (2010).

The political multiagent system becomes *intelligent* (in the AI sense of "smart") when it implements the previous features efficiently and effectively, minimising resources, wastefulness and errors, while maximising the returns of its actions.

The emergence of intelligent, political multiagent systems poses many serious questions. Some of them are worth reviewing here, even if only quickly: identity, cohesion, consent, social vs. political space, legitimacy, and transparency.

Identity Throughout modernity, the state has dealt with the problem of establishing and maintaining its own *identity* by working on the equation between state = nation. This has often been achieved through the legal means of citizenship and the narrative rhetoric of space (the Mother/Father Land) and time (story in the sense of traditions, recurrent celebrations of past nation-building events, etc.). Consider, for example, the invention of mandatory military service during the French Revolution, its increasing popularity in modern history, but then the decreasing number of sovereign states that still impose it nowadays (your author belongs to the last generation that had to serve in the Italian army for 12 months). Conscription transformed waging war from an eminently economic problem—Florentine bankers financed the English kings during the Hundred Years War (1337–1453), for example—into also a legal problem: the right of the state to send its citizens to die on its behalf. It thus made human life the *penultimate* value, available for the ultimate sacrifice, in the name of patriotism: 'for King and Country'. It is a sign of modern anachronism that, in moments of crisis, sovereign states still give in to the temptation of fuelling nationalism about meaningless, *geographical* spots, often some small islands unworthy of any human loss, including the Falkland Islands (UK) or Islas Malvinas (Argentina), the Senkaku (Japan) or Diaoyu (China) islands, and the Liancourt Rocks, also known as Dokdo (South Korea) or Takeshima (Japan).

Cohesion The equation between state, nation, citizenship and land/story had the further advantage of providing an answer to a second problem, that of *cohesion*. For the equation answered not only the question of who or what the state is, but also the question of who or what belongs to the state and hence may be subject to its norms, policies, and actions. New political multiagent systems cannot rely on the same solution. Indeed, they face the further problem of having to deal with the decoupling of their political identity and cohesion. The political identity of a multiagent system may be strong and yet unrelated to its temporary and rather loose cohesion, as is the case with the Tea Party movement in the US. Both identity and cohesion of a political multiagent system may be rather weak, as in the international Occupy movement. Or one may recognise a strong cohesion and yet an unclear or weak political identity, as with the population of tweeting individuals and their role during the Arab Spring. Both identity and cohesion of a political multiagent system are established and maintained through information sharing. The land is virtualised into the region of the infosphere in which the multiagent system operates. So memory (retrievable recordings) and coherence (reliable updates) of the information flow enable a political multiagent system to claim some identity and some cohesion, and therefore offer a sense of belonging. But it is, above all, the fact that the boundaries between the online and offline are disappearing, the appearance of the onlife

experience, and hence the fact that the virtual infosphere can affect politically the physical space, that reinforces the sense of the political multiagent system as a real agent. If *Anonymous* had only a virtual existence, its identity and cohesion would be much less strong. Deeds provide a vital counterpart to the virtual information flow to guarantee cohesion. Interactions become more fundamental than things, in a way that is coherent with interactability as a criterion of existence and the development of informational identities. With word play, we might say that *ings* (as in interact-*ing*, process-*ing*, network-*ing*, do-*ing*, be-*ing*, etc.) replace *things*.

Consent The breaking up of the equation 'political multiagent system = sovereign state, citizenship, land, story, nation' and the decoupling of identity and cohesion in a political multiagent system have a significant consequence. The age-old theoretical problem of how consent to be governed by a political authority arises is being turned on its head. In the historical framework of social contract theory, the presumed default position is that of a legal opt-out. There is some kind (to be specified) of original consent, allegedly given (for a variety of reasons) by any individual subject to the political state, to be governed by the latter and its laws. The problem is to understand how such consent is given and what happens when an agent, especially a citizen, opts out of it (the out-law). In the hyperhistorical framework, the expected default position is that of a social opt-in, which is exercised whenever the agent subjects itself to the political multiagent system conditionally, for a specific purpose. Simplifying, we are moving from being part of the political consensus to taking part in it, and such part-taking is increasingly 'just in time', 'on demand', 'goal-oriented', and anything but stable, permanent or long-term. If doing politics looks increasingly like doing business this is because, in both cases, the interlocutor, the citizen-customer, needs to be convinced to behave in a preferred way every time anew. Loyal membership is not the default position, and needs to be built and renewed around political and commercial products alike. Gathering consent around specific political issues becomes a continuous process of (re)engagement. It is not a matter of limited attention span. The generic complaint that 'new generations' cannot pay sustained attention to political problems any more is ill-founded. They are, after all, the generations that binge-watch TV. It is a matter of motivating interest again and again, without running into an inflation of information (one more crisis, one more emergency, one more revolution, one more…) and political fatigue (how many times do we need to intervene urgently?). Therefore, the problem is to understand what may motivate repeatedly or indeed force agents (again, not just individual human beings, but all kinds of agents) to give such consent and become engaged, and what happens when such agents, unengaged by default (note, not disengaged, for disengagement presupposes a previous state of engagement), prefer to stay away from the activities of the political multiagent system, inhabiting a social sphere of civil but apolitical identity.

Failing to grasp the previous transformation from historical opt-out to hyperhistorical opt-in means being less likely to understand the apparent inconsistency between the disenchantment of individuals with politics and the popularity of global movements, international mobilisations, activism, voluntarism, and other

social forces with huge political implications.[12] What is moribund is not politics *tout court*, but historical politics, that based on parties, classes, fixed social roles, political manifestos and programs, and the sovereign state, which sought political legitimacy only once and spent it until revoked. The inching towards the so-called centre by parties in liberal democracies around the world, as well as the 'get out the vote' strategies (the expression is used to describe the mobilisation of *voters as supporters* to ensure that those who can do vote) are evidence that engagement needs to be constantly renewed and expanded in order to win an election. Party (as well as Union) membership is a modern feature that is likely to become increasingly less common.

Social vs. Political Space In prehistory, the social and the political spaces overlap because, in a stateless society, there is no real difference between social and political relations and hence interactions. In history, the state tends to maintain such co-extensiveness by occupying, as an informational multiagent system, the social space politically, thus establishing the primacy of the political over the social. This trend, if unchecked and unbalanced, risks leading to totalitarianisms (consider for example the Italy of Mussolini), or at least broken democracies (consider next the Italy of Berlusconi). We have seen earlier that such co-extensiveness and its control may be based on normative or economic strategies, through the exercise of power, force, and rule-making. In hyperhistory, the social space is the original, default space from which agents may move to (consent to) join the political space. It is not accidental that concepts such as *civil society*,[13] *public sphere*,[14] and *community* become increasingly important the more we move into a hyperhistorical context. The problem is to understand and design such social space where agents of various kinds are supposed to be interacting and which give rise to the political multiagent system.

Each agent within the social space has some degrees of freedom. By this I do not mean liberty, autonomy or self-determination, but rather, in the robotic, more humble sense, some capacities or abilities, supported by the relevant resources, to engage in specific actions for a specific purpose. To use an elementary example, a coffee machine has only one degree of freedom: it can make coffee, once the right ingredients and energy are supplied. The sum of an agent's degrees of freedom is its 'agency'. When the agent is alone, there is of course only agency, but no social let alone political space. Imagine Robinson Crusoe on his 'Island of Despair'. However, as soon as there is another agent (Friday on the 'Island of Despair'), or indeed a group of agents (the native cannibals, the shipwrecked Spaniards, the English mutineers), agency acquires the further value of social interaction. Practices and then rules for coordination and constraint of the agents' degrees of freedom become

[12] On volunteerism see United Nations (2011). *State of the World's Volunteerism Report, 2011: Universal Values for Global Well-being*, United Nations Volunteers., on digital activism, the Digital Activism Research Project (http://digital-activism.org/) offers a wealth of information.

[13] I use the expression here in the post-Hegelian sense of non-political society.

[14] The social space where people can meet, identify and discuss societal problems, shaping political actions.

essential, initially for the well-being of the agents constituting the multiagent system, and then for the well-being of the multiagent system itself. Note the shift in the level of analysis: once the social space arises, we begin to consider the group as a group—e.g., as a family, or a community, or as a society—and the actions of the individual agents constituting it become elements that lead to the newly established degrees of freedom, or agency, of the multiagent system. The previous simple example may still help. Consider now a coffee machine and a timer: separately, they are two agents with different agency, but if they are properly joined and coordinated into a multiagent system, then the issuing agent has the new agency to make coffee at a set time. It is now the multiagent system that has a more complex capacity, and that may or may not work properly.

A social space is the totality of degrees of freedom of the inhabiting agents one wishes to take into consideration. In history, such consideration—which is really just another level of analysis—was largely determined physically and geographically, in terms of presence in a territory, and hence by a variety of forms of neighbourhood. In the previous example, all the agents interacting with Robinson Crusoe are taken into consideration because of their relations (interactive presence in terms of their degrees of freedom) to the same 'Island of Despair'. We saw that ICTs have changed all this. In hyperhistory, where to draw the line to include, or indeed exclude, the relevant agents whose degrees of freedom constitute the social space has become increasingly a matter of at least implicit choice, when not of explicit decision. The result is that the phenomenon of distributed morality, encompassing that of distributed responsibility, is becoming more and more common. In either case, history or hyperhistory, what counts as a social space may be a political move. Globalisation is a de-territorialisation in this political sense.

Turning now to the political space in which the new multiagent systems operate, it would be a mistake to consider it a separate space, over and above the social one. Both the social and the political space are determined by the same totality of the agents' degrees of freedom. The political space emerges when the complexity of the social space requires the prevention or resolution of potential *divergences* and coordination or collaboration about potential *convergences*. *Both* are crucial. And in each case information is required, in terms of representation and deliberation about a complex multitude of degrees of freedom.

Legitimacy It is when the agents in the social space agree to agree on how to deal with their divergences (conflicts) and convergences that the social space acquires the political dimension to which we are so used. Yet two potential mistakes await us here.

The first, call it Hobbesian, is to consider politics merely as the prevention of war by other means, to invert the famous phrase by von Clausewitz (1780–1831), according to whom 'war is the continuation of politics by other means'. This is an unsatisfactory view of politics, because even a complex society of angels would still require rules in order to further its harmony. Convergences too need politics. Politics is not just about conflicts due to the agents' exercises of their degrees of freedom when pursuing their goals. It is also, or at least it should be, above all, the

furthering of coordination and collaboration of degrees of freedom by means other than coercion and violence.

The second potential mistake, which may be called Rousseauian, is to misunderstand the political space as just that part of the social space organised by law. In this case, the mistake is subtler. We usually associate the political space with the rules or laws that regulate it but the latter are not constitutive, by themselves, of the political space. Compare two cases in which rules determine a game. In chess, the rules do not merely constrain the game; they are the game because they do not supervene on a previous activity. Rather, they are the *necessary and sufficient conditions* that determine all and only the moves that can be legally made. In football, however, the rules are supervening *constraints* because the agents enjoy a previous and basic degree of freedom, consisting in their capacity to kick a ball with the foot in order to score a goal, which the rules are supposed to regulate. Whereas it is physically possible, but makes no sense, to place two pawns on the same square of a chessboard, nothing impeded Maradona from scoring an infamous goal by using his hand in the Argentina vs. England football match (1986 FIFA World Cup), and that to be allowed by a referee who did not see the infringement. Now, the political space is not simply *constituted* by the laws that regulate it, as in the chess example. But it is not just the result of the *constraining* of the social space by means of laws either, as in the football example. The political space is that area of the social space *configured* by the agreement to agree on resolution of divergences and coordination of convergences. The analogy here is the formatting of a hard disk. This leads to a further consideration, concerning the transparent multiagent system, especially when, in this transition time, the multiagent system in question is still the state.

Transparency There are two senses in which the multiagent system can be transparent. They mean quite different things, and so they can be confusing. Unsurprisingly, both come from ICTs and computer science, one more case in which the information revolution is changing our conceptual framework.

On the one hand, the multiagent system (think of the sovereign state, and also of corporate agents, multinationals, or supranational institutions, etc.) can be transparent in the sense that it moves from being a black box to being a white box. Other agents (citizens, when the multiagent system is the state) not only can see inputs and outputs—for example, levels of tax revenue and public expenditure—they can also monitor how (in our running example, the state as) a multiagent system works internally. This is not a novelty at all. It was a principle already popularised in the 19th century. However, it has become a renewed feature of contemporary politics due to the possibilities opened up by ICTs. This kind of transparency is also known as *Open Government*.

On the other hand, and this is the more innovative sense that I wish to stress here, the multiagent system can be transparent in the sense of being ‘invisible’. This is the sense in which a technology (especially an interface) is transparent: not because it is not there, but because it delivers its services so efficiently, effectively, and reliably that its presence is imperceptible. When something works at its best, behind the scenes as it were, to make sure that we can operate as smoothly as possible,

then we have a transparent system. When the multiagent system in question is the state, this second sense of transparency should not be seen as a surreptitious way of introducing, with a different terminology, the concept of 'small state' or 'small governance'. On the contrary, in this second sense, the multiagent system (the state) is as transparent and as vital as the oxygen that we breathe. It strives to be the ideal butler. There is no standard terminology for this kind of transparent multiagent system that becomes perceivable only when it is absent. Perhaps one may speak of *Gentle Government*.

It seems that multiagent systems can increasingly support the right sort of ethical infrastructure (more on this later) the more transparently, that is, openly and gently, they play the negotiating game through which they take care of the *res publica*. When this negotiating game fails, the possible outcome is an increasingly violent conflict among the parties involved. It is a tragic possibility that ICTs have seriously reshaped.

All this is not to say that *opacity* does not have its virtues. Care should be exercised, lest the socio-political discourse is reduced to the nuances of higher quantity, quality, intelligibility, and usability of information and ICTs. The more the better is not the only, nor always the best, rule of thumb. For the withdrawal of information can often make a positive and significant difference. We already encountered Montesquieu's division of the state's political powers. Each of them may be informationally opaque in the right way to the other two. For one may need to lack (or intentionally preclude oneself from accessing) some information in order to achieve desirable goals, such as protecting anonymity, enhancing fair treatment, or implementing unbiased evaluation. Famously, Rawls' 'veil of ignorance' exploits precisely this aspect of information, in order to develop an impartial approach to justice.[15] Being informed is not always a blessing and might even be dangerous or wrong, distracting or crippling. The point about the value of transparency is that its opposite, informational opacity, cannot be assumed to be a good property of a political system unless it is adopted explicitly and consciously, by showing that it is a feature not a mere bug.

6.1.4 Infraethics

Part of the ethical efforts engendered by the fourth revolution concerns the design of environments that can facilitate ethical choices, actions, or process. This is not the same as *ethics by design*. It is rather *pro-ethical design*, as I hope will become clearer in the following pages. Both are liberal, but *ethics by design* may be mildly paternalistic, insofar as it privileges the facilitation of the *right* kind of choices, actions, process or interactions on behalf of the agents involved. Whereas *pro-ethical design* does not have to be paternalistic, insofar as it privileges the facilitation of *reflection* by the agents involved on their choices, actions, or process. For example,

[15] Rawls (1999).

strategies based on *ethics by design* may let you opt out of the *default* preference according to which, by obtaining a driving licence, you are also willing to be an organ donor. Strategies based on *pro-ethical design* may not allow you to obtain a driving license unless you have indicated whether you wish to be an organ donor, the unbiased choice is still all yours. In this section, I shall call environments that can facilitate ethical choices, actions, or process, the ethical infrastructure, or *infraethics*. The problem is how to design the right sort of infraethics. Clearly, in different cases, the design of a liberal infraethics may be more or less paternalistic. My argument is that it should be as little paternalistic as the circumstances permit, although no less.

It is a sign of the times that, when politicians speak of infrastructure nowadays, they often have in mind ICTs. They are not wrong. From business fortunes to conflicts, what makes contemporary societies work depends increasingly on bits rather than atoms. We already saw all this. What is less obvious, and intellectually more interesting, is that ICTs seem to have unveiled a new sort of ethical equation.

Consider the unprecedented emphasis that ICTs have placed on crucial phenomena such as trust, privacy, transparency, freedom of expression, openness, intellectual property rights, loyalty, respect, reliability, reputation, rule of law, and so forth. These are probably better understood in terms of an infrastructure that is there to facilitate or hinder (reflection upon) the im/moral behaviour of the agents involved. Thus, by placing our informational interactions at the centre of our lives, ICTs seem to have uncovered something that, of course, has always been there, but less visibly so: the fact that the moral behaviour of a society of agents is also a matter of 'ethical infrastructure' or simply *infraethics*. An important aspect of our moral lives has escaped much of our attention. Many concepts and related phenomena have been mistakenly treated as if they were only ethical, when in fact they are probably mostly infraethical. To use a term from the philosophy of technology, such concepts and the corresponding phenomena have a dual-use nature: they can be morally good, but also morally evil (more on this presently). The new equation indicates that, in the same way that, in an economically mature society, business and administration systems increasingly require infrastructures (transport, communication, services etc.) to prosper, so too, in an informationally mature society, multiagent systems' moral interactions increasingly require an infraethics to flourish.

The idea of an infraethics is simple, but can be misleading. The previous equation helps to clarify it. When economists and political scientists speak of a 'failed state', they may refer to the failure of a *state-as-a-structure* to fulfil its basic roles, such as exercising control over its borders, collecting taxes, enforcing laws, administering justice, providing schooling, and so forth. In other words, the state fails to provide *public goods*, such as defence and police, and *merit goods*, such as healthcare. Or (too often an inclusive and intertwined or) they may refer to the collapse of a *state-as-an-infrastructure* or environment, which makes possible and fosters the right sort of social interactions. This means that they may be referring to the collapse of a substratum of default expectations about economic, political and social conditions, such as the rule of law, respect for civil rights, a sense of political community, civilised dialogue among differently-minded people, ways to reach peaceful resolutions of ethnic, religious, or cultural tensions, and so forth. All these expectations,

attitudes, practices, in short such an implicit 'socio-political infrastructure', which one may take for granted, provides a vital ingredient for the success of any complex society. It plays a crucial role in human interactions, comparable to the one that we are now accustomed to attributing to physical infrastructures in economics.

Infraethics should not be understood in terms of Marxist theory, as if it were a mere update of the old 'base and superstructure' idea. The elements in question are entirely different: we are dealing with moral actions and not-yet-moral facilitators of such moral actions. Nor should it be understood in terms of a kind of second-order normative discourse on ethics. It is the not-yet-ethical framework of implicit expectations, attitudes, and practices that *can* facilitate and promote moral decisions and actions. At the same time, it would also be wrong to think that an infraethics is morally neutral. Rather, it has a dual-use nature, as I anticipated earlier: it can both facilitate and hinder morally good as well as evil actions, and do this in different degrees. At its best, it is the grease that lubricates the moral mechanism. This is more likely to happen whenever having a 'dual-use' nature does not mean that each use is equally likely, that is, that the infraethics in question is still not neutral, nor merely positive, but does have a bias to deliver more good than evil. If this is confusing, think of the dual-use nature not in terms of a state of equilibrium, like an ideal coin that can deliver both heads and tails, but in terms of a co-presence of two alternative outcomes, one of which is more likely than the other, as a biased coin more likely to turn heads than tails. When an infraethics has a 'biased dual-use' nature, it is easy to mistake the infraethical for the ethical, since whatever helps goodness to flourish or evil to take root partakes of their nature.

Any successful complex society, be this the City of Man or the City of God, relies on an implicit infraethics. This is dangerous, because the increasing importance of an infraethics may lead to the following risk: that the legitimization of the ethical discourse is based on the 'value' of the infraethics that is supposed to support it. *Supporting* is mistaken for *grounding*, and may even aspire to the role of *legitimizing*, leading to what the French philosopher Jean-François Lyotard (1924–1998) criticized as mere 'performativity' of the system, independently of the actual values cherished and pursued. As an example, think of a bureaucratic context in which some procedure, supposed to deliver some morally good behavior, through time becomes a value in itself, and ends giving ethical value to the behavior that was supposed to support. Infraethics is the vital syntax of a society, but it is not its semantics, to re-use a distinction we encountered when discussing artificial intelligence. It is about the structural form, not the meaningful contents.

We saw earlier that even a society in which the entire population consisted of angels, that is, perfectly moral agents, still needs norms for collaboration and coordination. Theoretically, a society may exist in which the entire population consisted of Nazi fanatics who could rely on high levels of trust, respect, reliability, loyalty, privacy, transparency, and even freedom of expression, openness, and fair competition. Clearly, what we want is not just the successful mechanism provided by the right infraethics, but also the coherent combination between it and morally good values, such as civil and political rights. This is why a balance between security and privacy, for example, is so difficult to achieve, unless we clarify first whether

we are dealing with a tension within ethics (security and privacy as moral rights), within infraethics (both are understood as not-yet-ethical facilitators), or between infraethics (security) and ethics (privacy), as I suspect. To rely on another analogy: the best pipes (infraethics) may improve the flow but do not improve the quality of the water (ethics); and water of the highest quality is wasted if the pipes are rusty or leaky. So creating the right sort of infraethics and maintaining it is one of the crucial challenges of our time, because an infraethics is not morally good in itself, but it is what is most likely to yield moral goodness if properly designed and combined with the right moral values. The right sort of infraethics should be there to support the right sort of values. It is certainly a constitutive part of the problem concerning the design of the right multiagent systems.

The more complex a society becomes, the more important and hence salient the role of a well-designed infraethics is, and yet this is exactly what we seem to be missing. Consider the recent Anti-Counterfeiting Trade Agreement (ACTA), a multinational treaty concerning the international standards for intellectual property rights.[16] By focusing on the enforcement of intellectual property rights (IPR), supporters of ACTA completely failed to perceive that it would have undermined the very infraethics that they hoped to foster, namely one promoting some of the best and most successful aspects of our information society. It would have promoted the structural inhibition of some of the most important individuals' positive liberties and their ability to participate in the information society, thus fulfilling their own potential as informational organisms. For lack of a better word, ACTA would have promoted a form of *informism*, comparable to other forms of social agency's inhibition such as classism, racism, and sexism. Sometimes a protection of liberalism may be inadvertently illiberal. If we want to do better, we need to grasp that issues such as IPR are part of the new infraethics for the information society, that their protection needs to find its carefully balanced place within a complex legal and ethical infrastructure that is already in place and constantly evolving, and that such a system must be put at the service of the right values and moral behaviours. This means finding a compromise, at the level of a liberal infraethics, between those who see new legislation (such as ACTA) as a simple fulfilment of existing ethical and legal obligations (in this case from trade agreements), and those who see it as a fundamental erosion of existing ethical and legal civil liberties.

In hyperhistorical societies, any regulation affecting how people deal with information is now bound to influence the whole infosphere and onlife habitat within which they live. So enforcing rights such as IPR becomes an environmental problem. This does not mean that any legislation is necessarily negative. The lesson here is one about complexity: since rights such as IPR are part of our infraethics and affect our whole environment understood as the infosphere, the intended and unintended consequences of their enforcement are widespread, interrelated, and far-reaching. These consequences need to be carefully considered, because mistakes will generate huge problems that will have cascading costs for future generations, both ethically and economically. The best way to deal with 'known unknowns' and

[16] For a more detailed analysis see Floridi (2012).

unintended consequences is to be careful, stay alert, monitor the development of the actions undertaken, and be ready to revise one's decision and strategy quickly, as soon as the wrong sort of effects start appearing. *Festina lente*, 'more haste, less speed' as the classic adage suggests. There is no perfect legislation but only legislation that can be perfected more or less easily. Good agreements about how to shape our infraethics should include clauses about their timely updating.

Finally, it is a mistake to think that we are like outsiders ruling over an environment different from the one we inhabit. Legal documents (such as ACTA) emerge from within the infosphere that they affect. We are building, restoring, and refurbishing the house from inside. Recall that we are repairing the raft while navigating on it, to use the metaphor introduced in the Preface. Precisely because the whole problem of respect, infringement, and enforcement of rights such as IPR is an infraethical and environmental problem for advanced information societies, the best thing we can do, in order to devise the right solution, is to apply to the process itself the very infraethical framework and ethical values that we would like to see promoted by it. This means that the infosphere should regulate itself from within, not from an impossible without.

6.1.5 Hyperhistorical Conflicts and Cyberwar

The story goes that when the Roman horsemen first saw Pyrrhus' twenty war elephants, at the battle of Heraclea (280 BC), they were so terrorised by these strange creatures, which they had never seen before, that they galloped away, and the Roman legions lost the battle. Today, the new elephants are digital. The phenomenon might have just begun to emerge in the public debate but, in hyperhistorical societies, ICTs are increasingly shaping armed conflicts.

Disputes become armed conflicts when politics fails. In hyperhistory, such armed conflicts have acquired a new informational nature. Cyberwar or information warfare is the continuation, and sometimes the replacement, of conflict by digital means, to rely once more on von Clausewitz's famous interpretation of war we encountered above. Four main changes are notable.

First, in terms of conventional military operations, ICTs have progressively revolutionized communications, making possible complex new modes of field operations. We saw this was already the case with the *Chappe* telegraph.

Second, ICTs have also made possible the swift analysis of vast amounts of data, enabling the military, intelligence and law enforcement communities to take action in ever more timely and targeted ways. ICTs and Big Data are also weapons.

Third, and even more significantly, battles are nowadays fought by highly mobile forces, armed with real-time ICT devices, satellites, battlefield sensors and so forth, as well as thousands of robots of all kinds.

And, finally, the growing dependence of societies and their militaries on advanced ICTs has led to strategic cyber-attacks, designed to cause costly and crippling disruption. Armies of human soldiers may no longer be needed. This creates

a stark contrast with suicide terrorism. On the one hand, human life can regain its ultimate value because the state no longer needs to trump it in favour of patriotism. Contrary to what we saw in the previous pages, drones do not die 'for King and Country'. Cyberwar is a hyperhistorical phenomenon. On the other hand, terrorists de-humanise individuals as mere delivery mechanisms. Suicide terrorism is a historical phenomenon, in which the technology in-between is the human body and a person becomes a 'living tool', using Aristotle's definition of a slave.

The old economic problem—how to finance war and its expensive high tech—is now joined by a new legal problem: how to reconcile a hyperhistorical kind of warfare with historical phenomena, such as the infringement of national sovereignty and respect for geographical borders. Furthermore, cyber-attacks can be undertaken by nations or networks, or even by small groups or individuals. ICTs have made asymmetric conflicts easier, and shifted the battleground more than an inch into the infosphere.

The scale of such transformations is staggering. For example, in 2003, at the beginning of the war in Iraq, US forces had no robotic systems on the ground. However, by 2004, they had already deployed 150 robots, in 2005 the number was 2,400; and by the end of 2008, about 12,000 robots of nearly two dozen varieties were operating on the ground.[17]

In 2010, Neelie Kroes, Vice-President of the European Commission, commenting on Cyber Europe 2010, the first pan-European cyber-attack simulation, said that:

> This exercise to test Europe's preparedness against cyber threats is an important first step towards working together to combat potential online threats to essential infrastructure and ensuring citizens and businesses feel safe and secure online.[18]

As you can see, the perspective could not be more hyperhistorical.

ICT-mediated modes of conflict pose a variety of ethical problems, for war-fighting militaries in the field, for intelligence gathering services, for policy makers, and for ethicists. They may be summarised as the three Rs: risks, rights and responsibilities.

Risks Cyberwar and information-based conflicts may increase risks, making 'soft' conflicts more likely and hence potentially increasing the number of casualties. Between 2004 and 2012, drones operated by the US' Central Intelligence Agency (CIA) killed more than 2,400 people in Pakistan, including 479 civilians, with 3 strikes in 2005 escalating to 76 strikes in 2011.[19] A troubling perspective is that ICTs might make unconventional conflicts more acceptable ethically, by stressing the less deadly outcome of military operations in cyberspace. However, this might be utterly illusory. Messing with ICT-infrastructures of hospitals and airports may easily cause the loss of human lives, even if in a less obvious way than bombs do.

[17] Source: *The New Atlantis* report, available online.

[18] Source: Press release, *Digital Agenda: cyber-security experts test defences in first pan-European simulation*, available online.

[19] The Economist (2012).

Despite this, the mistaken impression remains that we might be allegedly moving towards a more precise, surgical, bloodless way of handling violently our political disagreements.

Rights Cyberwar tends to erase the threshold between reality and simulation, between life and play, and between conventional conflicts, insurgencies or terrorist actions. This threatens to increase the potential tensions between fundamental rights: informational threats require higher levels of control, which may generate conflicts between individuals' rights (e.g. privacy) and community's rights (e.g. safety and security). A state's duty to protect its citizens may come to clash with its duty to prevent harm to its citizens, via an extended system of surveillance, which may easily end up infringing on citizens' privacy.

Responsibilities Cyberwar makes it more difficult to identify responsibilities that are reshaped and distributed. Because causal links are much less easily identifiable, it becomes much more difficult to establish who, or what, is accountable and responsible when software/robotic weapons and hybrid, man-machine systems are involved.

New risks, rights and responsibilities: in short, cyberwar is a new phenomenon, which has caught us by surprise. With hindsight, we should have known better, for at least three reasons.

Take the nature of our society first. When it was modern and industrial, conflicts had mechanised, second-order features. Engines, from battleships to tanks to aeroplanes, were weapons, and the coherent outcome was the emphasis on energy, petrol first and then nuclear power. There was an eerie analogy between assembly lines and warfare trenches, between working force and fighting force. Conventional warfare was kinetic warfare. We just did not know it, because the non-kinetic kind was not yet available. The Cold War and the emergence of asymmetric conflicts were part of a post-industrial transformation. Today, in a culture in which we have seen that the word 'engine' is more likely to be preceded by the verb 'search' than by the noun 'petrol', hyperhistorical societies are as likely to fight with digits as they are with bullets, with computers as well as guns, not least because digital systems tend to be in charge of analogue weapons. I am not referring to the use of intelligence, espionage, or cryptography, but to cyber attacks or to the extensive use of drones and other military robots in Iraq and Afghanistan. It is old news. On 27th of April 2007, about 1 million computers worldwide were used for DDOS (distributed denial of service) attacks on Estonian government and corporate web sites. A DDOS attack is a systematic attempt to make computer resources unavailable, at least temporarily, by forcing vital sites or services to reset or consume their resources, or by disrupting their communications so that they can no longer function properly. Russia was blamed but denied any involvement. In June 2010, Stuxnet, a sophisticated computer malware, sabotaged ca. 1000 Siemens centrifuges used in the Iranian nuclear power plant of Bushehr. That time, the US and Israel denied any involvement. At the time of writing, there is an on-going attack on US ICT infrastructure. This time China that denies any involvement. Then there are robotic weapons, which may be seen as the final stage in the industrialisation of warfare, or, more interestingly, as

the first step in the development of information conflicts, in which command and control as well as action and reaction become tele-concepts. Third-order technological conflicts in which humans are no longer in the loop have moved out of science fiction and into military scenarios. From software agents in cyberspace to robots in physical environments we should not be too optimistic about the non-violent nature of cyberwar. The more we rely on ICTs, the more we envelop the world, the more cyber attacks will become lethal. Soon, crippling an enemy's communication and information infrastructure will be like zapping its pacemaker rather than hacking its mobile.

Second, consider the nature of our environment. We have been talking about the internet and cyberspace for decades. We could have easily imagined that this would become the new frontier for human conflicts. Technologies have continuously expanded. We have been fighting each other on land, at sea, in the air, and in space for as long, and as soon as technologies made it possible. Predictably, the infosphere was never going to be an exception. Information is the fifth element,[20] and the military now speaks of cyberwarfare as 'the fifth domain of warfare'. The impression is that, in the future, such a fifth domain will end up dominating the others. The following two examples may help. On 13th of May 1999, arguably the first combat between an aircraft and an unmanned drone took place when an Iraqi MiG-25 shot down a US Air Force unmanned MQ-1 Predator drone. More than 360 drones have been built since 1995, for more than $ 2.38 billion. Second, since 2006, Samsung, the maker of smart phones and refrigerators, has also been producing the SGR-A1. It is a robot with a low-light camera and pattern recognition software to distinguish humans from animals or other objects. It patrols South Korea's border with North Korea and, if necessary, it can autonomously fire its built-in machine gun. It is increasingly hard to draw a clear distinction between cyberwarfare and conventional, kinetic warfare when some tele-warfare is in question.

Finally, think of the origin of cybernetics, the computer, the Internet, the Global Positioning System (GPS), and unmanned drones and vehicles. They all developed initially as part of wider military efforts. The history of computing is deeply rooted in the Second World War and Turing's work at Bletchley Park. Cybernetics, the ancestor of contemporary robotics, begun to develop as an engineering field in connection with applications for the automatic control of gun mounts and radar antenna, still during the Second World War. We know that the internet was the outcome of the arms race and of nuclear proliferation, but we were distracted by the development of the Web and its scientific origins, and forgot about the Defense Advanced Research Projects Agency (DARPA). The now ubiquitous GPS, which provides the satellite-based information for navigation systems, was created and developed by the US Department of Defence, one more case of the political importance of geography. It became freely available for civilian use only in 1983, after a Boeing 747 of the Korean Air Lines, with 269 people on board, was shot down because it had strayed into the USSR's prohibited airspace. Finally, the development of drones, mainly but not only by the US military, as well as autonomous vehicles (DARPA

[20] Floridi (1999).

again) and other robots, owes much to the conflicts in Iraq and Afghanistan and the fight against terrorism. In short, much of the history of digital ICTs spookily corresponds to the history of conflicts and the financial efforts behind them: Second World War, Cold War, First and Second Iraq War, War in Afghanistan, and various 'wars' on terrorist organisations around the world. Hyperhistory has merely caught up with us.

The previous outline should help one understand why cyberwar, or more generally information warfare, is causing radical transformations in our ways of thinking about military, political, and ethical issues. The concepts of state, war, and the distinction between civil society and military organisations are being affected. Are we going to see a new arms race, given the high rate at which cyber weapons "decay"? After all, you can use a piece of malware only once, for a patch will then become available, and often only within, and against, a specific technology that will soon be out of date. If cyber disarmament is ever going to be an option, how do you decommission cyber weapons? Digital systems can be hacked: will the Pony Express make a patriotic comeback in the near future as the last line of defence against an enemy that could tamper with anything digital and online? Some questions make one smile, but others are increasingly problematic. Let me highlight two sets of them that should be of more general interest.

The body of knowledge and discussion behind Just War Theory is detailed and extensive.[21] It is the result of centuries of refinements since Roman times. The methodological question we face today is whether information warfare is merely one more area of application, or whether it represents a disruptive novelty as well, which will require new developments of the theory itself. For example, within the *jus ad bellum*, which kind of authorities possesses the legitimacy to wage cyberwar? And how should a cyber attack be considered in terms of last resort, especially when a cyber attack could, allegedly, prevent more violent outcomes? And within the *jus in bello*, what level of proportionality should be attributed to a cyber attack? How do you surrender to cyber enemies, especially when their identities are unknown on purpose? Or how will robots deal with non-combatants or treat prisoners? Is it possible or even desirable to develop in-built 'ethical algorithms' when engineering robotic weapons?

Equally developed, in this case since Greek times, is our understanding of military virtue ethics. How is the latter going to be applied to phenomena that are actually reshaping the conditions of possibility of virtue ethics itself? Bear in mind that any virtue ethics presupposes a philosophical anthropology, that is, a view of the human nature that may be Aristotelian, Buddhist, Christian, Confucian, Fascist, Nietzschean, Spartan, and so forth. Information warfare is only part of the information revolution, which is also affecting our self-understanding as informational organisms. Take for example the classic virtue of courage: in what sense can someone be courageous when tele-manoeuvring a military robot? Indeed, will courage still rank so highly among the virtues when the capacity to evaluate and manage information

[21] For a study of how current international law applies to cyber conflicts and cyber warfare, see NATO Cooperative Cyber Defence Centre of Excellence (2013).

and act upon it wisely and promptly will seem to be a much more important trait of a soldier's character?

Similar questions seem to invite new theorising, rather than the mere application or adaptation of old ideas. ICTs have caused radical changes both in how societies may come into conflict and how they may manage it. At the same time, there is a policy and a conceptual deficit. For example, the US Department of Defence intends to replace a third of its armed vehicles and weaponry with robots by 2015, but it still lacks an ethical code for the deployment of these new, semi-autonomous weapons.[22] This is a global issue. The 2002 Prague Summit marked NATO's first attempt to address cyber-defence activities. Five years later, in 2007, there were already 42 countries working on military robotics, including Iran, China, Belarus, and Pakistan,[23] but not even a draft of an international agreement regarding their ethical deployment. There is a serious need for more descriptive and conceptual analyses of such a crucial area in applied ethics, and more assessment of the effectiveness of the initial measures that have been taken to deal with the increasing application of ICTs in armed conflicts. The issue could not be more pressing and there is a much felt and quickly escalating need to share information and coordinate ethical theorising. The goals should be sharing information and views about the current state of the ethics of information warfare, developing a comprehensive framework for a clear interpretation of the new aspects of cyberwar, building a critical consensus about the ethical deployment of e-weapons, and laying down the foundation for an ethical approach to information warfare. We experimented with chemical weapons, especially during the First World War, and with biological weapons, in particular during the Sino-Japanese War of 1931–1945. The horrific results led, in 1925, to the Geneva Protocol, prohibiting the use of chemical and biological weapons. In 1972, the Biological and Toxin Weapons Convention (BWC) banned the development, production and storage of bio-weapons. Since then, we have managed to restrain their use and, by and large, respect the BWC. Something similar happened with nuclear weapons. The hope is that information warfare and e-weapons will soon be equally regulated and constrained, without having to undergo any terrible and tragic lesson.

Let us return to the elephants. During the civil war, in the battle of Thapsus (46 BC), Julius Caesar's fifth legion was armed with axes and was ordered to strike at the legs of the enemy's elephants. The legion withstood the charge, and the elephant became its symbol. Interestingly, nobody at the time could even imagine that there might be an ethical problem in treating animals so cruelly. We should think ahead, because history occasionally is a bit petulant and likes to repeat itself. At a time when there is an exponential growth in R&D concerning ICT-based weapons and strategies, we should collaborate on the identification, discussion and resolution of the unprecedented ethical difficulties characterizing cyberwar. This is far from being premature. Perhaps, instead of updating our old ethical theories with more and more service packs, we might want to consider upgrading them by developing new ideas. Like the civilian uses of robots, information warfare calls for an information

[22] The Economist (2007).

[23] Source: *The Wilson Quarterly*, report available online.

ethics. After all, iRobot produces both the *Roomba 700* that vacuum cleans your floor and the *iRobot 710 Warrior* that disposes of your enemies' explosives.

6.1.6 Conclusion

Six thousand years ago, humanity witnessed the invention of writing and the emergence of the conditions of possibility that were going to lead to cities, kingdoms, empires, sovereign states, nations, and intergovernmental organisations. This is not accidental. Prehistoric societies are both ICT-less and stateless. The state is a typical historical phenomenon. It emerges when human groups stop living a hand-to-mouth existence in small communities and begin to live a mouth-to-hand. Large communities become political societies, with division of labour and specialised roles, organised under some form of government, which manages resources through the control of ICTs, including that special kind of information called 'money'. From taxes to legislation, from the administration of justice to military force, from census to social infrastructure, the state was for a long time the ultimate information agent and so history, and especially modernity, is the age of the state.

Almost halfway between the beginning of history and now, Plato was still trying to make sense of both radical changes: the encoding of memories through written symbols and the symbiotic interactions between the individual and the *polis*–state. In 50 years, our grandchildren may look at us as the last of the historical, state-organised generations, not so differently from the way we look at some Amazonian tribes, as the last of the prehistorical, stateless societies. It may take a long while before we come to understand in full such transformations.

References

Anderson, S., and J. Cavanagh. 2000. Top 200: The rise of corporate global power. Institute for Policy Studies 4.

Floridi, L. 1999. *Philosophy and computing: An introduction*. London: Routledge.

Floridi, L. 2012. ACTA—The ethical analysis of a failure, and its lessons. ECIPE working papers 04/2012.

Halper, S. A. 2010. *The Beijing consensus: How China's authoritarian model will dominate the twenty-first century*. New York: Basic Books.

NATO Cooperative Cyber Defence Centre of Excellence. 2013. *Tallinn manual on the international law applicable to cyber warfare: prepared by the international group of experts at the invitation of the NATO Cooperative Cyber Defence Centre of Excellence*. Cambridge: Cambridge University Press.

Ramo, J. C., and Foreign Policy Centre (London England). 2004. *The Beijing consensus*. London: Foreign Policy Centre.

Rawls, J. 1999. *A theory of justice*. Cambridge: Belknap Press of Harvard University Press.

The Economist. June 7 2007. Robot wars.

The Economist. June 2 2012. Morals and the machine.

United Nations. 2011. *State of the world's volunteerism report, 2011: Universal values for global well-being*. United Nations volunteers.
Williamson, J. 1993. Democracy and the "Washington consensus". *World Development* 21 (8): 1329–1336.
Williamson, J. 2012. Is the "Beijing Consensus" now dominant? *Asia Policy* 13 (1): 1–16.