MASTERARBEIT

Titel der Masterarbeit

Emerging trends in machine autonomy:

Defining and regulating intelligent machines

verfasst von

Alexander Kühne, BA

Angestrebter akademischer Grad

Master of Arts (MA)

Wien, 2014

Studienkennzahl lt. Studienblatt: A 066 906

Studienrichtung lt. Studienblatt: Masterstudium Science Technology/Society

Betreuerin / Betreuer: Univ. Prof. Dr. Ulrike Felt

Abstract

Machine intelligence and intelligent software are becoming increasingly integrated into our everyday systems and infrastructure. Today drones, robots and other forms of intelligent machines as well as intelligent software and algorithms are being used in a wide range of applications, ranging from simple maintenance tasks and administrative functions all the way to armed conflict and warfare. Increasingly, these applications replace the old forms of delegation of accountability and responsibility with new modes of delegating to non-human actors. What used to be under the control of a human actor, for instance an investment banker, can now be done by a sophisticated algorithm in a fraction of the time and cost (in terms of labour). The advances in computer technology, microchips, material science and miniaturization in the 20th century have led us to technologies which for the first time in human history truly offer the opportunity as well as the risk of delegation large aspects of our daily lives to non-human actors, be it individual machines or software running on a computer.

This form of machine autonomy has been the focus of popular science fiction for decades spawning many popular myths about intelligent machines such as the famous Artificial Intelligence HAL9000 in the movie 2001 or the intelligent machine network Skynet which according the the 1984 Scifi classic The Terminator would have launched and enslaved humanity in 2010. Other imaginations of autonomy are more realistic and work with the technology available today while yet others seek to develop more and more sophisticated systems in order to reach a certain level of autonomy. The thesis aims to present an overview of the most common and influential imaginations of what machine autonomy means. Ranging from remote controlled to fully autonomous and automatic. Depending on how someone (the public, individual actors) imagine and view autonomy the resulting conclusions, be they fears and hopes or rules and laws, will differ since each imagination of what autonomy means for a machine offer different assumptions and represent a different level of technological sophistication. Once determined, the second part of the thesis will come into play: The question whether a technology such as machine autonomy could and should be regulated. By comparing this technology to weapons of mass destruction (ABC) and by providing a short overview of how these are being regulated, the thesis will try to make a point for a regulatory model of machine autonomy as well as some of the particularities which would have to be taken into account. In the end this thesis will provide an overview over the most dominant views on machine

autonomy, its application to different aspects of society as well as an outlook on possible regulatory measures.

Zusammenfassung

Maschinenintelligenz und intelligente Software werden zunehmend in unseren alltags-Systeme und Infrastruktur integriert.Drohnen, Roboter und andere Formen der intelligenten Maschinen sowie intelligente Software und Algorithmen werden in einem breiten Spektrum von Anwendungen eingesetzt. von einfachen Wartungsarbeiten und administrativen Funktionen bis hin zu bewaffneten Konflikten und Kriegsführung. Zunehmend ersetzen diese Anwendungen die alten Formen der Delegation von Verantwortung und Rechenschaftspflicht an menschliche Akteure durch neue Modi zur Delegation an nicht-menschliche Akteure. Was früher unter der Kontrolle eines menschlichen Akteurs, zum Beispiel ein Investmentbanker, kann nun durch einen ausgeklügelten Algorithmus in einem Bruchteil der Zeit und Kosten (in Form von Arbeits) durchgeführt werden. Die Fortschritte in der Computertechnologie, Mikrochips, Materialwissenschaften und Miniaturisierung im 20. Jahrhundert haben es ermöglicht Technologien zu entwickelt die zum ersten Mal in der Geschichte der Menschheit wirklich die Möglichkeit bieten Arbeit an wirklich autonome oder teilautonome Systeme zu delegieren.

Diese Form der Maschinenautonomie ist seit Jahrzehnten Fokus vieler populärer Mythen über intelligente Maschinen wie die berühmte Künstliche Intelligenz HAL9000 in dem Film 2001 oder die intelligente Maschinennetzwerk Skynet. Andere Vorstellungen von Autonomie sind realistischer und arbeiten mit der heute verfügbaren Technologie, während wieder andere versuchen, mehr und mehr hochentwickelte Systeme zu entwickeln. Die Thesis zielt darauf ab, einen Überblick über die häufigsten und einflussreichsten Vorstellungen von dem, was Maschinenautonomie bedeutet, zu präsentieren. Je nachdem, wie jemand (der öffentlichen oder privaten Akteure) sich Autonomie vorstell werden die daraus resultierenden Schlussfolgerungen zu Ängsten und Hoffnungen oder Regeln und Gesetzen.Der zweite Teil der Thesis stellt die Frage Frage ob eine Technologie wie Maschinenautonomie geregelt werden sollte. Durch den Vergleich dieser Technologie mit Massenvernichtungswaffen (ABC) und durch die Bereitstellung eines kurzen Überblicks darüber, wie diese reguliert werden, wird die Thesis versuchen ein Plädolyer für ein Regulierungsmodell der Maschinenautonomie auszuarbeiten sowie einige der Besonderheiten, die würde berücksichtigt werden sollten. Am Ende wird diese Thesis einen Überblick über die dominierenden Ansichten über Maschinenautonomie, ihre Anwendung auf verschiedene Aspekte der Gesellschaft sowie ein Ausblick auf mögliche Regulierungsmaßnahmen bieten.

CV

Schulausbildung:

2002-2004 Albert Schweitzer Gymnasium, Eisenhüttenstadt
Abschluss: Abitur

1991-2002 Deutsche Schule Moskau, Russische Föderation

Studium:

2004 – 2011 Studium der Politikwissenschaft an der Universität Wien
(Diplomstudium)

2011 – 2014 Masterstudium Science Technology Society an der Universität Wien

Berufliche Tätigkeiten:

04/08 – 01/2010 Projektmanager emailcharity.com
Tätigkeiten:
Konzeption/Koordinierung/Abnahme/Weiterentwicklung
der emailcharity.com Plattform sowohl im Front-end als auch im Back-end
Verantwortlich für zeitweise 3 Mitarbeiter (2 Praktikanten und e in Programmierer)
Betreuung von Firmen/Organisationen/Sponsoren/Usern in allen Bereichen (Key Account Management)
Assistenz der Geschäftsführung
Adminstration/täglicher Betrieb der emailcharity.com Plattform

07/2007 – 12/2007 Praktikum in der Delegation der Europäischen Kommission
China (Peking)
Schwerpunkte:
Mitarbeiter der Press & Information Section:
Organisation von Pressekonferenzen und Events,
Betreuung von Medienvertretern,
Medienrecherche/Pressespiegel,
Redaktion des Delegations Newsletters sowie Webmaster der Delegationshomepage
Allgemeine Öffentlichkeitsarbeit
Design/Korrektur von EU Broschüren und Informationsmaterial für das chinesische Publikum

03/2007 – 07/2007 Redakteur für www.chilli.cc, Österreichs Jugendseite
Ressort: Gesellschaft
Themenschwerpunkte: Informationsgesellschaft, IT,
Computer, Infomationssicherheit, Internet, Online

Überwachung/Vorratsdatenspeicherung, Internet
Kriminalität, Russland

Table of contents

1) Introduction

“I know not with what weapons World War III will be fought, but World War IV will be fought with sticks and stones.”
— Albert Einstein

For as long as humans exist and form societies the matter of thinking about, creating and transforming our environment and our world has always been a matter of human decision making. Terms like accountability, responsibility, risk or danger are terms which we usually attach to human decisions and human made systems and societies. We trust that our governments will make decisions according to certain moral values and ethical considerations. Authors such as Machiavelli, Kant or Marx have attempted to theorize ways in which human societies would act in an effort to better understand what shapes us and our societies. Authors like Machiavelli tended to put their attention on systems of governance and the role of individual actors (for instance the sovereign) in order to better explain and understand their respective societies. With the rise of the industrial age however a new form of societal change had established itself, a force driven by human intellect and innovation, a force we now refer to as science and technology. Of course science and technology had already been part of the human experience for a long time, but it was the end of the renaissance era and the beginning of the industrial revolution when science and technology moved out or the shadows of organised religion and the churches and onto the main stage of society. More and more people started to understand that the technology you posses, how you use and and what you know and understand about the world will directly influence the power and influence you (your country, your government and so on) will be able to wield. The active development of science and technology had become a matter of survival. This became painfully obvious during the first and second world wars which are referred to the first real “wars of attrition” (Attrition Warfare. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Attrition_warfare) in history, where the use of material and modern weapons, rather than the size of armies or the amount of men, was key. It was also during that time the first real weapons of mass destruction were developed and used. Mustard gas and other forms of weaponized gas were used in massive amounts both against enemy troops as well as civilians – while the gas was not always very effective and subject to wind and weather conditions just the mention of a gas attack alone would be enough to strike fear among anyone. Additionally to other forms of weapons of mass destruction like nuclear and biological weapons/agents military technology in general made huge advances: From the development of advanced munitions for war (both anti-personnel and anti-material) to the development of new generations of guns and explosives and the mechanization of large parts of the armies (mechanization in the military sense refers to the use of mechanized infantry, that is troops supported or carried mainly by mechanical vehicles) - technology promised to transform the business of war entirely. Now, in the 21st century, a new form of technology again promises to reinvent what it means to be human and again will give humans new ways of organising themselves. While the technological advances of the 20th century mainly improved already existing systems this time around the outcome could be much more unexpected. With the development of machine learning, machine intelligence and AI humanity has opened up the opportunity to develop machines capable of taking decisions, to receive and understand complex tasks and to generally become more intelligent. How people imagine and enact this complex new technology will determine its outcome on such diverse factors as human/non-human interaction, societal risk factors and the delegation of tasks away from human actorship.

1.1 Research question

Drawing from this introduction and the reasoning it brought up the following research questions can be specified:

1.) What characterizes the most common imaginations of machine autonomy

2.) Can / should there be a regulation of machine autonomy

The first question will help to determine if and which differing imaginations of autonomy exist and which actors are depending on them. As already mentioned in the introduction autonomous machines have been the subject of imagination for some time now and the idea of autonomy itself can be interpreted in different ways. To determine which interpretations dominate the debate is an important part of this thesis. Closely related to the first question is the question which of the defined imaginations of machine autonomy are being realized on a broad scale, meaning which can be expected to be adopted by important decision-makers and actors such as governments and large institutions. The second part of the overarching research on the emerging technology of machine autonomy is the question whether a technology such as this could be regulated in any form. The aim of this thesis is to answer both of these questions by studying the current literature about the differing degrees of machine autonomy and what is being worked on in terms of regulation by institutions such as the UN, the EU or international NGOs such as Human Rights Watch. To provide a clear idea of how autonomous machines and intelligent software are being imagined and realised by different actors is an important first step in understanding the changes that autonomy can potentially bring to our way of life, it will also help to highlight potential risks (imagined and real) and potential uses. This would include such aspects as military vs civilian use of drones and other machines with degrees of autonomy or the debate around the domestic use of such technologies for other purposes they were originally intended for. The insights gained in the first part of the research question will help in thinking about the potential regulation of machine autonomy and highlight some of the particularities this would involve, such as the dual use character of many of the related technologies and technological objects.

1.2 Chapters

The thesis is sectioned in four chapters not including this introduction, each dealing with one aspect of the thesis.

First the underlying theoretical framing and methodology are specified. This includes explanations of the theories, methods and sensitizing concepts/approaches used as well as the selection of sources employed throughout the thesis.

The next chapter deals with the state of the art and the debate around machine autonomy. Here the current development of the technology will be assessed as well as the current dominant imaginations of it. This will also include explanations of relevant actors and how their understanding of machine autonomy relates to the state of the art.

The question of a potential regulation of machine autonomy will be the focus of the next chapter. Here machine autonomy will be compared to earlier technologies with a similar impact. This chapter will try to make a point for a regulation of machine autonomy and some of the particularities of that might come up will be highlighted.

The last chapter will look deeper into some of the underlying issues concerning the regulation and use of machine intelligence – aspects such as the delegation of responsibility, accountability and risk management and communication will play a large role in this chapter.

The conclusion will give a final overview over the different imaginations of machine autonomy and will again try to make a point about the regulation of this technology.

2) Theoretical framing

This thesis will deal with the aspect of machine autonomy: What imaginations of autonomous machines are out there, how do they influence the debate around machine autonomy and how can we as a society regulate the development and use of this powerful technology?
Intelligent machines and artificial intelligence in the form of software are becoming so widespread now that it is time to think about the delegation of responsibility and accountability to non-human actors that these technologies bring. What this means for the risk assessment of such technologies and what this would mean for regulatory approaches is an important aspect. From warfare (Marra and Mckneil, 2012) to our financial systems (Chaboud, Chiquoine, Hjalmarsson and Vega, 2014, p. 2045-2084) down to everyday tasks (Fiorini and Prassler, 2000, p. 227-235) and social responsibilities (Broekens, Heerink, and Rosendal, 2009, p. 94-103) machine autonomy is having a deep impact. Intelligent software and algorithms govern large parts of our modern infrastructures from basic utilities such as electricity and water to public transport (Wren, A. and Wren, D., 1995, p. 101-110), traffic management (Kato, Tsugawa, Tokuda, Matsui and Fujii, 2002, p. 155-161) or even the planning of police patrols and actions. As a society, we are already relying on a large amount of delegation to non-human actors such as these intelligent machines and software. Therefore it is important to understand the technological capabilities and opportunities as well as the potential risks and conflicts of interest that certain imaginations and practises of machine autonomy can bring with them. The aim of this thesis is to present an overview of the current debate around machine autonomy and to present a case for the regulation of machine autonomy on an international level. The field of Science/Technology&Society provides a useful framework for analysing the underlying mechanisms of agency and the move towards non-human agency, the different imaginations of machine autonomy and how they influence actors and decisions. The possibility for a regulatory framework around intelligent machines is another important aspect and will be talked about in the second part of the thesis. According to Sheila Jasanoff,
"the proposition that the ways in which we know and represent the world (both nature and society) are inseparable from the ways we choose to live in it" (Jasanoff, 2004)

The co-production of technology and society, in terms of how we design and develop machines capable of autonomous actions shapes our view of reality when it comes to machine autonomy. The same is true for other aspects of life such as transportation, logistics or basic infrastructure. It has become a negotiation of sorts between what is technically possible, what is morally and ethically acceptable to us and what can be realized economically. Another important aspect here would be the concept of prescription and description brought forth by authors such as Akrich or Bijker. It claims that technologies have a certain set of assumptions that are prescribed to them. For example, a technology might assume a certain type of use or user which reflects in the way this technology is developed and applied. In the same way the theory implies a phase of description to technology which signifies the way the technology is actually used and adopted by whoever is relevant to it.

"A large part of the work of innovators is that of "inscribing" this vision of (or prediction about) the world in the technical content of the new object. I will call the end product of this work a "script" or a "scenario." (Bijker and Law, 1992)

Developers of intelligent machines and software prescribe a certain way these technologies should be used and more importantly by whom they should be used.

"Thus, if we are interested in technical objects and not in chimerae, we cannot be satisfied methodologically with the designer's or user's point of view alone. Instead we have to go back and forth continually between the designer and the user, between the designer's projected user and the real user, between the world inscribed in the object and the world described by its displacement. " (Bijker and Law, 1992)

They might assume that these technologies will become legitimate parts of a nations arsenal and that they would be used according to rules and regulations (which would have to be invented first).

"Once technical objects are stabilized, they become instruments of knowledge... If it also chooses categories used in other socioeconomico-political network, then the knowledge it produces can be "exported." - "Data" can thus be drawn from the network and transmitted elsewhere... However, the conversion of sociotechnical facts into facts pure and simple depends on the ability to turn technical objects into black boxes. In other words, as they become indispensable, objects also have to efface themselves. " (Bijker and Law, 1992)

In presuming this the makers of these technologies prescribe these technological artefacts as surveillance technologies, or unmanned patrols or even warriors - on the other hand, the "end users" might describe the same artefacts very differently, modifying, enhancing or simply using them in unforeseen scenarios. The prescriptions themselves do not come out of thin air either, they are also based on a co-production process.

Popular movies and science fiction have already defined numerous instances of autonomous machines, shaping a certain perspective on these systems. Developers, politicians, the public all have a certain view of how these technologies are supposed to operate. This translates into concrete policies and opinions about the use of machine autonomy. After the militarization and weaponization of drone technologies we are now facing the widespread introduction of drones into civilian life be it as service drones, security drones or for science and research – this offers us the opportunity of not only studying the emergence and shift of a arguably transformative technology but also the luxury of anticipating and counteracting potential harm by pro-actively thinking about regulation and safeguards. In order to avoid the mistakes of the past where new generations of military technologies where adopted and utilized without really planning for the consequences, one has to be able to anticipate potential outcomes (International Red Cross, 2003).

One thing that clearly distinguishes humans from almost every other animal on this planet is our capability to envision, imagine and manipulate our physical reality through the use of tools, technology and science. From the earliest examples of tool use by our ancestors to the first homosapiens using traditional techniques and methods to hunt or farm, establishing the first ideas of how "things could or should be" to ancient civilisations building elaborate societies, the imagination and use of technology has always played a central part in our development as a species.

Technology as a term is so widely used that it is hard to find a definition of it that most people would agree on, but at its core technology is a mechanism which we use to alter, improve or protect our own natural abilities and bodies and which can be described as the imagined future realised by physical objects and applied knowledge. When our ancestors first discovered tool use the imagination of being able to augment ones own bodies in order to better hunt, kill or protect was novel. Simply picking up a stone and using it was what gave our ancestors a considerable edge over all other animals. This of course also required the ability to envision ones own future and deduct future events from clues and information of today, an ability that only few other organisms on earth posses(Self awareness. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Self-awareness#Developmental_stages). From simple tool use to simple technologies of tradition and traditional hand-craft more elaborate ideas started to emerge.

Individuals and institutions would emerge who would take it on them to further develop, guide and improve whatever tradition of technology and science a social group would have, increasingly gaining influence over the centuries and millennia of human development (Library of Alexandria. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Library_of_Alexandria).

"STS scholarship allows us to analyse the complex interplay of scientific and technological developments with other dimensions of social life. Sociotechnical imaginaries call attention to the fact that visions of future developments in S&T almost inevitably bring with them wider visions of social futures, of risk and benefit, and of the collective good. In this respect, sociotechnical imaginaries are instruments of co-production." (Harvard Kennedy School. (n.d.). Frequently Asked Questions about Sociotechnical Imaginaries. Retrieved November 6, 2014, from http://sts.hks.harvard.edu/research/platforms/imaginaries/imaginaries-faqs/)

At its core STS research seeks to uncover the invisible links and interactions between what we call "science", what we make and build as "technology" and the way we live and imagine to live in the future as "society" – STS, in a nutshell, seeks to make visible how we shape our life, societies and environment.

“Sociotechnical imaginaries are at once descriptive of attainable futures and prescriptive of the kinds of futures that ought to be attained. As an influential part of the currency of contemporary politics, these imaginaries have the power to shape technological design, channel public expenditures, and justify the inclusion or exclusion of citizens with respect to the presumed benefits of technological progress. Given the political salience of such imaginaries, and the risks and instabilities that inevitably accompany their realization, understanding how they are formed and implemented is necessary to any serious exploration of what the sociologist Ulrich Beck has called a “cosmopolitan” vision of inter-cultural collaboration and coexistence" (National Science Fund, 2007)

Autonomous machines imply a delegation of tasks to non-human agents, in this case robots and other forms of machine intelligence. Tasks can range from mundane or dangerous things such as disaster relief and clean-up to more demanding tasks such as security and patrolling up to very complex tasks such as full on warfare. The delegation of accountability away from human towards non-human actors is another important aspect of this thesis. The question if a regulation of machine autonomy is feasible will serve to illustrate some of the aspects that such regulation would bring up. A non-human actor can be defined as any non-human entity capable of taking action and making decisions which can also include software, algorithms, objects and machines. When applying the basic idea of the sociotechnical imaginary, the question "how do we want to live in the future?" to the topic of this thesis questions about accountability, risk and responsibility become even more apparent. Throughout the centuries our societies have developed elaborate rules and laws concerning the delegation of responsibility and therefore the acceptance of accountability to other human actors or human institutions. The risks of delegating tasks to other human beings is understood and accepted, regulation and lawmaking provide the framework which allows for the transfer of responsibility and accountability without risking ones own legal security. If, as some imaginations of machine autonomy suggest, technological development will progress in the direction it is now we will have to deal with questions of delegation to non-human actors sooner or later. But while delegating to other humans can be argued as being part of our basic make-up or just how humans work together, delegating to machines increasingly complex tasks presents a new challenge for regulators and policy makers.

2.1 Current debates around machine autonomy

"STS, as practised in academia today, merges two broad streams of scholarship. The first consists of research on the nature and practices of science and technology (S&T). Studies in this genre approach S&T as social institutions possessing distinctive structures, commitments, practices, and discourses that vary across cultures and change over time. This line of work addresses questions like the following: is there a scientific method; what makes scientific facts credible; how do new disciplines emerge; and how does science relate to religion? The second stream concerns itself more with the impacts and control of science and technology, with particular focus on the risks that S&T may pose to peace, security, community, democracy, environmental sustainability, and human values. Driving this body of research are questions like the following: how should states set priorities for research funding; who should participate, and how, in technological decision-making; should life forms be patented; how should societies measure risks and set safety standards; and how should experts communicate the reasons for their judgements to the public?" (Harvard Kennedy School. (n.d.). What is STS. Retrieved November 6, 2014, from http://sts.hks.harvard.edu/about/whatissts.html)

Theoretical frameworks such as the Actor-Network approach by Latour tend to focus on the first "flavour" of STS, uncovering the underlying connections and principles that guide how we create "facts", amass knowledge and how science works as an institution in itself as well as part of a larger system of institutions. On the other hand thinkers such as Sheila Jasanoff have coined the theoretical framework of "sociotechnical imaginaries", a theory better suited for analysing more concrete and tangible issues such as how individual key technologies were developed, what kind of ideas and fears surround certain science and technologies or how a certain technology or scientific practise become accepted into society. Other approaches such as the pre and description of technological objects attempt to analyse the process of how technologies or even individual inventions are seen and understood during their conception, their development and afterwards how the users and the public understand and use or "describe" such an object.

In the case of this thesis the latter more technology centred approach is highly favoured since the debate around machine autonomy requires a deep understanding of the technological principles and capabilities involved in the research and development of autonomous machines, software and AI.

As mentioned before in the previous section it is important to note that there are many different imaginations of the term autonomy and each imagination comes with its set of pre and descriptions of technologies, political, economical and ethical considerations (or lack thereof) which effect the way machine autonomy is enacted by whoever holds a certain imagination of it.

Since these approaches rely on analysing the conflict between human vs non-human actor and agency the question to what degree a technology can be considered to be dominated by the one or the other becomes a major issue. For instance, while the functions and features of a drone might be predetermined by regulations which rely on political, social and ethical considerations while these considerations are based on their imagination of machine autonomy before they are translated into concrete policy. Additionally this section will include an analysis of effects the differing degrees of autonomy might have on some central issues concerning any military technology, namely accountability, risk-management and communication, transparency and other factors. Sensitizing concepts and approaches provide valuable perspectives to further analyse the issues surrounding machine autonomy. The question who can be made accountable when a semi or fully autonomous drone malfunctions or otherwise does something unintended to cause harm is a novum as its analysis has to include human as well as non-human actors sometimes as abstract as an algorithm. Questions of accountability will become more important as machine autonomy gets more pervasive. Who is to blame when something goes wrong, will it be the engineers who build the robot, the programmer responsible for the algorithms, the company producing the machine or maybe the military official responsible for deploying the machine in the field?

This is closely tied to risk-management and risk communication, as well as transparency issues. Any government or military has to justify the use of such technologies or hide it from its population. While drone technologies can be used quietly for military purposes in distant regions, using them on your own population will incur much stronger responses and fears.

Emerging technologies tend to have a few defining aspects which characterize how the technology progresses or in which “direction”. When computers where developed in the late 1940s and early 1950s, they where considered large scale machinery, it was expected that only a few mainframe-size machines would be needed to completely satisfy all computational needs which might arise -

"There is no reason anyone would want a computer in their home." -- Ken Olson, president, chairman and founder of Digital Equipment Corp., 1977

Olson, who at this time was one of the most influential business leaders in computers and office

equipment failed to see one distinctive trend in computers: the ongoing miniaturization and advances in semiconductor technology made it possible to build and offer more compact and affordable personal computers – something which we now take for granted.
When the internet was developed, it was never intended to completely replace all international communication infrastructure – it was merely designed as a fail.-proof communication network for a limited number of participants, first only military and later academic. (Leiner, B., Cerf, V., Clark, D., Kahn, R., Kleinrock, L., Lynch, D., ... Stephen, W. (n.d.). Origins of the Internet. Retrieved November 6, 2014, from http://www.internetsociety.org/internet/what-internet/history-internet/brief-history- internet#Origins). Yet again with the introduction of personal computers and affordable communication hardware and infrastructure (modems, public service providers, private service providers) the internet became the most important and efficient means of communication within 10 years of its popularization in the early 1990s. For machine intelligence, one of these defining trends is autonomy. The term autonomy itself is a rather vague one and its definition depends on who you ask and what exactly autonomy refers to. While autonomy might be associated with freedom, self sustaining behaviour or independence in the context of relationships or formal agreements between nations, if used in a different context the term autonomy can carry very different notions. Autonomy in a technological sense refers to certain functionalities a machine can have which allow said machine to act on its "own". What does that mean exactly though, can a machine really act independently from any prior instructions or programming? In the world of drones and other related machines autonomy can have different degrees of meaning to different actors and audiences. It is not only interesting to look at what exactly these different degrees of autonomy are on a technological basis but rather to look at how different groups or actors in society frame the issue of machine autonomy. There is a distinction to be made in the differing imaginations of autonomy that exist in society. For example there is a clear difference between how the topic of autonomy is approached by the mainstream media and the public as opposed to actors actually involved in the development of said technologies. While popular accounts of machine autonomy often emphasize potential risks and speculate on future developments of the technology to allow for human-level intelligence or autonomy and behaviour comparable to a human being, current technological development is clearly favouring a compromise between "full" autonomy and "semi" autonomy. "Semi" in this context refers to certain security precautions or features which get implemented into a system which in most cases means having a human actor "on-the-loop" - the human actor is able to check, interrupt or launch a machines actions instead of having the machine

do it completely without any human interaction.

Another important distinction to be made on machine autonomy would be between functional autonomy and autonomy that affects certain core issues or features of a machine such as its weapon systems. Many systems under development at the moment do not offer a fully autonomous solution but rather offer to automate certain tasks – this might be things that could free up human actors or reduce the risk of human error during high precision tasks. Navigation and reading sensor data comes to mind, this would be considered a functional type of autonomy as it allows the machine to function without direct human interaction but still retains final control over the machines consequential actions. This kind of technology would be employed with drones tasked to patrol an area or provide sensory data for example about weather conditions or as sea&air rescue devices capable of detecting survivors at sea. A different case would be a system designed with an action in mind that would generally be considered dangerous, such as controlling a weapon system. In this case autonomy suddenly becomes a very different issue in itself as it no longer functions as a technology designed and enacted to provide help and assistance to humans as envisioned by the famous author Asimov (Three laws of robotics. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Three_Laws_of_Robotics). Instead of simply freeing up a human actor from mundane or dangerous tasks a machine with this decisive type of autonomy would be regarded very differently. While functional autonomy is often seen and understood in a very optimistic utopian idea of intelligent helper-machines, autonomy designed for warfare and control is generally seen under a very dystopian light heavily influenced by popular science fiction.

2.1.1 Selection of sources

Source and materials used in this thesis range from official government reports and studies to interviews and analyses of experts on the topics of machine autonomy, robotics and AI as well as studies published by non-governmental organizations, advocacy and lobby groups and articles published in the media. A major focus in the selection of source material was the question of the representation of framing of autonomy in the source. Since one aim of this thesis is to understand better the different framing of autonomy, focusing on this aspect makes sense. This selection of sources also represents a wide selection of involved actors and decision makers and allows us to analyse an actors stance towards the topic. This becomes even more interesting and useful if we combine this with the analysis of other factors involved in the debate on machine autonomy such as accountability or risk. The sources also give insight into how different actors portray the involved risks, what risks are being promoted or not, how the question of accountability is discussed or which "framing" of machine autonomy an actor prefers. It should also be noted that some of the sources are aimed at certain audiences, stakeholder or decision makers while others are aimed at a larger public. For example the report by Human Rights Watch, which is one of the bases for the categorization of autonomy into three distinct sub-categories, is aimed at politicians and decision makers. It allows us to understand the current state of the debate (at the time of publishing) and the concerns and risks which are being addressed. Another report, created for the EU FP7 programme, focuses on the potential uses of drone technologies and autonomous machines for the EUROSUR border program – in this case the report assumes a certain definition of machine autonomy and already prescribes certain uses of the technology. On the other side of the spectrum media articles and popular science /scifi depictions of machine autonomy tend to focus on either the potential use or the potential risks of such technologies – both utopian and dystopian scenarios are popular and can be argued to a certain extent. Analysing such sources is useful because it shows us the fears and hopes that are associated with machine autonomy as well as some of the potential pre and descriptions of individual technological objects such as the US predator drone.

The imagination of autonomy and autonomous machines in the mainstream society is heavily influenced by popular science fiction such as the HAL9000 computer from the movie 2001: A Space Odyssey where a human crew gets trapped by their on board artificial intelligence controlling their spaceship and has to fight it to survive – in this case, the AI is portrayed as being ruthless and efficient without any empathy. Another popular depiction of fully autonomous machines taking over humanity is the classic 1980s movie “Terminator” which came out in 1984 and presented the idea of fully autonomous robots in humanoid form and appearance combined with a powerful AI called Skynet. On a similar note, the 1999 movie “the Matrix” shows a dystopian vision of the future where humanity has been enslaved by its own intelligent machines due to the fact that at some point their AI decided that humans would be wasteful and could be used more efficiently if harvested as energy sources.

What all of these Scifi imaginations of machine autonomy and AI have in common is the fear of unforeseen consequences, an overpowering of the human race by the machines and a lack of empathy and “humanity”. These fears get reflected on a regular basis in popular media accounts of autonomy which often emphasize potential for fully autonomous machines even if the technology is not there yet.

“Popular understandings of technologies, especially in the field of robotics, are quite different from the on-the-ground reality of technological capabilities. Lucy Suchman’s blog, Robot Futures*, has looked at* numerous examples *of such misconceptions. Nor is it unexpected that the terms “autonomous” and “unmanned,” so frequently applied to robotic technologies, are misnomers and obscure the very real human labour involved, from producing and operating hardware (from* flying *to* interpreting data*) to coding software. From Marx’s Capital to the more recent work of Shoshana Zuboff, Lucy Suchmann and many others, research has shown that advances in automation and robotics do not so much do away with the human but rather obscure the ways in which human labour and social relations are reconfigured."* (Elish, 2012)

Another large part of the sources used in this thesis are sources produced by different non-governmental organisations such as Human Rights Watch or different research institutions such as Stanford Law School or the NYU. These sources include reports from regions where today’s semi-autonomous and remote controlled drones are being used such as Pakistan, Yemen or Afghanistan amongst others. These sources tend to be more balanced in the way that they portray the technological capabilities and the degree of autonomy actually present in the machines. These sources also include interviews with experts in the fields of robotics, machine learning, military strategy, technology and risk assessment and others. One thing that becomes apparent quite quickly when studying these interviews is that most current development of machine autonomy seems to favour the middle-approach of developing semi-autonomous systems. Instead of creating the fully autonomous machine portrayed in the popular media it seems to be the case that most developers try to include at least some form of human agency in their systems, mostly choosing to include a “veto-power” which allows a human operator to interrupt a machine at any point or might even require the monitoring human agent to actively allow certain actions such as firing a weapon or entering certain areas.

The third category of sources used in this thesis are sources taken from official government bodies, mostly from the US and the EU due to the fact that these parties have a relatively open policy of access to information and are very active in developing drone programmes. The US for instance are mostly concentrating their efforts on developing military-grade drones and robots for a range of applications and with the aim of relieving human soldiers of dangerous and mundane tasks and extending the strategic and tactical capabilities of the armed forces. The EU is currently concentrating their efforts on developing drones and legislation aimed at using theses machines to protect the European Unions borders – especially after the human cost of migration into the EU becomes ever more apparent EU officials are pushing for a system of drones patrolling the Eus outer borders with a focus on the Mediterranean Sea. According to the project EUROSUR (DPA, 2013) and PERSEUS the aim is to develop a fleet of drones capable or patrolling the sea region and detecting refugee boats before they even get close to the Eus coasts – this will also include cooperation with border nations to the EU especially in the northern.-African region.

“PERSEUS addresses the call for an integrated European system for maritime border control. Its purpose is to build and demonstrate an EU maritime surveillance system integrating existing national and communitarian installations and enhancing them with innovative technologies. By means of two large scale demonstrations PERSEUS will prove its feasibility and will set the standards and grounds for the future development of EU maritime surveillance systems." (PERSEUS-FP7. (n.d.). Retrieved November 6, 2014, from http://www.perseus-fp7.eu/?page_id=17)

Looking at the different sources and analysing the different framing angles and prescribed values communicated by the sources can help in highlighting the conflict lines between the different imaginations of machine autonomy. Just as opponents of such technologies tend to emphasize the dangers and potential risks of full autonomy the supporters and/or makers of drone technologies and associated technologies will not hesitate to point out that full machine autonomy is in fact not what they are developing and that their machines would merely help the human actor do his/her job more efficiently. Caught in the middle are governmental institutions and regulatory bodies which depend on both sides to provide them with factual information in order to be able to manufacture some sort of consent between all parties on how exactly machine autonomy should be used in our societies and how this could or should be regulated. The methods employed and the type of sources used put a focus on the analysis of the imaginations of different key terms and technologies such as autonomy, risk, responsibility, agency or delegation. Therefore it is important to understand that none of these terms provide clear definitions per se in every case, rather these terms are used and framed in different ways according to who is using them for what purpose. A government report might use a different definition of the term ”autonomy” when they talk about the drones they use then a technologist talking about the same drones, who would then add that the term “autonomy” is rather loose and also includes remote controlled machines or machines which are not actually autonomous but simply feature a limited auto pilot. The government report might have certain goals to achieve or a certain image to uphold, while a report of an NGO in the same issue might be focusing on the imagination of definition of “risk” when it comes to the use of autonomous machines. In that case, the term “autonomy” could be framed or imagined in a much broader way to suggest to the reader a much broader range of potential risks then a report would suggest that had a more realistic imagination of autonomy. This is not to say that certain imaginations are better or worse or more or less “scientific” or “real”, they are simply different because they are co-produced out of different necessities and motivations by a range of different actors.

2.1.2 Methodology

As already mentioned in the previous chapter on the selection of sources most of the material can be divided into categories of how the source frames autonomy and what other key aspects are mentioned in the document. Certain terms such as “autonomy”, “regulation”, “risk” or “responsibility” where defined as key terms and the sources included in this thesis chosen accordingly. During the research it became clear that the most common representation of autonomy can be divided in three distinct categories or the degree of automation/autonomy. Source material was therefore analysed under this aspect in order to create a broad understanding of how autonomy is framed in different communities and for different actors. In addition to the general framing of autonomy focus was put on the key terms talked about earlier in order to further understand the defined key terms relevance. This was later used to reflect upon the theoretical approaches and sensitizing concepts that were choosen for this thesis. By combining the framing angles of autonomy with some of the associated secondary factors such as risk/opportunity I was able to further reflect on the aspects of machine autonomy and create an argument for the regulation of machine autonomy. This argument is based on observations made in researching the regulation of ABC weapons and other types of regulated ammunitions and reflecting the observations made on the catalogue of framing/secondary factors associated with autonomy. The conclusion further reflects upon the theories and STS concepts applied throughout the thesis and attempts to create a list of particular aspects when considering the regulation of machine autonomy with a special STS perspective in mind. To analyse the source material and reflect upon its content the idea or concept of engagement in STS proved to be very helpful. The approach emphasizes a method of analytical thinking which at the same time is forward-thinking and has social implications and social change in mind. While other approaches tend to focus on the purely academic pursuits of analysing and working with materials the engagement approach allows for a broader approach with a focus on actively thinking about the future.

Autonomy is the central point of conflict for most of the debates around machine autonomy as it raises a host of secondary issues: Autonomy implies that an actor is able to make decisions and take action on its own, without needing confirmation or being controlled. The term autonomy can be interpreted in several very different ways which each have different implications on regulation and risk assessment. A major part of this thesis will be to define the ways in which autonomy is imagined and how these imaginations differ from one another. From a widely science fiction influenced idea of machine autonomy which is widespread in the mainstream media and public debate to more realistic and "down to earth" imaginations of autonomy as an emerging technology and up to wide ranging imaginations of a future society using a wide range of intelligent machines. Autonomy is seen and debated on very different levels in diverse communities. One of the major concerns in all of the debates however is the underlying question of agency, particularly human vs non-human agency as defined by scholars such as Bruno Latour or John Law in their ANT approach. Law describes the approach as

> *"…a disparate family of material-semiotic tools, sensibilities and methods of analysis that treat everything in the social and natural worlds as a continuously generated effect of the webs of relations within which they are located. It assumes that nothing has reality or form outside the enactment of those relations. Its studies explore and characterise the webs and the practices that carry them."* (Law, 2008)

For this reason official documentation, government and NGO reports as well as expert analyses and interviews are used to create three distinct categories or machine autonomy. Further these degrees are analysed by applying the theoretical frameworks discussed in this section with a focus on matters of human agency, responsibility and accountability as well as risk. The final part of the thesis will attempt to make a point for a international model of regulation of machine autonomy and will include a short comparison of regulatory models already in existence for weapons of mass destruction and some thoughts on regulating machine autonomy.

2.2 Sensitizing Concepts / Approaches

This section of my thesis will deal with the different concepts, theories and approaches which can lend a hand at analysing the issue of machine autonomy. Many of the introduced approaches have similar trains of thought on questions of human vs non-human agency. The sociotechnical imaginaries approach can help to analyse complex systems of technological, social, cultural, ethical, political and economical aspects which in combination form a holistic way of thinking about a certain technology. One example of this would be the dual use of drone technologies as both military technology and civilian applications. On the matter of semi-autonomous and fully autonomous drone platforms a similar pattern can now be observed as more and more components for a fully autonomous drone become commercially available.

"...new technologies may not only lead to new arrangements of people and things. They may, in addition, generate and "naturalize" new forms and orders of causality and, indeed, new forms of knowledge about the world... technologies may generate both forms of knowledge and moral judgements." (Akrich, 1992, p. 205-224)

A approach which combines many of the ideas of agency, responsibility, delegation and co-construction are the socio technological imaginaries of which Sheila Jasanoff is a well known proponent.

"As an analytic concept, "sociotechnical imaginary" straddles the line between structure and agency: it combines some of the subjective and psychological dimensions of agency with the relative hardness of technological systems, policy styles, organizational behaviours, and political culture. The methods best suited to the study of sociotechnical imaginaries therefore are the interpretive research and analytic methods in STS that help illuminate the structure-agency relationship. Although these methods are not specific to the analysis of imaginaries, they can be used in ways that are especially attuned to this concept, e.g., to the ways in which imaginaries frame and represent futures, relate past and future time, enable or restrict action, and naturalize certain ways of thinking about possible worlds. " (Harvard Kennedy School. (n.d.). STS Methods. Retrieved November 6, 2014, from

http://sts.hks.harvard.edu/research/platforms/imaginaries/ii.methods/methodological-pointers/)

The materials and sources used throughout this thesis all fall into one of these categories – official documents, studies and NGO documents are great sources for analysing the discourse and framing around machine autonomy by providing examples of both official representations of the issue as well as analyses of these representations by third parties. For instance while many of the US army documents talk about the delegation risk and responsibility as a challenge they still maintain the position that the claimed features and capabilities of their drones such as accuracy of attacks are correct. Documents by an NGO on the same kind of US drones will also talk about the delegation of risk and responsibility but will cite cases where things have gone wrong or where risks have not been met with proper precautions and will paint a very different picture of what is in fact the same situation. The choice of content of discourses and the way different content is weighted also gives insight into how an issue is being "imagined" by an actor or a whole group of actors – politicians will tend to "imagine" a use-case scenario where the technology simply works as in a black box model and will communicate this to their constituents while an analyst would approach the issue with a different set of goals such as finding potential risks. It is important to note that "imagination" in this sense does not mean the actual mental representation of the issue in a persons head but rather the way an issue is framed in speech and imagery to create a certain common imagination. In this sense sociotechnical imaginaries are closer to a communication device then what the name might suggest a way of planning technological innovation.

2.2.1 Engagement in STS

A significant part of this thesis deals with the question of the role of STS in public engagement and in dealing with the fallout of emerging technologies. Machine autonomy conjures a whole set of topics to be discussed ranging from the risk assessment and communication to the delegation of responsibility and accountability up to the question of political agenda setting. When science gets involved in politics or individual scientists become involved this is often met with criticism and scepticism by both the scientific community and the larger public – there are however also proponents of a more engaged approach to social science and in particular STS research. In his 2001 speech addressed to the president's plenary at the 2001 annual meeting of the Society for Social Studies of Science in Cambridge, Massachusetts, Bijker makes a broader argument for the engagement of STS scholars in the public sphere.

"That science and society movement fruitfully merged with the socio-logical studies of science in society, roughly at the same moment that the crucial impulse from the strong program and the sociology of scientific knowledge (SSK) came. This merger moved the agenda along from studying science (as a subsystem or function) in society to studying the culture of science and technology. Over the past two decades, the resulting research has yielded many insights into the detailed processes of scientific knowledge production and technological development. But at the same time, much of this work has not addressed the normative, political, or practical consequences of these insights. To put it strongly: most STS researchers have not raised the question of what to do with these insights. The STS agenda has been largely agnostic as to the normative and political issues related to the application of STS insights." (Bijker, 2003)

Bijker continuous to explain the importance of public involvement of STS scholars as public intellectuals. Specifically Bijkers mentions the potential for STS to tackle the some of the most complex global issues that we face today. Bijker even goes a step further by arguing that STS research now has the opportunity to move from a purely descriptive and theoretical approach towards a more applied and normative way of doing STS.

"Current societal problems urge a broadening of the STS agenda. The big issues of social order, international peace, local and social security, national and religious identity, and democracy

should be addressed again, but now on the basis of detailed insights in scientific knowledge and technical machines: the agenda can shift from studying the culture of science and technology to studying technological culture. So when I argued previously that STS can contribute to the agenda of democratizing technological culture, this really was an understatement. I would rather observe that STS has created the basic ingredients of this agenda and then propose that it now should work to realize its potential." (Bijker, 2003)

Bijkers employs the phrase "intellectual" and more specifically "public intellectual" to describe what he understands as the ideal role of the STS researcher. Drawing on writings of Bourdieu Bijker re-enforces his argument that STS and STS researchers should and could function as a modern form of public intellectualism. This would also include the involvement of such influences in the political sphere of agenda seting and policy making. What Bijker means by that a stronger normative involvement of science and scientists as public figures that provide guidance, orientation and discuss the “what do we want” and “how do we want to achieve it” of societal and scientific progress.

"In his analysis of the history of intellectuals, Bourdieu (1991) saw philosophers, artists, and scientists moving between the two poles of the ivory tower and politics: he sketched such pendulum movements from the period of the Enlightenment to the beginning of the twentieth century. The modern intellectual, according to Bourdieu, emerges only when writers, artists, and scientists discover that they can engage in politics without immediately threatening the autonomy of their own worlds (of science and art)." (Bijker, 2003)

Bijkers understands STS research as a multi-purpose approach which can be utilized for different goals: Be it academic, focused on scholarly and theoretical pursuits or more focused on the policy making and analysis work and even STS research which Bijkers demands should be geared towards long term societal goals, agenda setting and directly influencing the public sphere.

"STS now has three distinct routes to the future. First, there is the Academic Highway, with scholarly journals, monograph series, chairs, undergraduate programs, and graduate schools. Second, there is Policy Street, with potentially income-generating STS work, directly useful to

the public and private sectors. Third, I make a plea for Democratization Boulevard, on which studies are carried out that combine long-term academic agendas with clear political and societal engagement. This route leads to less immediately applicable studies than work on Policy Street, but the studies are more directly inspired by political concerns than work on the Academic Highway." (Bijker, 2003)

Reflecting on Bijkers position on the engagement of STS scholars as actors of social change and policy setting, the topic of this thesis should be understood as the third approach described by Bijker. Public engagement of STS scholars is not new either, it has happened before in other emerging fields such as nanotechnology. (Delgado, Kjølberg, and Wickson, 2011, p. 826-845) In a similar fashion this thesis aims to uncover the particularities of the emerging technology of machine autonomy as well as assess some of its potential political, social and economical consequences.

2.2.2 Framing angles & framing of autonomy

In a social science and STS context, framing is the social construction of certain phenomenon in society. This involves how the phenomenon is presented ("framed") by mass media, lobby groups, politicians, civil society or other important actors.

An important aspect of framing is the control over individuals understandings of certain words or phrases, which can influence a persons mental imagination of an issue or can influence public discourse about an issue. Framing can both be used to simplify, interpret or steer opinion as well as to frame communication practises and content between actors. Different interpretations of framing also exist, for instance n psychology where framing presents equal alternatives rather than encouraging or discouraging certain viewpoints as is practised in the social science/ communication context. A "frame" in the social science context is a collection of (mental)images, stereotypes (sometimes re-enforced by the framing) and anecdotal evidence that can lead an individual, institutional actor or whole parts of society to accept certain interpretations of an issue. Framing modifies the imagination of an issue, it can also pre-condition certain responses and reactions to it. (Druckman, 2001, p. 225-256)

The term autonomy is a good example of how different framing of the issue can lead to very different outcomes. Most peoples first contact with the concept of autonomous machines is most likely framing used by the mainstream media and popular culture. This is heavily influenced by science fiction, philosophy and local cultural and social conventions. The understanding and expectations towards robots and other machines is very different in the US-American public versus the Japanese public, where the former tends to focus on the development of robots and other machines for military and security purposes while the latter is actively funding development of robots and other (semi)-autonomous machines for expressed civilian purposes and so called companionship robots used to care for the elderly or disabled. The framing of what autonomy means to people in those societies, while based on the same sources (media, popular culture, social norms) is very different from each other.

Once an individual starts researching the vast field of robotics and machine autonomy one will develop a more nuanced understanding of its central term autonomy. By studying documents published by developers and researchers involved in developing new generations of autonomous machines it becomes apparent that the framing of these machines in the “professional” sphere tends to be more realistic and feature-oriented. Instead of focusing on a rather broad definition of machine autonomy as a black box, the prevalent framing here makes clear distinctions between different forms of autonomy and usually tends to emphasize the need for control. Especially when communication with the larger public through interviews and other forms of media involvement, developers and researchers make sure to make a point of discouraging the development of fully autonomous systems – a very interesting distinction from the "mainstream" definition of autonomy which often paints the picture of the fully autonomous robot capable of becoming a real danger to human society. Similar things can be said about the third big area of framing the subject, which is the sphere of governance, policy making and regulation. Policy makers and regulators are interested in the details, so framing in the regard takes into account the three degrees of autonomy defined as "man on", "man-in" and "man-out-of" the loops systems where the systems are designed with or without a human operator in mind capable of interrupting or at least monitoring the machines actions. The emphasis on a clear definition of a machines capabilities away from marketing claims or framing intended to sway the public or regulators is very important here as well as the message to the public that their government or their institutions are aware of the issue and are considering the different scenarios with care.

2.2.3 Risk analysis / Risk communication

A major factor for technology assessment, the emerging discipline of future studies and policy advisers in general is the question of risk: Risk is a diffuse term and can be approached from various perspectives, for instance how risk is being communicated, what is being seen / accepted as risky, what actually poses risk on a rational level but is not seen as risky and so on. Risk is both a very “firm” term in the sense that everybody can understands what risk means, that it involves danger to themselves, their surroundings or their way of life – at the same time, risk is being seen differently in severity, consequence and relevance to ones own life-circumstances. This makes risk communication a very complex issue as it not only needs to include rational factors such as hard scientific data but at the same time account for subjective factors such as the psychology of risk assessment, standards in media communication and how it handles risk (for instance a authoritarian state will have different ways to communicate with its population than a democracy). In the concrete cases of future studies/emerging technologies/scenario building risk of course plays a central role which has links to and influences all other factors such as public reactions, policy requirements, potential fall backs or fail safes and many more. Risk, how it is being calculated and how it is being dealt with defines to a large degree how we as a society deal with the phenomena that future studies try to understand.

Paul Slovic tries to address this complexity in his 1999 paper “Trust, Emotion, Sex, Politics, and Science: Surveying the Risk-Assessment Battlefield" published in the journal "Risk analysis".

“Risk management has become increasingly politicized and contentious. Polarized views, controversy, and conflict have become pervasive. Research has begun to provide a new perspective on this problem by demonstrating the complexity of the concept “risk” and the inadequacies of the traditional view of risk assessment as a purely scientific enterprise.” (Slovic, 1999, p. 689-701)

Slovic goes on to argue that while there are real dangers out there that societies have to deal with, the associated risk or how this risk is being seen and dealt with is entirely socially constructed. Slovic describes risk assessment as “inherently subjective” pointing out that not only “rational” facts play a role but that psychological, social, cultural and political factors are just as important to consider when analysing risk assessment/risk communication. Slovic also makes the important observation that democratic institutions tend to breed a certain amount of distrust – making risk communication in a democratic society even more problematic.

The definition of risk plays another very important role. As Slovic point out, whoever is in control of defining what risk is, what it is not will be in control of of the issue is being tackled on a rational level. Depending on the agenda of this stakeholder, a solution might favour security, risk mitigation, cost efficiency or even adhere to popular or populist political demands against better judgement.

"Defining risk is thus an exercise in power. Scientific literacy and public education are important, but they are not central to risk controversies. The public is not irrational. Their judgements about risk are influenced by emotion and affect in a way that is both simple and sophisticated. The same holds true for scientists. Public views are also influenced by world-views, ideologies, and values; so are scientists' views, particularly when they are working at the limits of their expertise." (Slovic, 1999, p. 689-701)

In order to better be able to deal with risk in a society, Slovic argues for more public participation and a more transparent decision-making process. Slovic also acknowledges the limitations of a science based approach and and points out that the importance of participation and openness is especially relevant when dealing with public acceptance and trust in the decision being taken. (Slovic, 1999, p. 689-701)
In a related approach, "The Social Amplification of Risk: A Conceptual Framework" by Kasperson et al asks why risk events gain more attention then others often unrelated to the actual rational risk. In some cases, risk-events which are relatively unlikely to occur (for instance an extinction level event such as a meteor strike) gain widespread media and public attention sometimes even demanding direct action from politicians. At the same time, other risk events which do have a high a likelihood of impacting society on a broad scale (such as climate change, overpopulation or resource shortages) do not seem to elicit as much reactions.

"One of the most perplexing problems in risk analysis is why some relatively minor risks or risk events, as assessed by technical experts, often elicit strong public concerns and result in substantial impacts upon society and economy." (Kasperson et al., 1988, p. 177-187)

Similar to Slovics argumentation in”Trust, Emotion, Sex, Politics, and Science: Surveying the Risk-Assessment Battlefield”, the authors also refer to the subjectivity of risk, its involved factors such as psychology of risk assessment, sociological and cultural particularities or risk perception. The authors introduce the concept of amplification which they describe as a two level process consisting of the transfer of information about the risk event and the response mechanisms in place in a society to react on such information.

“The main thesis is that hazards interact with psychological, social, institutional, and cultural processes in ways that may amplify or attenuate public responses to the risk or risk event. A structural description of the social amplification of risk is now possible. Amplification occurs at two stages: in the transfer of information about the risk, and in the response mechanisms of society.” (Kasperson et al., 1988, p. 177-187)

In her 1998 book titled "Risk (Key Areas) Deborah Lupton gives an introduction into risk perception by comparing it to the medieval term and understanding. During the period, risk was seen as a natural occurrence similar to a natural disaster or the concept of "fate" (a predetermined path) or "fortune" up to the present day where risk is categorized in numerous categories such as the area the risk concerns (ecological, security, health and so on) and risk is largely understood as something that can be minimized or even avoided given enough knowledge, preparation or security. (Lupton, 1999) Lupton continues by looking at a few approaches to risk in the current social/scientific context, focusing on the cultural-symbolic approach of Mary Douglas, the "risk society" approach of Anthony Giddens (Giddens, 1990) and Ulrich Beck (Beck, 1992) as well as Michel Foucaults perspective on risk which he refers to as governmentality perspective. Lupton attempts to categorize these approaches into "weak" up to "strong" social constructionist approaches, a weak position being described as a "realist view of risk" – an emphasis on the perception of the risk and the "real" risk. On the other hand, the relativist position also referred to as the strong constructivist position focused on risk as a subjective phenomena influenced by each persons individual surroundings, social contacts, power relations and the social and societal mechanism that inform of risk, thereby "creating" the awareness in the first place.
Beck (Beck, 1992) and Giddens (Giddens, 1990) coin the phrase of the "risk society", an approach that tends to be in the middle of the realistic or weak position and the largely subjective strong position. In the risk-society, risks and hazards have become so common that they become accepted

parts of our everyday life. This is compared to early modernity, where most peoples lifes in Europe depended on industrial processes and the division of labour. Individuals were usually not expected to know or care about risks, there were little or no insurances, work safety or other safety regulations that we take for granted. Today, though, an individual is expected to care for himself and minimize a whole range or perceived risks.

Some of these risks are very real, such as the risk to develop cancer in ones lifetime, while other might be rather unrealistic but most people would still fear them more, for example dying in a plane crash. A modern individual (speaking of western/European values) nowadays needs to have insurances, save some money for bad times, wear a seat belt, own a mobile phone (so you can always be reached) and so on. We perceive such risks as real and relevant, since they affect us now but what about risks that will likely affect us in a few years, or even just our children?

Lupton describes the relationship between risk and subjectivity by dealing with the issue of lay people and how they construct risk according to what they know and in the context or their own life and setting. Lupton argues that risk may be perceived and dealt with differently depending on where it is being discussed. In the public sphere, the media, politics and so on risk has already been studied in extent, however the personal subjective perspective of risk for each individual is the focus of Luptons work.

2.2.4 Applied socio-ethics

The study of robotics and other emerging technologies already has a certain tradition within the framework of the European Union. Areas of research include bionics, the field of replacing and enhancing parts of the human body, robotics (robots as companions, but also imagined forms of interaction and relationships between human/non-human actors). A special focus is also put on robots in combination with automation, autonomy and the militarization of such technologies. What makes the approach brought forth in this EU-Deliverable so useful for this thesis is the concept used to study these technologies. The authors employ a methodology hey describe as "applied socio-ethics" which can be grouped with other similar approaches more known in the STS sphere such as ANT or the assemblage approach. Technology in the socio-ethics approach is being understood as a complex system or artefact, similar to the notion of the assemblage. Ideas of pre/description also play a role in their definition, as does co-construction and human&non-human agency. What separates the applied socio-ethics approach from other more theoretical approaches is the "applied" part of the concept, which seeks to translate theoretical assumptions on the role of technology, social conventions and ethical beliefs into real policy and tries to compare it to existing real world examples of regulation, policy making and agenda setting.

"On the one hand, one has more general assumptions about e.g. our way of living, our shared values and future perspectives, the role of technology in society, about the relation of society and technology, of human beings and machines. Think, in this connection, of all assumptions made in the process of technology development and design which are linked to our ideas of a good way of living, about a desirable work(place), the proper way of conducting warfare, about the compatibility of work and private life, the boundaries between public and private. An important question concerns the compatibility of these underlying, and often only implicit, assumptions with the EU Charter of Fundamental Rights, with its appeal to notions of liberty, human dignity, personal identity, moral responsibility and freedom. And how compatible are these assumptions with notions of social responsibility and justice as well as solidarity? Do concepts and models of technology exclude or disadvantage human beings with regard to their gender, age, ethnicity, educational background or sexual orientation?" (ETHICBOTS, 2008)

The study took place in the framework of the EU "Ethicbots" project seeking to research ethical and

moral issues when developing, regulation and using robots in all areas of life including companionship, health care, senior care, security, surveillance and warfare.
In the first part of the project titled “Analysis of the State of the Art in emerging technologies for the integration of human and artificial entities" (D1) a more general approach to emerging technologies is taken in order to determine relevant fields of study, part 2 of the project carries the title "Methodology for the identification and analysis of techno-ethical issues" and seeks to establish theoretical frameworks and methodology for analysing such emerging technologies, in the fourth (part sequence seems to skip part 3) part "Analysis of national and international EU regulations and ethical councils opinions related with technologies for the integration of human and artificial entities" existing regulations and policies are presented and analysed while the fifth part the authors focus on actual technologies and their place in society. (ETHICBOTS, 2008)

“(...)This also includes the need to analyse those concepts, theories, models and approaches in robotics, AI systems and bionics which are used to model the relation between user and machine (e.g. master-slave, caregiver-infant, partner, pet-owner). Their underlying and often hidden assumptions have to be critically evaluated with respect to their potential cultural and societal impact. Here, a basis for assessment is provided by the EU Charter of Fundamental Rights but also by socio-ethical reflections on social responsibility and possible disadvantages for groups of human beings.” (ETHICBOTS, 2008)

The authors pay special attention to the concept of mutual co-construction and tackle some of the central questions brought up by most STS research, namely how the work of scientists and technology researchers is influenced by their socio-economic background, the frameworks and structural constrains thy might face as well as what part notions of gender, education or age play

“How far are assumptions underlying technological developments impregnated by the experiences, interests and by the social, cultural and educational background of technology researchers? In which way are they influenced by research policies and agendas, etc.? How do these assumptions fare vis-à-vis notions of social responsibility, exclusion or disadvantage issues arising in connection with, say, gender, age, ethnicity, educational background or sexual orientation?” (ETHICBOTS, 2008)

At the end of their methodological approach the authors emphasize again the importance of understanding robotics as devices in the making rather than ready-made and determined in purpose. In accordance with the concept of mutual co-construction and similar to Akrichs Pre/Description concept the authors also make the point that technology as complex as this is a product of social negotiations, ethical and moral categories and development of such technologies should include analysis of these factors to help develop technologies that are beneficial to society.

"This analysis should also be an intrinsic part of ethical work on the level of technology design. As robots are not regarded as ready-made products of engineers but as contested devices and technologies in the making we need ethical reflections to support us in the development of technologies which will support our common values and the prosperity of our social and political life. In this context we need to ask: How far do the assumptions made in the research and development process reflect on the needs and values of the EU citizens, of the EU everyday users?" (ETHICBOTS, 2008)

2.2.5 Sociotechnical imaginaries

A theoretical approach that combines many of the before mentioned aspects and concepts is the approach made popular by such authors as Sheila Jasanoff. In this approach, the process imagining, producing and establishing a certain idea of the future, certain technologies or fields or research is analysed and retold as a puzzle of co-produced ideas about the potential uses, risks, costs etc. of a technology or technological progress. By combining the imagined ideas of several actors, by analysing what actors are involved in a sociotechnical imagination of an issue it becomes possible to paint a clearer picture of how "we" as a society in general negotiate technological progress and change.

"(...)*In 2005, the international biomedical research community was shaken by the announcement that Dr. Hwang Woo-Suk, a celebrated South Korean scientist, had fabricated results and violated ethical norms in work that had led to highly acclaimed publications in leading journals such as Science. Hwang, who initially came to fame for animal cloning research, hit global headlines when he announced that his research team had succeeded in growing human blastocysts via nuclear transfer, and had also created 11 patient-tailored stem cell lines from them in a remarkably efficient manner (Hwang et al.2004, 2005). These studies quickly won international recognition as a major scientific breakthrough with significant economic and social implications. Koreans, however, saw them above all as demonstrating South Korea's world-class scientific capabilities, paving the way to the country's "next-generation growth engine industries." Even after it was revealed that Hwang and his co-workers had committed serious scientific and ethical misconduct (SNU-IC 2006; KNBC 2006), one poll indicated that nearly 70 percent of South Koreans still wanted to give them a second chance if they had the "indigenous technology" to produce human blastocysts from cloned embryos (CBS Radio 2006)."* (National Science Fund, 2007)

The idea of sociotechnical imaginaries is that technologies are developed and used according to certain imagined futures in which such technologies would exist or specifically not exist in some cases. This can include imaginations of potential uses, involved risks (imagined and real), cultural fears and particularities, costs and investments needed, questions of sustainability of a technology or even moral and ethical considerations. STI also analyses imaginations to better understand how involved actors can shape these imaginations through risk communication, public awareness campaigns and other forms of engaging the public. Finally, STI offers explanations and theories on how we negotiate technological change in our societies and how our societies deal with the changes science and technology brings. Sociotechnical imaginaries differs from other common social science concepts used in studying collective beliefs in a few key terms and concepts. Differences between theoretical concepts and theories tend to be hard to define, still Jasanoff and other author have created their model of sociotechnical imaginaries by employing methods of boundary work and defining certain terms.

"

- *Master narrative: possibly the term that comes closest to sociotechnical imaginaries, but more monolithic and unchangeable, and more closely bound to a re-narration of history; not tied to a notion of specific goals to be achieved.*
- *Discourse: mainly focused on language and lacking the normative/prescriptive dimension of sociotechnical imaginaries.*
- *Ideology: more attuned to power and social structure than sociotechnical imaginaries, but also more static and lacking the connotations of reaching and striving; usually does not include material constructs.*
- *Policy: refers to formal or tacit programs of action, not to the underlying rationale or justification, which is often provided by sociotechnical imaginaries.*
- *Plan: has the intentionality of sociotechnical imaginaries, but usually refers to near-term futures with specific, designated goals (e.g., a plan to build a highway) and is usually a product of formal institutional authority.*
- *Project: usually involves a single, targeted, technological endpoint such as the moon landing, the "cure for cancer," or the sequencing of the human genome; such projects reflect underlying sociotechnical imaginaries.*
- *Public reason: tends to be explicit rather than tacit, and is shaped by institutionalized relations between political authorities and citizens, in contrast to sociotechnical imaginaries, which can be articulated and pressed for from below. "* (Harvard Kennedy School. (n.d.). Frequently Asked Questions about Sociotechnical Imaginaries. Retrieved November 6, 2014, from http://sts.hks.harvard.edu/research/platforms/imaginaries/imaginaries-faqs/)

3) State of the art: Types of autonomy

3.1 Relevant categories of machine autonomy

The way in which an actor imagines a technology, how it works, what it can and cannot do, how it should or should not be employed largely determines the theoretical approach this actor develops towards the technology. If the question concerns the regulation of a technology or its use, imagined risks, cultural fears, moral and ethical considerations all play a role in forming a framework of values and variables which in the end form a whole “imagination” - for example, an imagination of what an autonomous machine is, what it can do and what risks are involved. An actor tasked with writing a regulatory framework on machine autonomy would come to very different results each time if a different imagination of what autonomy is on a technical level was used. Just by assuming different levels of sophistication in the technology the analysis would require a completely different focus on things like involved risks, fears, cost or potential misuse. A regulatory model based on a conservative imagination of autonomy, based on the “man-In-the-loop” assumption for example would not consider the risk of an (military) autonomous machine causing collateral damage to be of high priority because the assumption that an autonomous machine is in fact not fully autonomous but still controlled by a human actor in is most vital functions such as firing a weapon. This would mean that in this version of imagined machine autonomy, the risk of a out of control machine would be considered minimal due to the fact that autonomy itself is assumed to have certain safeguards implemented. If the regulator however employs a more open-ended imagination of the term machine autonomy, which includes true autonomy without human oversight, the risk of a machine getting out of control would be considered differently. This is why is is very important for the understanding of this thesis to be aware of the different degrees of machine autonomy in debate and their differences in regard to human intervention.

Fully autonomous weapon system do not exist yet on the battlefield but they are being actively developed by a number of national and private enterprises. A fully autonomous weapon would be able to launch, navigate, acquire targets and launch attacks at them without any intervention from human actors either acting by a pre-programmed scheme or fully autonomous by using advanced algorithms. The difference between a robot that has been pre-programmed and one that is using advanced algorithms is that the first can only function in pre-defined contexts and environments

while the latter could function in unknown terrain and situations by dynamically interpreting sensor information and using its algorithms as "logic circuits" to reach decisions.
According to a model by Human Rights Watch and the International Human Rights Clinic when we talk about remote controlled weapon systems of today we have to make 3 distinct categories: (Human Rights Watch & International Human Rights Clinic, 2012)

3.1.1 "Man-IN-the-loop" systems

The first category or remote controlled weapons are weapons system designed with a certain degree of automation, such as an autopilot, but still require human interaction to fire a weapon for example. These systems are referred to as "man-IN-the-loop" systems because the human actor has to actively trigger the weapon. Such is the case with the Predator and Reaper drones currently in use by the Us in Afghanistan, Pakistan, Yemen and possibly other regions. Target identification with these systems is generally based on the pilots impressions rather than automated systems.
When talking about these kinds of "unmanned" systems, we are actually talking about remote controlled weapons rather than autonomous systems. Since these kinds of systems differ little from older established weapon platform which use automation (Cruise missile. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Cruise_missile) the perceived risk of using such technologies seems to be relatively small. The fact that there are no human pilots on board make these kinds of remote controlled drones attractive from various perspectives: For the military, the risk of loosing pilots is practically zero, which also means a lower risk of investing in training and hardware. This also means a lower rate of reported casualties to the own population as well as less media coverage compared to troops on the ground. All of this implies a lower risk assessment by the general public of a nation that uses such a technology as very little consequences become apparent and there does not seem to be any relevant threat of malfunction from a remote controlled machine compared to more sophisticated autonomous systems as it is essentially still being controlled by a human actor. (Deri, 2012) Tyler Wall and Torin Monahan deal with the issues such drones bring up in their paper "Surveillance and violence from afar: The politics of drones and liminal security-scapes" by taking a closer look at the deployment of aforementioned Predator and Reaper drones in regions such as Pakistan, Afghanistan, Yemen, some border regions of the US and urban environments.

“As surveillance and military devices, drones—or ‘unmanned aerial vehicles’—offer a prism for theorizing the technological politics of warfare and governance. This prism reveals some violent articulations of US imperialism and nationalism, the dehumanizing translation of bodies into ‘targets’ for remote monitoring and destruction, and the insidious application of militarized systems and rationalities to domestic territories and populations. In this article, we analyse the deployment of drones within warzones in Afghanistan, Iraq, and Pakistan and border zones and urban areas in the USA. What we call ‘the drone stare’ is a type of surveillance that abstracts people from contexts, thereby reducing variation, difference, and noise that may impede action or introduce moral ambiguity. Through these processes, drones further normalize the ongoing subjugation of those marked as Other.” (Wall and Monahan, 2011, p. 239-254)

Tyler and Monahan describe the these drones as part of a broader surveillance system or logic which is based on the collection of mass data on a large number of subjects, often without any clear indicators on what to scan for or what is defined as problematic. The authors understand drones as complex systems rather that a singular technology which involves a certain form of surveillance which they describe as a form of risk management. In collecting data and monitoring targets the drones create a logical network , enabling interpretation of certain targets as network nodes that pose a increased risk in whatever context the surveillance is taking place. So for instance, during a large demonstration or social unrest drones would be able to hover over a area detecting movement in the crowd, targeting especially active individuals and thus giving operators, police or whoever else would be involved a hint at who might be considered a leader in such a spontaneous movement. Engaging this “risky” node then would serve to collapse the whole network around it or at least help in slowing down communication, actions or collaboration between “nodes” should this be desirable for the security forces (or whoever else holds the monopoly on force). Tyler and Monahan describe this logic as similar to existing systems of positive identification (of targets) as known threats. Tyler and Monahan also understand drone technology in the context of militarization, especially in combination with surveillance which they describe as a dynamic between technological myths and efficiency. The Predator and Reaper drones used by the US are being presented as technologically superior, preciser and accurate which in turn helps in their rapid adoption by security forces around the world. The way these technologies are used also depends where they are being used, Tyler and Monahan describe these areas as “security landscapes” which all have their own rules of engagement, social, ethical and political particularities and goals which in turn reflects on targeting practises, the readiness to engage a target or the acceptance of the use of drones in general. For instance a drone deployed in the border areas or Afghanistan and Pakistan might be witness to a group of people, seemingly men, gathering in an open space – the drone or the operator of the drone might interpret this as a "risky" node in this risk network and order a drone strike on said group. In reality, gatherings like this are usually called "Jirgas" and are meetings of village elders and other influential people of the region and serve as political meetings for solving issues of the community. At the same time, the US targeting practises emphasize the risk of larger gatherings of people anywhere in the area as a risk to US and allied troops which has led to several unfortunate incidents. As a reaction, such Jirgas are now often announced to the US forces to avoid such attacks and as a reaction to civilian casualties the US has decided to reframe their statistics to include all male individuals within a certain age as potential "enemy combatants" rather then civilians. (Becker and Shane, 2012)

Monahan and Tyler compare the recent increase in the usage of drone technologies to the established military strategy of "verticality" or "technologies of verticality". According to this principle, a military force should always strive for a tactical/strategic advantage by seeking the means to attain higher ground (quite figuratively) over its enemies. In the past this has been done in many different form and with many different technologies throughout the ages, Castles, Walls, Turrets and other structures come to mind but also planes, rockets, helicopters, satellites etc. and most recently drones. Compared to the air superiority planes and bombers provide drones can be used on a smaller scale to control smaller regions or regions which are too densely populated and require more localised approaches. Drones combine flexibility, affordability and efficiency – a very attractive package for any military actor.

While conflicts of the past where based on a variety of factors such as land mass, number of available soldiers, resources or geography in todays modern technology dominated world the only way to assert state superiority in military terms is with the use of superior technology. This combined with the need for to answer new tactics of asymmetric warfare that more and more conflicts involve require a technological solution that is quick to deploy, can scale up or down according to the current situation and ideally does not require much man-power to operate.

In addition to to combat missions drones can and are being used for other purposes. One major additional use of drone technology is in the security sector basically replacing human actors as guards and security personnel for mundane and repetitive tasks such as patrolling or standing guard.

With communication technology included in such drones, an operator can easily control vast stretches of a border and even interact with potential border crossers. This is especially relevant in the US context where the United States are more and more determined to guard their southern border to fight of illegal immigration. In a similar fashion, drones can be deployed in the urban landscape as guards and patrols especially in areas that are too dangerous for humans actors to police or not relevant enough such as industrial areas or business parks.

An important observation of Tyler and Monahan is the “risk-network” approach in which they describe the surveillance logic of a drone based surveillance network. In this system, drones serve as actors who collect information on single or group targets which is used to visualize or otherwise accumulate a “risk-network” of individual nodes which can be interpreted as more or less risky according to the criteria in place. So for example, a network of drones can monitor an area in Afghanistan of Pakistan watching individuals and groups on the ground. According to the US targeting practises for their so called “surgical strikes” and larger gathering of people poses an increased risk to US and allied troops, as do individuals being seen carrying “cone shaped objects” which can be interpreted as weapons such as rocket-propelled grenades (RPGs). A deployed drone network can determine such risk-nodes dynamically and on the spot or simply act as a sensory network for a remote base of operations. Tyler and Monahan make the point that this creates a situation where, in any situation, individuals and groups of individuals are singled out and determined as risky on the base of singular observations and narrow observations which could also simply fit in the context of the situation, for instance a group of males could be seen carrying cone shaped objects determined as weapons which would provoke a drone strike on the location when in reality the cone shaped objects could be building materials or other harmless objects.

3.1.2 "Man-ON-the-loop" systems

Secondly, weapon systems can be designed with an enhanced amount of automation and autonomy by enabling them to acquire targets autonomously. In this case, the weapon system has the capabilities to identify and track targets on its own. In order to fire a weapon, the operator has to give his or her OK – such systems are referred to as "man-ON-the-loop" systems since the operator can interfere with the systems actions at any moment (veto power)
An example of such a system would be the sentry robots used by the South Korean forces at the demilitarized zone (DMZ) on the border of North/South Korea. (Samsung Techwin SGR-A1 Sentry Guard Robot. (n.d.). Retrieved November 6, 2014, from http://www.globalsecurity.org/military/world/rok/sgr-a1.htm). These sentry robots are able to scan for targets using high-tech cameras, far surpassing anything the human eye would be capable of. Targets can be automatically aquired and tracked up to a distance of 2km away. In the event of a detection the sentry is able to alarm an operator who then decides if the sentry bot should fire at the target, warn it or take no action – the sentry as an effective range of 2km with its 5.5mm machine gun and 40mm grenade launchers.
Israel uses another variant of a man-on-the-loop system, its Iron Dome is a network of radar and anti-munitions systems capable of intercepting rockets and artillery shells launched at Israel from the Gaza strip. The Israeli government claims that each intercepting missile has to be launched manually by an operator, experts however doubt this as this would require the operator to react and order a launch within a few seconds of the detection of an approaching munition (Broad, 2013).
Compared to the first category of autonomy in unmanned vehicles, the second degree of autonomy involves much more technology/algorithm based solutions and less involvement of a human actor. Most notably movement and target acquisition are automated , targeting principles are no longer based on a humans perception of a situation but on pre-defined factors and patterns. This could include movement patterns, certain shapes such as cone-shaped objects which could resemble a weapon, detection of heat signatures to identify hidden targets (for example under a roof or behind a wall) or other triggers – it should be noted however that many of these targeting practises are being used today by human pilots of remotely controlled systems as well. (Stanford International Human Rights and Conflict Resolution Clinic (IHRCRC) & Global Justice Clinic (GJC) at NYU School of Law, 2012) The importance this case lies on the algorithms used, other technical equipment and depending on the veto-powers can still remain relatively human-centric in its functioning.

Analysing such systems and theorizing rational and potential subjective risk-assessments of 2nd degree autonomous systems is a much more complex task compared to assessing "simple" remote controlled systems. There is a wide range of accompanying factors at play in this very complex assemblage of a technology, many of which are entirely subjective and emotionally laden. Questions of accountability come into mind: In a case where a system acts at least partially autonomous to outside stimuli such as movement, the problem of potential failures and accidents arises. When a remote controlled (human controlled) weapon system is involved in some unfortunate incident, its quite straight forward to blame the pilot or the commanding official – a partially autonomous system however has more "actors" inscribed into its function as a remote controlled system.

An "inscribed actor", in this case, could be part of developing the algorithm used to govern such a system: From software engineers, algorithm designers, the host company, the contract-giver to the person actually putting such a system into an active state, it is a very complex task to determine who has what influence over how an algorithm or associated technology will work and who should be blamed for potential errors or unforeseen consequences. There are no precedents in human history where we allowed a technology to take over such a complex task as manoeuvring a battlefield and possibly identifying and killing anything that is programmed as "the enemy".

While most popular media tends to portray machine autonomy as going in the direction of full autonomy (man-out-of-the-loop systems) allowing to a fully autonomous machine capable of acting on its own inscribed instructions while utilizing any degrees of autonomy and machine learning the machine is outfitted with, most current development seems to focus on developing machines which assist the human operator rather than making him or her completely obsolete. Instead of having fully autonomous drones and other machines which carry a high risk of malfunction and unintended behaviour current and next generation autonomous machines are developed to allow for less human interaction while still retaining a minimum of human oversight and in most cases a veto power over the machines actions by a human operator. Instead of automating critical tasks such as firing a weapon thus far development has focused on automating navigation (autopilot, GPS assisted navigation), assist in targeting and many other tasks which take up valuable time and resources from a human pilot and often carry a risk of error which in part can be avoided by utilizing a machine. On the other hand, the fully autonomous machine framing popular in the media does not seem to be the goal of most developers according to the information available on current development. On the other hand fully autonomous systems are absolutely in the realm of the technologically possible and are being worked on on a limited scale. There are many applications which would benefit from using fully autonomous machines especially in an environment not suitable for humans and out of reach from remote control (either because the connection is not possible at all or too slow or unresponsive) – space exploration comes to mind as well as deep sea exploration, in fact companies like Planetary Resources are already developing concepts for space vehicles capable of exploring our solar system and mapping out asteroids for later mining missions. (Technology - Planetary Resources. (n.d.). Retrieved November 6, 2014, from http://www.planetaryresources.com/technology/). While the most concerning developments of military drones and drones used for surveillance and security purposes will be found in the field of semi-autonomous machines in the near future, full autonomy is just around the corner and is something that needs to be closely monitored to avoid any unwise developments towards creating fully autonomous weapon systems.

3.1.3 "Man-OUT-OF-the-loop" systems

Finally there are those systems that would allow for completely autonomous decision-making. Such systems are still under development and are far from perfect. Such systems would be able to autonomously launch, navigate, identify and acquire targets and also fire the weapons. They would only be controlled by the programming and whatever fail safes that might be put into the system. The operator would not be part of the decision-making any more and would only serve as a passive observer at best. The drones used by the US administration in their "targeted killing" campaign and "surgical strikes" clearly belong to the first category since they require active participation of a human actor to launch, navigate and fire their weapons. Targets are chosen by the operators, with the drones only providing technical help in tracking/identifying targets. The decision to fire is reserved for the human actor who is governed by rules of conduct, national and international laws and regulations and can be held accountable if necessary.

If, as is to be expected, governments start using fully autonomous systems in their military and security forces, the situation changes dramatically and so does the potential risk analysis and risk perception and communication. Without any clear rules or precedents for accountability and transparency, the use of fully autonomous drones is an unknown risk to both experts and lay-people alike. Technically, a system that can automatically manoeuvre and navigate is nothing special nowadays, fully autonomous systems however would add functionalities which would allow these weapons platforms to independently attack and kill or destroy targets – depending on the purpose of a system, this can mean killing humans or destroying non-human targets such as telecommunication infrastructure or the energy grid. As opposed to first and second level systems, the third degree of autonomy brings to the table a range of unknowns: Since such platforms are based on artificial intelligence governed by algorithms, their behaviour might not be fully predictable. Such an algorithm would be extremely complex, would have to be able to dynamically adapt to unforeseen situation on the battlefield and still provide reliable results. At the same time, the algorithm would act as an inscription device for the drone – since the drone is not piloted by a human actor, all considerations concerning the co-construction of moral/ethical/legal boundaries of such a technology would have to be inscribed into this algorithm, as it would be the only actor available to control the drone. Apart from the actual construction of the system (its physical components) its programming, the algorithm, would be the only thing tying the weapon platform to any form of human accountability.

While still in a developmental stage, the robots shown at the 2013 International Aerial Robotics Competition show the direction in which autonomous weapon platforms, drones and other system could go in the coming years.

From the Wikipedia Page on the IARC:

"The International Aerial Robotics Competition (IARC) began in 1991 on the campus of the Georgia Institute of Technology and is the longest running university-based robotics competition in the world. Since 1991, collegiate teams with the backing of industry and government have fielded autonomous flying robots in an attempt to perform missions requiring robotic behaviours never before exhibited by a flying machine.[1] In 1990, the term "aerial robotics" was coined by competition creator Robert Michelson to describe a new class of small highly intelligent flying machines.[2] The successive years of competition saw these aerial robots grow in their capabilities from vehicles that could at first barely maintain themselves in the air, to the most recent automatons which are self-stable, self-navigating, and able to interact with their environment—especially objects on the ground. The primary goal of the competition has been to provide a reason for the state-of-the art in aerial robotics to move forward.[3] Challenges set before the international collegiate community have been geared towards producing advances in the state-of-the-art at an increasingly aggressive pace. From 1991 through 2009, a total of six missions have been proposed. Each of them involved fully autonomous robotic behaviour that was undemonstrated at the time and

impossible for any robotic system fielded anywhere in the world, even by the most sophisticated military robots belonging to the super powers." (International Aerial Robotics Competition. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/International_Aerial_Robotics_Competition)

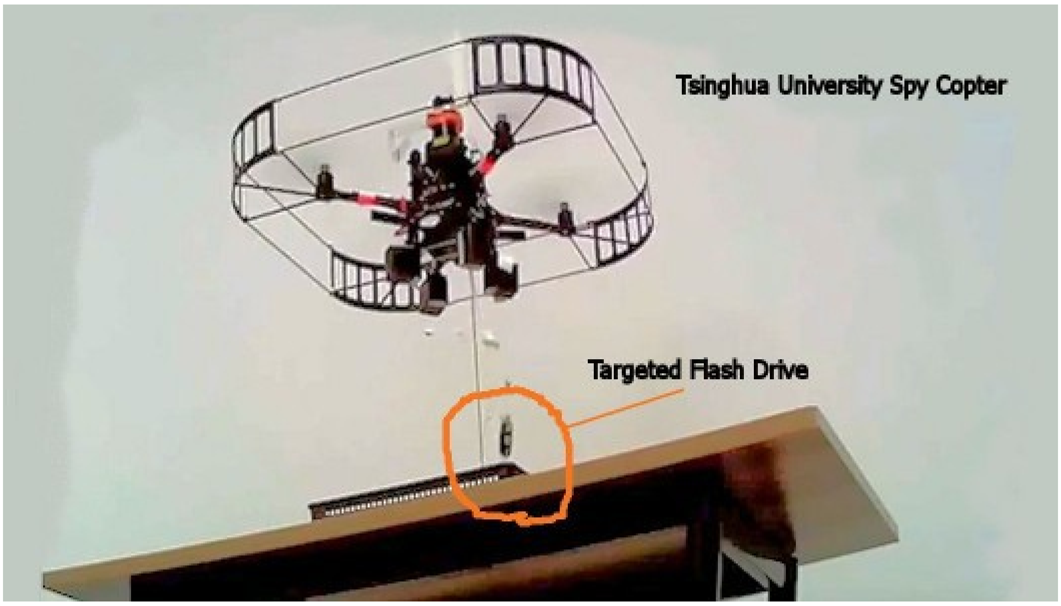

Illustration 1: Winning quadcopter design from Beijing Tsinghua University was able to navigate fully autonomous in close quarter space

At this international robotics competition which has been taking place since 23 years student from all over the world are tasked to develop, build and demonstrate innovative solutions in robotics for certain specific tasks. In this years competition, which took place in the US and China the team from Beijing Tsinghua University won the event by constructing a fully autonomous quad-copter like drone which was able to fulfil the set mission.

This years competition scenario called for the construction of a flying drone platform which would have to enter a fictional building representing a bank in the European Union. The drone would have to enter the building through a broken window and navigate a set of office rooms and corridors autonomously without any remote control or corrections by the construction team whatsoever. Since this mission represented a “break and enter” mission, the goal was for the drone to obtain a USB-Key located in the office and replace it with a similar looking key to avoid suspicions of theft. The key was said to contain sensitive data which, in the wrong hand,s would lead to the collapse of the banking system. In addition to being able to navigate and locate the Usb Key, while carrying a replacement key, the quadcopter drone would also have to detect and avoid several security measures in the office including laser-beams and motion detectors. The performance of the drone and the successful completion of the mission can be seen on video (Knightkimi [Knightkimi]. (2013, August 4th). *IARC 2013 mission 6 completed – Tsinghua University*. Retrieved from https://www.youtube.com/watch?v=B-iZE_Nn52w&feature=player_embedded).

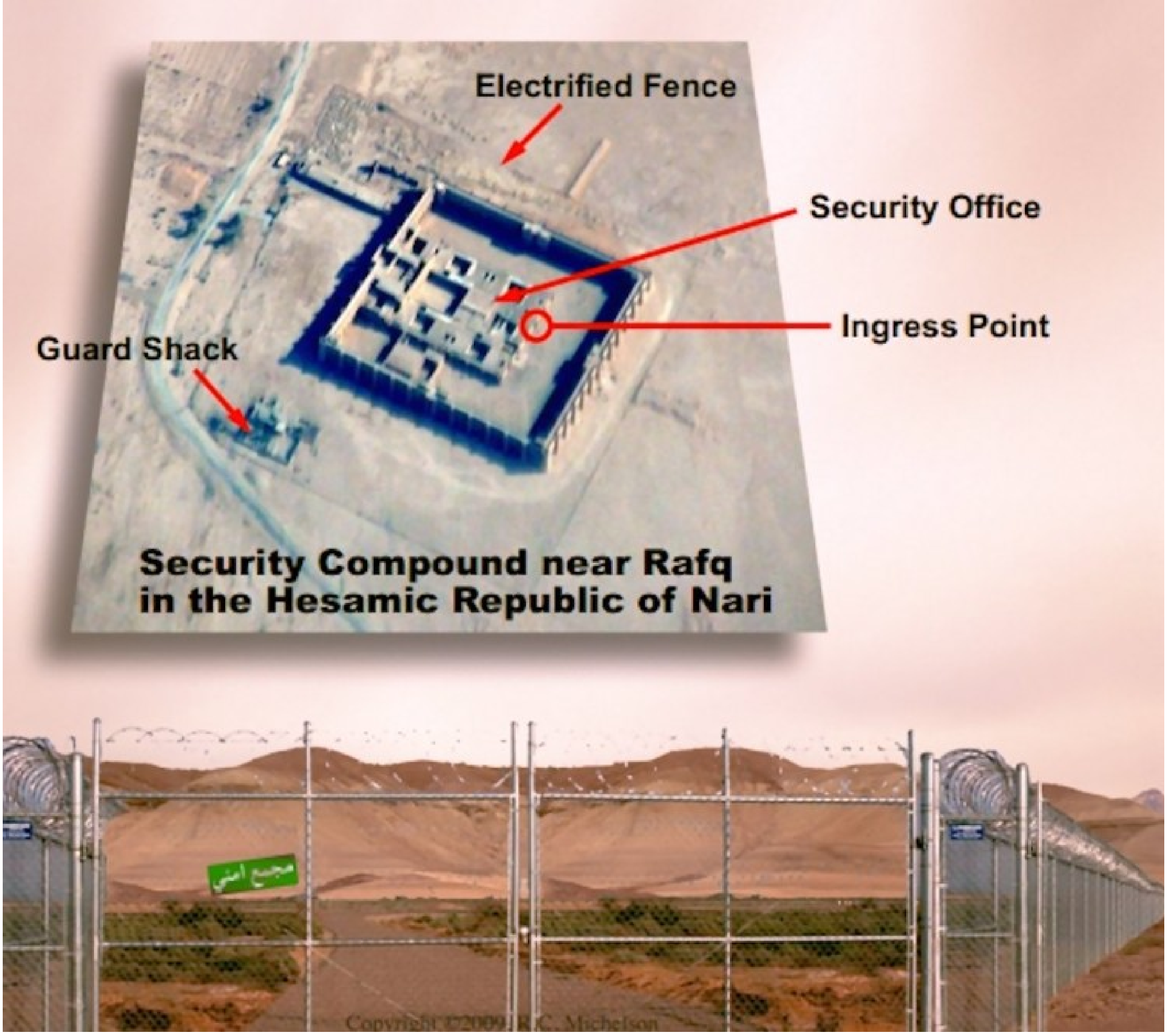
Electrified Fence
Security Office
Ingress Point
Guard Shack
Security Compound near Rafq
in the Hesamic Republic of Nari
مجمع امني
Copyright ©2009 R.C. Michelson

Illustration 2: Artist depiction of the IARC 2013 mission which included avoiding several security measures

The video starts with the drone taking off from the ground and entering the building through the broken window – although seemingly at a slow pace the quadcopter reaches and enters the building at 0:16 without any problems, eliciting the first amazed cheers from the audience.

The drone then continues to fly down a corridor, moving forward at a steady pace and slightly controlling for height and scanning its surroundings for any obstacles.

At the 1:00 mark of the video the quadcopter can be seen crossing the first room and taking a corner into the second room of the mission building – thus far, there have been no incidents or collisions despite the drone navigating a close quarter environment. Additionally, the drone is carrying a USB-Key similar to the one left in the building for a later exchange. In the background, nervous and excited chatter can be heard, each success of the robot is commented by cheers and clapping of the crowd. At 1:45 the crowd can be heard cheering as the drone passes into the second room around a corner and through a particularly narrow doorway. From a technical perspective it is quite astonishing to watch a quadcopter fly at such a steady pace and height in a a small room, something that remote controlled drones in the past have struggled with – clearly not relying on a human pilot allows for much more precise movement. At the 2:50 mark the robot can be seen approaching and positioning itself over the mission goal, a normal USB-Key placed on a small table. As the robot positions itself over the mission goal, the audience cheers in surprise and astonishment.

Dropping the fake USB-Key and picking up the real one proves to be a challenge for the drone as it hovers over the table, pacing its way carefully towards the mission object. Since the drone is not allowed to land on or touch the table, this part of the mission is the most sensitive one as the drone could easily be detected at this point. After about 2 minutes the quadcopter finally latches on the the USB-Key and starts its way back out of the office, the audience cheering and clapping in amazement. On its way back out of the office building the drones path finding seems to have improved it takes less time to navigate the corners and corridors – only the final window seems to be an issue as the drone hovers and scans it for a period of of time until it finally exits and hovers over the starting area where a member of the construction team is able to take the stolen Usb-Key with sensitive data on it which could jeopardize the international financial markets according to the scenario given out by the competitions organizers.

The robot demonstrated during this test flight shows some very desirable characteristics for a military and security-service application of the technology. First of all it demonstrates the advances made in sensor technology and the accompanying software algorithms needed to interpret such sensor information in a way that is useful for the robot. Second the robot exhibits fully autonomous behaviour once set into motion, completely forgoing the need for a human operator. Third, the size of the robot is approaching a range in which it becomes interesting for stealth-use and surveillance missions for instance as a border control agent, security drone at restricted facilities or as an aggressive assault drone. Additionally, the robot build by the Tsinghua students was able to detect and avoid security measures which would have presented great difficulties for any human trying to enter the building unnoticed – last but not least it should be clear that this was only a prototype with much potential for optimisation.

These 3 different imaginations of autonomous machines provide a different base for analysing the issue of a potential regulatory model for the use of autonomous machines in the future. Only by determining exactly the way in which we imagine a potential future to happen can we imagine a possible model of regulation and control. A model based on a unrealistic imagination of full machine autonomy could lead to overegulation and irrational policies based on imagined risks and fears. Governments could struggle with public opinion demanding stricter control due to fear based arguments. On the other side of the spectrum a model based on a very liberal interpretation of machine autonomy could lead to regulation which is too lax and does not take into account fears held by the population or actual technological risks. Only a careful consideration of the different possible imaginations of the central issue, "how" are the machines autonomous, can provide us with a realistic imagination of the technology and its involved risks.

As the previous section has shown the imagination of autonomy most present in society differ quite a bit in some core aspects. While NGOs and other actors of civil society emphasize the potential risks of full autonomy and the military use of drones stating the US drone programme and relying on imaginations of autonomy which are not based on current technological development but rather on popular fears and assumptions other actors such as manufacturers of such technologies have an active interest in portraying machine autonomy as something that is being developed cautiously and with the human actor in mind. Stating that a machine will still be monitored or controlled by a human actor is a central argument for the further development of autonomous machines brought forth by most manufacturers and supporters of liberal policies.

4) Regulation of machine autonomy

This section argues that in order to safely develop machine autonomy and minimize the risks of accidents and misuse a regulatory model for the development and use of autonomous machines should be developed and ratified by as many countries as possible. Similar to agreements on the non-proliferation of nuclear, chemical or biological weapons the military use of drones and other autonomous machines can be regulated, monitored and sanctioned if necessary. This section will argue why a model of regulation is important and how this potential model can be seen in a similar tradition as the regulation of weapons of mass destruction in the 20th century. Finally some of the most important factors in regulation this technology will be discussed and their potential impact on a regulatory model will be assessed.

4.1 General thoughts on the regulation of weapons and technology

Over the course of the 20th century humanity has developed the most deadly and efficient weapon systems ever devised. Never before has it been possible to kill, injure or otherwise incapacitate so many individual human being and to destroy such large region as with the beginning of the 20th century, when the first ever truly industrialised conflict took place and the first ever weapons of mass destruction where developed and used.

4.1.1 Chemical weapons and delivery systems

World War I was dominated by vast and ongoing battles between fixed lines of demarcation, battles would often only decide who took the next hill or the next trench a few more meters ahead. Warfare had become a matter of economics and logistics, whoever was able to build more machine guns, artillery and field more men would likely win a battle. Due to the proliferation of weapon technologies some enemy armies would even use the same weapon systems from the same manufacturers. World War I was later dubbed a war or attrition in which the main goal was to hold your own lines and weaken the enemies defence to a point where it could not be sustained any further. This massing of troops and materials called for new military tactics to gain the upper hand of a battlefield – tactics that would allow to damage the enemy without exposing ones own weaknesses – and so the first true weapon of mas destruction was developed: gas.

“The sulphur mustards, or sulphur mustards,[2] commonly known as mustard gas, are a class of related cytotoxic and vesicant chemical warfare agents with the ability to form large blisters on the exposed skin and in the lungs. Pure sulphur mustards are colorless, viscous liquids at room temperature. (...) Mustard gas was originally assigned the name LOST, after the scientists Wilhelm Lommel and Wilhelm Steinkopf, who developed a method for the large-scale production of mustard gas for the Imperial German Army in 1916.[3] Mustard agents are regulated under the 1993 Chemical Weapons Convention (CWC). Three classes of chemicals are monitored under this Convention, with sulphur and nitrogen mustard grouped in Schedule 1, as substances with no use other than in chemical warfare. Mustard agents could be deployed on the battlefield by means of artillery shells, aerial bombs, rockets, or by spraying from warplanes." (Mustard Gas. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Sulfur_mustard)

Chemical weapon system where the first weapons of mass destruction and their use quickly became a matter of debate. It became ever more apparent that war would become more and more brutal and affect ever larger parts of the civilian population unless decisive steps to regulate the use of certain weapon types and tactics would be taken.

4.1.2 Nuclear weapons and delivery systems

The end of the second world war brought the Manhattan Project, a concentrated effort by the US to develop and build the first nuclear bomb – the US was not the only power developing these weapons but the Sowjets would take a few years longer. With the use of the first two nuclear bombs in Hiroshima and Nagasaki the world witnesses for the first time a destructive power that was able to devastate whole cities – not just for a period of time, but the long-time effects would continue to damage the health of the people living in that region for decades to come. Nuclear weapons where a whole new level of escalation of military technology, for the first time in human history it was technically feasible to literally destroy the whole planet and all life on it. Nuclear weapons became the defining technology of the cold war, allowing for an uneasy peace and a balance of power which, as some claim, has kept the two major powers of the time the US and the Sowjets to invade each other. As with chemical weapons before, nuclear weapons where seen with even more scepticism due to their huge potential for disaster. When the US stationed nuclear missiles in Europe, demonstrations and political actions would manifest and demand the removal of such weapons from their territories. At the same time, the civilian use of nuclear power was and is heavily influenced by the military use of nuclear power as all civilian nuclear programmes are subject to monitoring by the international community.

4.1.3 Biological weapons and delivery systems

The third major class of regulated weapons are biological combat agents. Although examples of biological warfare can be found throughout history, it was the advances in bacteriology and biology at the turn of the 19th/20th century that have led to the development of sophisticated biological agents used to kill or otherwise harm enemy troops, populations but also as area of effect weapons which could be used to control larger areas – an example of the use of a biological of chemical agents for this purpose would be the United States use of Agent Orange during the Vietnam war to defoliate large areas of jungle ion order to gain better vision over the battlefield and to be better able to spot enemies hiding on the ground. The agent was later tied to a large number of side effects including generations of children affected by the agent in their development causing large number of birth defects and other disabilities.

"Biological warfare (BW)—also known as germ warfare—is the use of biological toxins or infectious agents such as bacteria, viruses, and fungi with intent to kill or incapacitate humans, animals or plants as an act of war. Biological weapons (often termed "bio-weapons", "biological threat agents", or "bio-agents") are living organisms or replicating entities (viruses, which are not universally considered "alive") that reproduce or replicate within their host victims. Entomological (insect) warfare is also considered a type of biological weapon. Biological weapons may be employed in various ways to gain a strategic or tactical advantage over an adversary, either by threats or by actual deployments. Like some of the chemical weapons, biological weapons may also be useful as area denial weapons. These agents may be lethal or non-lethal, and may be targeted

against a single individual, a group of people, or even an entire population. They may be developed, acquired, stockpiled or deployed by nation states or by non-national groups. In the latter case, or if a nation-state uses it clandestinely, it may also be considered bioterrorism. There is an overlap between BW and chemical warfare, as the use of toxins produced by living organisms is considered under the provisions of both the Biological Weapons Convention and the Chemical Weapons Convention. Toxins and Psychochemical weapons are often referred to as midspectrum agents. Unlike bioweapons, these midspectrum agents do not reproduce in their host and are typically characterized by shorter incubation periods" (Biological warfare. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Biological_warfare)

In addition to these broad categories of regulated weapons other types of weapons developed during the 20th century have been found to be too dangerous and/or cruel to be used in warfare and the need to regulate them has arisen. Examples include anti-personnel mines, certain types of ammunition such as so-called "Dumdum" expanding bullets which are made to cause large wounds and internal injuries or cluster-bomb ammunition and some forms of incendiary type bombs. Although not all major powers are in favour of regulating or forbidding the use of these weapons many nations do see the need to regulate them and have signed according treaties. Nations who have not signed the international ban-treaty on anti-personnel mines include Russia and the United States. (Ottawa treaty. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Ottawa_Treaty#Signatories).

4.2 Overview of known regulatory models for weapons of mass destruction

Regulation in most cases only took place after the weapon systems in question where used at least once against enemy troops or civilians. The expanding bullet was banned in 1899 at the Hague Convention (Wikipedia, "Expanding bullet")

Biological weapons were banned under the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on their Destruction

"(...) This includes all microbial and other biological agents or toxins and their means of delivery (with exceptions for medical and defensive purposes in small quantities). Subsequent Review Conferences have reaffirmed that the general purpose criterion encompasses all future scientific and technological developments relevant to the Convention. It is not the objects themselves (biological agents or toxins), but rather certain purposes for which they may be employed which are prohibited; similar to Art.II, 1 in the Chemical Weapons Convention *(CWC). Permitted purposes under the BWC are defined as prophylactic, protective and other peaceful purposes. The objects may not be retained in quantities that have no justification or which are inconsistent with the permitted purposes. " (Biological weapons convention. (n.d.). In Wikipedia. Retrieved November 06, 2014, from* http://en.wikipedia.org/wiki/Expanding_bullet#International_law)

"

Article I: Never under any circumstances to acquire or retain biological weapons.

- *Article II: To destroy or divert to peaceful purposes biological weapons and associated resources prior to joining.*
- *Article III: Not to transfer, or in any way assist, encourage or induce anyone else to acquire or retain biological weapons.*
- *Article IV: To take any national measures necessary to implement the provisions of the BWC domestically.*
- *Article V: To consult bilaterally and multilaterally to solve any problems with the implementation of the BWC.*
- *Article VI: To request the UN Security Council to investigate alleged breaches of the BWC and to comply with its subsequent decisions.*
- *Article VII: To assist States which have been exposed to a danger as a result of a violation of the BWC.*
- *Article X: To do all of the above in a way that encourages the peaceful uses of biological science and technology. "* (Biological weapons convention. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Expanding_bullet#International_law)

Similar to biological weapons another broad category of weapons of mass destruction developed in the 20th century is subject to heavy regulation and intervention by the international community.

Although under regulation, these kinds of weapons have been used in the past in armed conflicts for instance during the Gulf War or allegedly by the Assad troops during the Syrian war in 2013.

"The main obligation under the convention is the prohibition of use and production of chemical weapons, as well as the destruction of all chemical weapons. The destruction activities are verified by the OPCW. As of January 2013, around 78% of the (declared) stockpile of chemical weapons has thus been destroyed. The convention also has provisions for systematic evaluation of chemical and military plants, as well as for investigations of allegations of use and production of chemical weapons based on intelligence of other state parties.

As of September 2013, 189 states are party to the CWC, and another two countries (Israel and Myanmar) have signed but not ratified the convention. In 14 September 2013, Syria deposited its instrument of accession to the CWC and agreed to provisional application pending its entry into force effective 14 October 2013" (Chemical weapons convention. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Chemical_Weapons_Convention)

Perhaps the most well known technology for weapons of mass destruction are nuclear weapons. Biological and chemical weapons have revolutionized military tactics and escalated the dangers or warfare for the civilian population but these technologies have never reached the level of sophistication and efficiency that a nuclear bomb can provide to a military. Nuclear weapons were developed at the end of the second world war as a means to dominate and punish the enemy – the use of such weapons not only brings great destruction and suffering to an area as large as a city, it also lingers and has decade long after-effects on the population. Arguable there has never been another weapon develops in human history which is both so feared and so prevalent at the same time. Nuclear proliferation and non-proliferation has been the focus point of the cold war, the cold war itself being held stable by the uneasy coexistence of two major nuclear powers capable of annihilating each other at the press of a button. The nuclear arms race has led to an enormous use or natural and economical resources, created vast stockpiles of hazardous materials and waste and created a situation where even today at any point in time nuclear weapons could destroy the planet entirely many times over – a phenomenon referred to as "overkill" (Overkill. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Overkill_%28term%29#Nuclear_weapons). There are numerous treaties dealing with nuclear armaments some of which regional and some of which global, but the most important one to date is the "Treaty on the

Non-Proliferation of Nuclear Weapons”

"Opened for signature in 1968, the Treaty entered into force in 1970. On 11 May 1995, the Treaty was extended indefinitely. A total of 190 parties have joined the Treaty, with five states being recognized as nuclear-weapon states: the United States, Russia, the United Kingdom, France, and China (also the five permanent members of the United Nations Security Council). More countries have ratified the NPT than any other arms limitation and disarmament agreement, a testament to the Treaty's significance. Four non-parties to the treaty are known or believed to possess nuclear weapons: India, Pakistan and North Korea have openly tested and declared that they possess nuclear weapons, while Israel has had a policy of opacity regarding its own nuclear weapons program. North Korea acceded to the treaty in 1985, but never came into compliance, and announced its withdrawal in 2003" (Nuclear non-proliferation treaty. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Nuclear_Non-Proliferation_Treaty).

What all of these weapon technologies have in common is that each of them represents a new approach, a new step in using technology to increase ones own capabilities of warfare and exerting violence. IN World War 1, a conflict dominated by fierce trench warfare, battles that would last days and include much man-to-man combat in the trenches, a weapon capable of "flushing out" troops from the trenches was the most efficient means of fighting enemies in trenches, emplacements and bunkers. Since the use or airplanes was just being developed using gas to cover large areas of the battlefield would prove beneficial on a tactical level even though gas was hard to control and could easily injure or kill ones own troops if the wind changed suddenly.

Biological weapons were a little more targeted in the sense that they could be developed for certain purposes such as defoliating trees as was the case with agent orange. Biological warfare and bioweapons are also easier to manufacture as are chemical weapons are less ingredients are needed – this of course means that such weapons can in theory be produced by anyone with the knowledge of materials to do so, as has been the case with attempts of bio-terrorism.

Chemical agents are harder to produce and control safely so the main concern with chemical agents is the use by governments and other comparable groups in terms of resourcefulness. Used mainly as weapons of mass destruction or area of effect weapons chemical agents are controlled heavily around the world with many countries monitoring and or banning their production and many ingredients for producing such agents on a watch list.

Nuclear weapons represent an even larger investment a stakeholder has to make in order to develop, acquire and maintain a nuclear arsenal – not mentioning the strict rules of non-proliferation and the nuclear test ban treaties - nuclear weapons are certainly the most fearsome and destructive weapons known to mankind today but with rigorous regulation we as a global community have decided to limit the development of this technology to mostly civilian purposes.

But how do drones, how do semi and fully autonomous machines compare to these weapons of mass destruction and can they even be considered to similar to these WMD?

4.3 Comparing classic WMDs to machine autonomy

Drones and other autonomous weapon platforms can be developed by any stakeholder capable of developing computers and machinery – no banned materials are needed, no controlled or monitored substances are required to develop, build or maintain a fleet of drones. This means that virtually anyone with enough financial and material resources is theoretically capable of developing and using such technologies. Not only that, but investment would be relatively small compared to other weapon classes especially if one factors in the cost-benefit analysis. Many parts of a drone can be ordered ready made, software, AI and other components can also be outsourced to the lowest bidder and in many cases components, software and other technological aspects are interchangeable and reusable for different models or purposes. IN short, building a drone has become such a mundane task for a skilled engineering department that it is now common in student competitions and the hobbyist sphere all around the world.

Using a non-proliferation approach to drone technologies or autonomous machines would be rather futile because most components are commonplace and can be used for a range of different purposes unrelated to constructing a drone, components such as software are not even material goods and can be distributed in the blink of an eye over the internet and due to the popularity of hobby and DIY use of drone technologies. Regulating the spread and use of drone technologies and autonomous actors will be much more complex than simply banning a few key components or substances and/or monitoring their use and proliferation – especially with other disruptive technologies such as 3d-printing give anyone advanced production capabilities.

Being relatively easy to develop, produce and maintain drones offer another important advantage which makes them harder to regulate: Using a drone does not require much training, if any. Most consumer level commercial drones are today fitted with advanced flight controls, stabilizers and software that allows the pilot to focus on the actual navigation rather than struggling to maintain flying levels and so on – this technology has become to common and so cheap to produce that it is available to anyone over the internet in the form of small flight computers that cost the fraction of anything needed to control a sophisticated weapons platform such as any ABC weapon type. This also means that training people how to fly a drone or how to operate any other vehicle is basically unnecessary as they literally start to fly and drive themselves. Nobody needs to maintain training centres for pilots or scientists for maintain a fleet of drones or other autonomous or semi-autonomous machines after an initial training – they will work as long as they are maintained in working order, a task which any capable engineer should be fully qualified to fulfil.

An important aspect in the analysis of new technologies used for military purposes is the question whether the technology offers new tactical capabilities and opens up new strategies to whoever employs the technology. This was certainly true for biological weapons and chemical weapons in the first half of the 21st century – both types of weapon allowed for large area of effect attacks covering whole regions of land in some cases. Biological as well as chemical agents can also be used to attack enemies which have bunkered down or are hiding from plain sight. While this was also possible with artillery and other forms of explosives, an attack using any form of bio or chemical agent was more likely to hit and injure or kill the target – an artillery shell could land meters from the target and miss it while the newly developed biological and chemical weapons were capable of spreading out and creeping into bunkers, trenches and almost any other form of fortification common at that time.

These new weapon types also involved another important aspect for military strategists: With weapons of mass destruction a reality war and conflict became even more of a way to deter an enemy or influence a weaker opponent – the fear of the effects of mustard gas was enough to scare most people and offered the added benefit of damaging the enemies morale and determination during a battle. With the development of nuclear weapons in the late 1940s and the proliferation of weapon platforms capable of delivering these types of bombs to virtually any place on earth within a few days or hours such as the ICBM (Intercontinental Ballistic Missile) or long range bombers such as the B117 a new tactical aspect arose: The real chance of a world-wide nuclear war which would have devastating consequences not only to the victim of the attack but the aggressor at the same time – which led to a system of mutually assured destruction which dominated and dominates global politics and geostrategy for the rest of the century and beyond. Just like the technologies developed before it, drone technology and machine autonomy bring a whole new set of rules and tactical capabilities to the table which will most likely define the way armed conflicts are carried out in the coming decades. The purely military aspect of drone technology is just one factor however, as this technology goes far beyond what former emerging technologies such as ABC achieved. Drone technologies do not only have the potential to dominate armed conflict and warfare, they also offer unprecedented tactical and strategic capabilities to any domestic police force, security agency, secret service, private and corporate interests and more – in fact, the technology is so broad and transformative in some areas that it is hard to predict where this wont make an impact of some sort. Drones offer a degree of flexibility and efficiency unparalleled to any other military technology so far – they represent the combination of several technological innovations in the late 20th century such as the miniaturization of key components such as computer chips and power sources, advances in material science allowing for extremely light yet sturdy constructions of frames and other components, advances in artificial intelligence, machine learning and algorithms, advances in sensor technology for increases situational awareness or the availability of a global communication network as well as a global system for navigation. By combining all these technological advances in one technology, one assemblage of technological capabilities, social and ethical norms and political and economical requirements, drones have become more than just a type of weapon or weapon platform. Unlike ABC weapons drones and autonomous machines can also be used for a wide range of civilian and paramilitary purposes either in a occupied territory or domestically making a drone programme not only a very attractive solution for any military force but also for any government interested in keeping a tight grip on its own population or certain parts of a population. Of course, autonomous robots can also be used for tasks that would clearly benefit a society such as search and rescue operations, security tasks or simply as companion and caretakers or the elderly and disabled.

4.3.1 Differences in drone technologies and its application

The term “drone” is usually used to describe an unmanned aerial vehicle such as the Predator or Reaper drones the US forces are using in locations such as Pakistan or Yemen. A drone can however also be any other unmanned, remote controlled or semi/fully autonomous system be it mobile or stationary, land, sea or air based. A drone, according to this definition, is simply a vehicle that has a certain amount of autonomy and does not require any human pilots or operators to be present with the vehicle. Drones can have many different purposes and designs as well, not only for purely military purposes. Drones are used by by many governments today for a wide range of uses. For instance, drones can be used to police borders or other sensitive regions or territory such as power plants or factory areas, in this case acting as security guards (Microdrones GmbH. (n.d.). Applications for microdrones Aerial Platforms. Retrieved November 6, 2014, from http://www.microdrones.com/applications/applications.php). Such drones would differ from military drones for example by not carrying any weapons or only using less-than-lethal weapons, they could also be build much smaller since no payload is required. Other uses of drones include monitoring large urban centres such as cities, especially during mass events such as sporting events, demonstrations or unrest – the EU is specifically developing programmes for the purpose of mass control and detection of unwanted behaviour in urban areas. (INDECT. (n.d.). FAQ. Retrieved November 6, 2014, from http://www.indect-project.eu/faq#Q1.2)· Other uses of drones include emergency relief, search and rescue missions in rural areas (woods, mountain regions, sea) and recreational uses such as toys or hobby. A drone, apparently, can be seen and understood in many different ways according to what the analyst is looking for. One forum for the development and use of drones and other robots for search and rescue operations is the Eurathlon contest. Generally, robots developed for search and rescue operations are non or semi autonomous platforms with a large degree of human-user interaction and remote control. These robots are mainly specialised for a single environment such as air, water or urban environments and buildings and mostly have specialised tasks such as detecting victims after an earthquake, fighting fires, gas leaks , and other hazardous elements or in theory could even act as automated ambulances if build into a proper vehicle. The Eurathon competition held each year in Germany is a place where search and rescue robots get developed and tested:

“EURATHLON is a new robot competition supported by the European Commission in the FP7.

The vision of EURATHLON is to provide real-world robotics challenges that will test the intelligence and autonomy of outdoor/off-road robots in demanding mock disaster-response scenarios.

Inspired by the Fukushima disaster we envision a competition that requires autonomous flying, land and underwater robots acting together to survey the disaster, collect environmental data, and identify critical hazards. In 2013 the competition will focus on land based vehicles. While in 2014 the competition addresses sea based vehicles, in 2015 it will be held as a joined land, sea and air event." (Eurathlon 2013. (n.d.). Retrieved November 6, 2014, from http://www.eurathlon2013.eu/eurathlon_2013.html)

Robots for search and rescue purposes are not entirely new as already mentioned and were primarily based on direct remote control is the past. With advances in miniaturization, communication technologies, software and algorithms the direction of development is now being set to creating autonomous search and rescue robots capable of cooperation both with human and non-human actors. For instance, the 2014 Eurathlon will be held under the premise of creating swarms of robots capable of cooperation.

Autonomy, however, is also being seen critically by some researchers. In an interview given to german language tech-news portal Heise.de Michael Gustmann of the Nuclear Helpservice (Kerntechnischer Hilfsdienst) voices his opinion on autonomy in robots and what it could mean for search and rescue missions and beyond. From the interview:

"*heise online: "During the opening ceremony an emphasis on the development of cognitive and of autonomous robots was given. This could take stress of the operator, but also increase the risk of malfunction. What is your opinion on this?"*

Gustmann: "The one aspect Ill agree to right away. Supporting the operator is necessary. ON the aspect of autonomy however I see some big risks because we are facing very unforeseeable situations which do not allow for any planing. The operator has to receive as much information about the environment of the robot through its sensors or any additional systems and has to be supported while carrying out his task. We are strictly using the Man-IN-the-Loop principle, the operator monitors the machine at all times. We do not let our machines act alone and autonomously" *heise online: "All in all you are sceptical of autonomy?"*

Gustmann: "Yes, because its also a question of accountability and responsibility. Who is accountable in the end for a certain movement of the robot? This is where we like to see a person and in especially difficult situations even a gremium of people for decision-making and not an autonomous decision by the vehicle" (Marsiske, September 04 2013)

In another interview given by Shinji Kawatsuma of the Japanese atomic energy agency (JAEA) given during the Eurathlon competition the JAEA chief speaks about Fukushima and the role robots played and still play in the recovery operations. Fukushima is an excellent example of how ever increasing complexity in our technological systems, in this case energy infrastructure, are leading to situations and risks humans are not capable of dealing with adequately. In this case having human actors do the tasks of the rescue and survey robots would have been impossible or too dangerous due to the high radiation – compared to Chernobyl in 1984 where the soviets rushed in thousands of unsuspecting labourers to clean up the first aftermath of the catastrophe. Back then the death toll amongst workers and soldiers was high as can be expected when sending humans into a nuclear disaster zone mostly without any adequate protection. Two decades later, Fukishima shows the kind of progress we have made in technologies such as robotics, software engineering, artificial intelligence, sensor technology, material science, mobile energy supplies and much more – it is all of those advances that make a search and rescue robot possible.

Despite this however the first reaction in Fukishima was not to use robots for the recovery operations. As Kawatsuma points out, there were no robots available from the beginning that would have fitted the requirements of the operation – prior generations of robots did not have this kind of operations in mind and thus were not designed with any real life experience in search and rescue operations. There was also no feedback from human recovery workers which could have been implemented into the design process. Additionally missing rescue plans made it harder for the developers of the robots to plan and anticipate the working environment that such robots would face. Although Kawatsuma points out that a decision to create a specialised agency tasked with developing rescue robots was made in 1999, no real progress had been made so far due to lack of support and interest. Kawatsuma also adds that although operators have lost 4 robots in Fukishima, these issues where due to communication breakdown, one faulty propulsion system but never due to the high radiation.

Despite relying on machines for their recovery operations to a large degree, robots are seen as part of a larger system in Fukishima. The main lesson Kawatsuma points to is the need for a specialised organisation tasked with defining needs and requirements to such types of robots and to define concrete rescue plans. Before Fukushima, there was no such organisation and therefore no one felt responsible for creating contingency plans or defining design guidelines for search and rescue robots. The second lesson Kawatsuma draws from Fukushima is that while a single robot can be very useful it is the complex system behind such an operation that will provide the real use. A robot can only be as useful as the supporting infrastructure around it. The integration into such a complex system has to take place in times of peace and training for operators is of the utmost importance. Thirdly, Kawatsuma points out, real life tests are required to further improve the design and functionality of the robots but most importantly improve the whole system of rules, plans, design principles and training to create a robust infrastructure for such robots to function efficiently

On the issue of autonomy Kawatsuma has a clear message to developers and designers: Designers and developers have to remind themselves that in the end the operator of the robot has to be responsible for its actions. Therefore the operators should be included into the development at an early stage. Developers should not just develop single robots but robot-systems which has to include operator training. Robots should also be customizable for different missions, this customization should be easy and affordable to do by the rescue workers. According to Kawatsuma it is impossible to create hundreds of different robot types for all sorts of situations therefore robots have to be able to adapt and be adapted for a wide range of purposes. (Marsiske, September 25 2013). The development of new military technology can be seen as a combination of what the technological capabilities allow, what was requested by “the market” (the market being regular armies and government agencies) and what innovation weapon designers and manufacturers could come up with. Up to the mid 20th century, the development of military technology was limited to a few fields defined as “military” such as materials science, rocketry or space technology. With the development of computers, the microchip and ensuing miniaturization and popularization of digital technologies innovation in military technology research also increased. Today the private sector has developed into a huge industry capable of developing innovative products on their own, without having to rely on prior funding or requests. This shift away from a demand based economy to a supply based economy of weapons marketing has also led to a large increase of influence the private sector seeks over governments and their agencies. Lobbying and advertising new innovations or products to their customers has become the new norm, just like with other markets.

The private sector has a vested interest in being able to develop increasingly powerful and feature-rich unmanned platforms ("drones") not only because its what these companies do, but also because they now follow the same basic market principles as any other market – innovate or vanish, hold your market share or be dominated by the competition. Additionally, these companies are also subject to the rules of the financial markets and their shareholders.

It can thus be argued that due to these ground rules, the development of drone technology in the private sector not only follows the classic lines of military technology (receiving concrete orders or requests from governments) but also "normal" market rules where the seller has to advertise, market and lobby for his product in order to make a profit.

A good example of the kind of products available to the private individual by private industry are the Linux based Autopilot-Computers by Airware.

These mini-computers range from 32grams to over 200grams of weight and can be attached to most types of commercially available drones (fixed- wing drones, helicopters and multicopters) and provide a fully autonomous autopilot with advanced navigational and other capabilities. Using these computers any drone can be turned into a fully autonomous machine capable of fulfilling a wide range of tasks including surveillance, exploration, tracking of and observation of single targets and much more – since the flight computers are based on the open source Linux Operating System the computers flight algorithms can be altered and enhanced by anyone capable of programming it, adding the potential for a huge range of applications developed by a hobbyist community.

Combining these devices with a quadcopter capable of carrying a certain payload anyone can easily build a personal surveillance drone capable of autonomous flight and recording any object or individual in the open, a level of technology which was previously only available to law enforcement and military agencies or some specialised professions. (Brown, E., 2013)

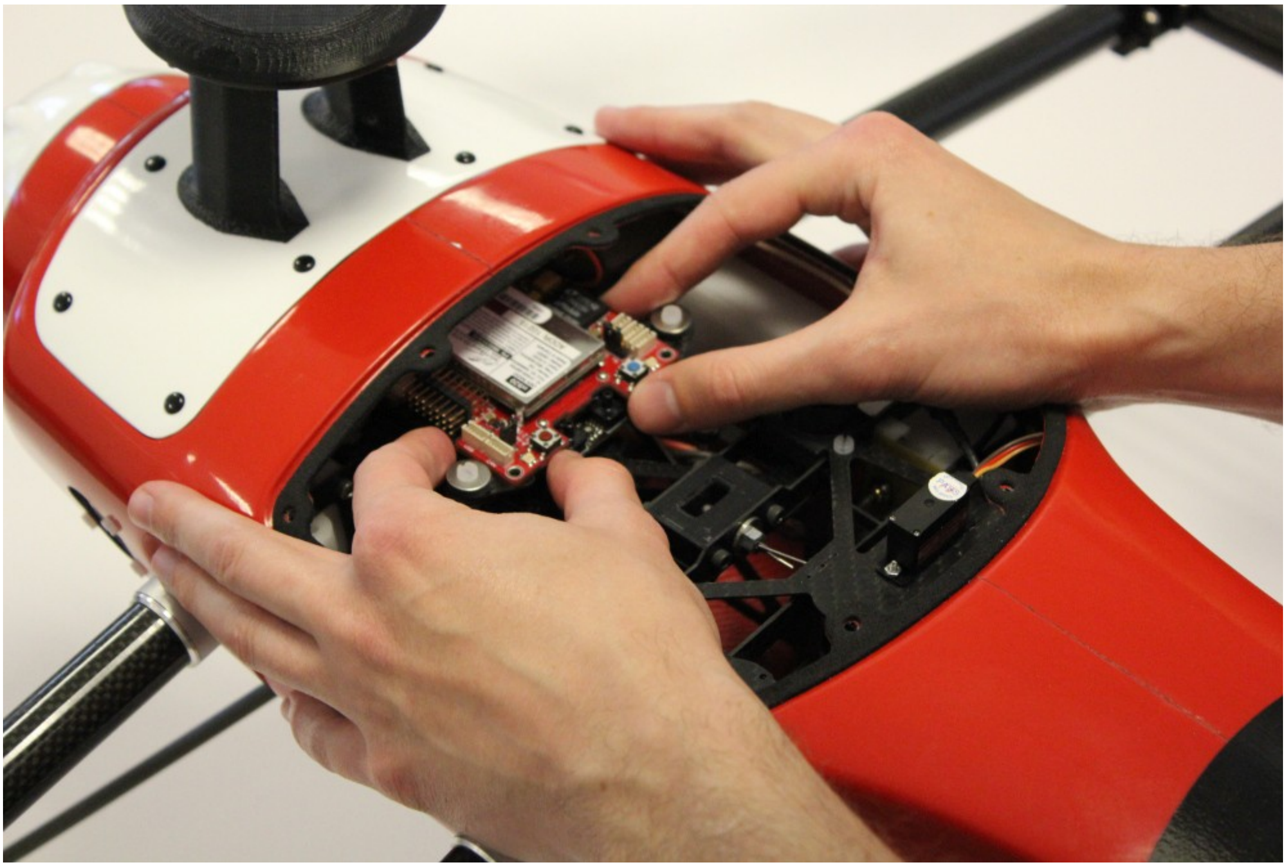

Illustration 3: Not much larger than a matchbox, the flight computer can be installed by anyone on almost any kind of commercially available drone to provide full autonomy

5) Analysing the machine autonomy assemblage

In the last section of the thesis the three distinct levels of autonomy that a remote controlled weapon system can achieve have been shown and analysed ranging from simple remote controlled flying to semi-autonomous navigation, detection and targeting of targets to the fully autonomous system able to navigate, detect, target and engage targets according to its programming or artificial intelligence and build in algorithms. The more the system gains autonomy, the more the human actor becomes obsolete – not only that, but the system becomes increasingly complex and thus harder to understand and anticipate as well as introducing a certain level of black boxing into its internal processes. These changes have a large impact on a number or moral, ethical, legal and economical issues which in return play a role in analysing and dissecting the drone assemblage.
An assemblage can be described as a bridging object between different disciplines, actors, approaches etc. Commonly used in geology, palaeontology, archaeology and art, recently it regains popularity in different fields (political sciences - Manuel DeLanda, science studies – Bruno Latour, cultural studies - Brian Massumi).

"In an assemblage, nothing explains it all: not the sciences, not the social sciences, not the human sciences. There isn't anything that is first or fundamental in an assemblage—nature, language, culture, institutions, whatever—it's all at once, and we with our questions come after it. Meaning that we are both assembled by it, and in pursuit of it. Even though we're consigned to come after the assemblage has been assembled, both with and without our intentionality, that doesn't stop us from going after it, too." (Fortun and Bernstein, 1998)

In the context of this thesis the assemblage approach allows us to look at the different actors involved in thinking about, using and making the intelligent machines. By including many different actors in our analysis we can gain a more complete understanding of social and cultural particularities, fears and risks (both true and imagined) as well as political and economical considerations and ambitions. In understanding the motivations behind the development of different kinds of autonomous machines we can also better understand and anticipate what needs to be regulated. Indeed the question of regulation itself can be argued by analysing the involved actors and realizing the complexity of the issue as well as the possible future development of this emerging technology.

5.1 Transparency & Regulation

Transparency is a very broad term used in politics, economics and the social sciences on a broad level. The term has gained popularity with the rise of the internet and new media as the driving factors in an increased demand for access to information and decision making. In the political context term transparency was popularized by such actors as Wikileaks and other whistle-blowers that have made an impact in recent years but also by new start-up political parties such as the international movement of Pirate Parties, a political platform dedicated to bringing more accountability and transparency into the political process and to establish a culture of open access and open data. Generally speaking, the growing demand for transparency is driving its inflationary use in political rhetoric but at the same time at least partially the concept is taking shape in actual legislation and rule-making on a larger societal level. On the other hand, the demand for more access to information and can also be easily silenced by an over-abundance of available data especially if no context or means of interpretation are given.

In the case of this thesis however transparency takes another role, since we are not talking about general political agenda but very specific technologies and their use in very sensitive contexts such as national security, defence and intelligence gathering. Transparency and access to information has very immediate effects on things like regulation, the assessment of the technological risks associated with using drone technologies, its representation in the media and top the public and on further research and development. For regulators, it is important to be able to know and understand as much precise technical information on a system that needs regulation as is possible – in the case of cutting edge military technology, that usually tends to be not that much or not enough to be able to fully anticipate and regulate certain risks of a technology. Any agency tasked with coming up with some sort of regulatory framework for the use of drone technologies would also have to have access to the tactics and strategies used with theses machines, another thing no military or security organisation is willing to give up easily. While for example the use of nuclear weapons is regulated and policed globally – the technology is well known and its principles and particularities well known – drone technology is not simply a single purpose technology but can be used and modified in numerous ways, making regulation even more complicated but necessary at the same time. Regulation of drone technology would have to deal with such questions as who is allowed to develop such technology (or if there should be restrictions on that at all), who can deploy drones and where, what purposes are legitimate for the use or drones – for example one might find a regulation that would allow for the use of drones as guards/sentries, but not allow for autonomous weapons fire requiring a veto agent (a human overseeing the drones actions). Regulation would also need to find binding rules on where drones can be deployed: With ongoing miniaturization it is to be expected that drone technologies able to enter private properties undetected will become readily available to law enforcement and security agencies – how will the use of such potentially invasive technology be regulated? Another important question would be the domestic use of drone technology against the own population, for instance in the case of social unrest or riots. Drones are very much suited for crowd-control tasks, for example flying over large crowds of protesters releasing a mist of some sort of crowd-control agent such as capsaicin or simply tasking close-up images of individual protesters for later use by law enforcement. The use of drone technologies for crowd control would dramatically cut down on personnel requirements for law enforcement and similarly to the use of drones instead of human soldiers on a war, the use of machines instead of humans against the own population would also incur less negative reactions as long as the use could be kept unknown or at least not challenged by the mass media.

In a scenario where the use of drone technology is widespread, another aspect of transparency and accountability comes into play: If drones are used in similar ways as human actors as guards, patrol units, police aids or weapons then being able to identify such a machine would be an important aspect in its acceptance in the public sphere. Where a human actor working as a guard can be asked who his or her employer is and what tasks he or she has, a sentry drone or similar machine would have to have some sort of identification method in order to determine its manufacturer, owner or whoever else is deemed accountable for the machines actions. If such drones should be used in public view, designing them to correspond to a certain image and making sure people can readily identify and understand their purpose would probably help in minimizing public backlash and perceived risk. For example, humans tend to react positively to human-shaped objects and attribute human characteristics to them this also happens when they resemble the proportions of children – a phenomena known as anthropomorphism (Anthromorphism. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Anthropomorphism) and “Kindchenschema” (scheme of childlike proportions). Robots designed in Japan for the care of elderly people tend to be designed to closely resemble human shape and that of children in order to reduce anxiety and of the elderly patients and facilitate acceptance of such non-human actors to replace traditionally human tasks (nursing or monitoring a sick person). (Valery, 2013)

A machine used to guard or patrol an area could easily be designed in a way that would minimize its potential to receive any damage, make it more or less invisible to the human eye and let it operate almost with no noise. Such a robot would probably not resemble any human shape and would be received with much scepticism and fear, especially if it was to act autonomously and might even carry a form or armament. Technically such a robot would probably be superior to a machine which would have to rely on human features such as legs or arms to function, but as mentioned before in order to avoid some of the public backlash the design of such machines should probably take into account the fears and assumed risks of drone technologies and create a more "appealing" package by adherent to some of the mentioned design principles. This is not only to avoid backlash or have a more positive risk-assessment but more importantly to allow for better regulation – if the use of such robots is widespread and more or less accepted, calls for regulation will be more frequent and the chance to build a robust regulatory framework increases – on the other hand, machines which are forced on a population, scare it or otherwise elicit negative reactions would probably be more subject to secrecy and thus make it harder to properly regulate their use and ownership. Design principles can also differ from culture to culture, a good example being the "robot cultures" of the US versus the Japanese robot industry. While the US model is centred on the military and security use of robots and drones, Japanese robots tend to be manufactured for industrial application or as a replacement for service personnel. Another issue having to do with the design of autonomous machines for the use amongst humans is the question whether people should be able to identify the purpose of such a machine in any way, and if yes, how. Our behaviour in the public sphere is heavily influenced by the way we perceive our surroundings and the people that are around us. People have developed uniforms, logos and other signs or belonging or purpose in order to be able to gauge and anticipate how the different human actors around us will act and influence our own daily life. For instance, we all know that the postal worker in his or her work uniform will bring the mail, we understand that a paramedic is someone with medical training capable of helping and rescuing injured people or that a fire-fighter can be recognized by his or her distinctive uniform. A autonomous robot, on the other side, does not need any of this. It does not need any clothes to cover itself, it does not need any external cues such as a uniform or logo to recognize its "peers" or "colleagues" and it does not need any protective wear – it can be constructed in a way to withstand anything it is supposed to withstand such as high temperatures or water. This leads us to the problem that in theory, one could easily design any robot with any purpose and make it look completely inconspicuous, harmless and benign. A robot can carry concealed weapon systems, advanced sensors and surveillance equipment and means for fast navigation and still look like a “black box” - in this case, quite literally because the purpose of a machine can be hidden behind simple metal plating. A robot deployed in a urban environment as a guard could turn out to be more than just a guard, that automated car on the street could just be a robot-taxi or it could be a mobile surveillance station housing a large array of advanced sensors and cameras to scan a whole neighbourhood for unwanted activity. The point here is, while we are able to see and understand to a certain extend what a human actor in any given situation is supposed to do and is capable of doing, we cannot judge with an anonymous machine designed to be inconspicuous and look benign. While this might seem unimportant for the military application of drones and robots, the ever growing domestic use of such systems would probably benefit from a clear standard for identification – just like certain vehicle types are marked and certain professions carry uniforms such machines could also be painted or otherwise marked with their purpose, functions and anything else that might be vital to dealing with such a machine. Only if people will be able to judge and understand a drones purpose in their daily lifes will the technology be accepted on a larger scale and will not be seen as invasive and alien.

5.2 Accountability

Legally and on questions of accountability, this poses a big issue: Who is responsible for malfunctions and accidents? What happens when civilians suffer, who can be held accountable before court? As mentioned before, any algorithm written by human actors would be written according to some basic ethical/moral considerations – such “ground rules” would have to be agreed upon, possibly in a international agreement as proposed by the Harvard Law School Human Rights Clinic and Human Rights Watch in their report

For the general population, accepting the use of fully autonomous systems might depend on the “how and where” - the domestic use of such systems would no doubt be seen more critical then the use on the battlefield in a far away region of the world. But even then, public reactions to the use and potential issues would play a major role on open societies, it would play a smaller role in more secretive societies where the use of such systems might be kept secret to the population at large as was the case with the US drone programme until recently. (Anders, 2013)

5.3 Machine learning

The analysis of an emerging technology can sometimes include aspects of it which are still not understood very well or which still pose a large potential for development and thus are harder to assess and work with. One such aspect in the debate about autonomy in machines is the question if and how such robots are capable of so called machine learning, the process of developing and expanding its own programming to dynamically react to a changing environment or mission.

One of the reason machine learning is still not well understood is its complexity and its dependence on rapidly improving computational capabilities and ongoing miniaturization of hardware. While the development of sensor technology, navigational systems or materials used took some decades the implementation of first autonomy and later machine learning supported autonomy capable of dynamically reacting to events only took place in the last decade and is still rapidly developing.

Safety quickly becomes a concern with machine learning as is introduces a factor of uncertainty into a technological field which was so far associated with a sense of determinism and predictability – a robot could malfunction, but only in a few predictable ways for which risk could be managed and minimized. With the introduction of machine learning into the behavioural routines of a robot however, the robot becomes a whole other class of non-human actor – it gains a level of agency which was unknown to non-human actors before.

In "Emerging Technoethics of Human Interaction with Communication, Bionic and Robotic Systems" the authors analyse the long term and short term ethical issues involved in machine learning in robots. They mainly focus on the development of civilian use robots such as robots for the assistance of the elderly of disabled or robots as companions for hobby and professionally.

An important point the authors make right in the beginning of the analysis is that in the past, robots and similar machines where developed under the assumption of a very static environment where they would have to function, this being especially true for industrial robots. Risks to humans in interaction with these machines were minimized by defining clear security rules for the behaviour towards such robots by humans and also towards the capabilities and functions of the robot. The authors refer to this practise as "segregation" and make the point that such a policy is likely to fail with more complex robots capable of acting in a dynamic environment with machine learning enabled algorithms giving the machine the capability to interpret any situation on a dynamic scale rather than with a fixed reaction. In the past the use of robots in more dynamic environments populated by humans was regulated by developing these robots as specialized machines with single purpose in mind, often minimizing the interaction or contact the machine would have with humans in order to minimize risk of accidents. Robots designed to interact with humans however will need to be very different however, being able to interact with humans within a acceptable range or risk of malfunctions (ETHICBOTS, 2008). If the development of non-military robots should follow the same trajectory it is taking at the moment, then autonomous robots will soon be acting as guards, search and rescue agents, hospital robots for daily tasks which require little human oversight, care for the elderly or infants. Obviously the old rule of trying to avid as much contact with humans as possible to minimize the risk of accidents or malfunctions and to avoid having to create complex security mechanisms. For this new mode of human-machine interaction which is no longer a separation of labour but a acceptance of robots as non-human actors capable of fulfilling complex tasks new rules and regulations will have to be defined to govern the way these new kinds of robots should interact with us – or should not interact with us, depending on the application. For example, do we want a robot to be able to administer drugs to a patient if it thinks the patient needs them, or should it just alert a human doctor and wait for instructions? What if the patient is in pain or in immediate need of assistance, what level of response will we allow a robot to take and what would happen if the decision ends up harming or even killing the patient? This conflict would already be true for a robot simply capable of fulfilling certain tasks, but a machine learning enabled robot would not just simply fulfil its tasks it would learn new information about its patients, environment and possibly even new drugs or instruments to use. A robot could decide that a new drug is better suited for a patient, even without a doctors consent if the programming allows for it, and administer it to a patient.

Expanding this onto the use of autonomous drones, UAVs and robots for military and associated security purposes, machine learning would introduce a new level of risk and uncertainty into the planning of military actions. A robot capable of dynamically reacting to a changing situation would either have to have very precise and unchangeable mission parameters or a very flexible set of rules of conduct in order to function effectively. So either such a machine would have to be build and programmed with very strict rules on the amount of flexibility in its behavioural routines and algorithms or these algorithms and routines need to be so sophisticated as to allow to complex ethical and moral considerations, maybe even ad-hoc during a mission. For example, if the rules on engagement list a cone shaped object as a potential weapon and threat as they do now, a drone capable of engaging a target or otherwise harming it would have to be able to either understand that not every cone shaped object is a weapon and thus consider alternatives or it would need to be able to train its sensors to better and better detect false positives. Such routines require much more complex programming and carry and increased risk of malfunctions and bugs, thereby increasing the risk.-factor of the technology. In their EU Deliverable the authors make the point that while many successful machine learning applications have already been deployed learning still plays a limited role in current robots – it has to be noted that this document is from 2008 and due to the very rapid development of this technology significant advances have to be taken into account. Generally speaking they describe a learning machine as a robot which has been programmed with a set or rules and assumptions about its working environment and is capable of applying these and additionally re-interpreting them in accordance with data input from the actual mission or task. Another method is a predetermined (by an external human engineer for example) set of alternative actions and responses to anticipated variations in the robots tasks and demands on it. (ETHICBOTS, 2008)

"A learning robot acting on the basis of background conjectural assumptions or biases about a partly unknown environment may try and get additional information by deploying learning algorithms that either change its "control policy" on the basis of "on-line" responses from the environment or enable one to identify inductive hypotheses on the basis of "off-line"

training data provided by some instructor. Surprisingly enough, however, one finds that learning plays a limited role in current robotic systems, as far as the adaptation of overall behavioural responses of a real robot during task execution is concerned. Even though a variety of both supervised and unsupervised learning approaches are being pursued, and a plethora of successful applications have been reported, none of these approaches and applications is easily adjusted for the purpose of achieving "autonomous learning" in robots." (ETHICBOTS, 2008)

In both cases the mutual co-construction of rules based on cultural, social, ethical, political and economical assumptions and positions plays an important rule. A robot, even a learning one, has to be programmed with a set of rules, assumptions and predetermined reactions to a standard set of potential issues. For instance Asimovs Laws on Robotics (Three laws of robotics. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Three_Laws_of_Robotics) state that no robot should be able to harm any human, no matter what. So a robot physically capable of causing harm would be constructed or programmed in a way which would make it impossible for it to harm a human, either by limiting its capabilities or by installing some sort of fail-safe mechanisms, security overrides or other forms of emergency control. But this is only true if the actors involved in producing such a technology are all determined to uphold these laws. Predetermined rules of conduct carry a heavy load or prescribed (human) values and requirements and play a leading role in the development of new robot (drone, UAV) technologies. Then there are pre-programmed reactions to potential new situations which require a certain amount behavioural flexibility on the machines part but which are still predetermined by a human actor. This, however, is not machine learning in the true sense as the machines reactions can still be anticipated to a large degree and accordingly risk of such technologies can be managed more easily. Machine learning, at this stage, is still a heavily human assisted process in which a human actor still has to help the machine determine meaningful relationships between objects and data basically acting as a trainer.

Machine learning can roughly be divided into the three areas of supervised computational learning unsupervised realtime-learning and reinforced learning. While the first process requires a "trainer" which can either be a human actor or another specialised algorithm and cannot be done "on the fly" in real time as it requires samples of the input and output values to be provided by a "trainer". These input and output values need to be pointed out to the machine, which means that in the case of a human trainer ethical/cultural/social/political/economical bias plays a role when giving instructions or pre-determining certain reactions and actions. Unsupervised, or real time learning can be achieved by teaching the machine to detect irregularities in its environment, however this still requires semantic analysis by another actor in most cases human. Reinforced learning does not require any supervision same as real-time learning but it can be used on a broader scale and more flexible to learn a wide variety of tasks on a trial and error basis using simple reward signal as the "motivator". In this case, reinforced learning is the closest to what most people understand when they hear "autonomous learning". It is of course not fully autonomous but requires a large amount of preparation and optimization which has to be done by human actors before learning can even take place. Since RL is based on a rather crude trial and error process this also means that any machine utilizing this technique will need to undergo extensive testing and "training" to fully develop its potential and minimize the risk of potential malfunctions, bugs or other unforeseen risk-factors. One also need to take into account the relatively large cost of time and computational power required to this process to take place especially when the machine has to interact with its environment which could be changing dynamically as well.

One major difficulty when trying to asses the risk of a machine-learning based actor is the fact that such a machine does not know if or what mistake it has potentially made. If a machine is supposed to recognize certain people or types of objects, an operator working with the machine will not know what that machine is "thinking" or "seeing" - for that, interfaces such as GUI (graphical user interface) or certain statistical data can be used but this only allows for some interpretation of the data the machine is acting on. So fixing or improving a machine can be potentially a very long and involved process since it can be hard to actually identify the reason for a machines misbehaviour. When dealing with a human actor such as a guard or other security personnel one can always simply ask for the reasons the human actor was performing a certain action or how he or she assess a situation, with a machine even a autonomous, machine learning enabled one, communication and interaction will always be more prone to mistakes as it requires additional steps of translation and interpretation on both sides of the communication.

5.4 Black boxing

How to deal with emerging technologies and regulate them is one of the major and central concerns of any government and also an important part of STS research, at least in its more modern and applied variant. In the case of drones and other autonomous machines regulation will most likely have to include solutions for things like accountability, transparency, safety or the protection of civil and human rights from being interfered by this technology.

One major challenge for this is the tendency for “black boxing” large parts of drone technologies to the public and concerned scientists. “Black boxing” refers to the process of hiding complex systems and procedures behind simple looking solutions and can be expanded to political or social solutions as well. For example, in drone technology not only the research and development of such technologies is largely seen as a trade secret or security risk and therefore not accessible to anyone not involved in the process. This makes the process of how new drone technologies are researched, developed and tested into a “black box” since we can only see the input (what requirements and prescriptions are used) and the finished output, not however the methodology and process used to achieve the goal. This is also true when dealing with the autonomy aspect of the drone technology as is also involves certain input and requirements and an output in the form of a piece of code, controlling microchip or even just a subroutine of a larger algorithm. The process how this software is being constructed is largely hidden to us up to the question if human or non-human actors are involved in training and further developing algorithms.

So the two main processes of how drones and other associated technologies are developed and produced, the hardware aspect as well as the software aspect are mostly black boxed complex systems which of course makes a technology and risk assessment very challenging to say the least. Black boxing not only affects the understanding of how drones, UAVs and other forms or autonomous robots are made, it also involves to large degree the ongoing research on the features and capabilities of these systems. Since many of the developed technologies are applied by military and or security forces, secrecy and information management play an important role in the risk communication on the issue. For example, media reports seldom target modern semi or fully autonomous systems but rather well known and established technologies such as the Predator and Reaper drones which are remote controlled. Not knowing or having the complete range or features and capabilities or a technology however makes it very difficult for researchers and regulators to assess and deal with the social and political implications. An example here would be the degree of autonomy and machine learning capabilities, understandably a very sensitive issue for manufacturers and users of the technology but at the same time of great concern to a larger public. A good example of what can happen when a new technology is largely black boxed is the decision of the german minister of defence to invest a substantial sum into the development and acquisition of drones capable of surveillance missions. During the phase of acquisition it was either ignored or not made clear to other involved stakeholders by the defence ministry that these drones would not feature a system for collision control and therefore would not be licensed to fly in EU airspace. (Schröder, 2013) Additionally, it was not clear that theses drones also include sophisticated equipment for intercepting communication which led to the situation that no rules for privacy or the protection of civil rights were even considered. (Bockenheimer, 2013)

While some of the issues in this case are likely due to political bargaining and strategies of avoiding blame and accountability, the underlying problem of not sharing or not making available important information of the complex system of the Eurohawk drone has led to a situation where the public is now very much opposed to the acquisition and use of these drones while the minister has lost a major part of his credibility.

5.5 Private / Civilian use of drone technologies

Drone technologies are increasingly being used by private individuals for many different purposes. Many hobbyist like to fly their drones and quad copters like model planes while others build sophisticated custom solutions to any issues they might try to tackle – this includes custom solution for taking aerial pictures and video for instance for journalistic purposes, the development of commercial services using drones technologies such as private health care (search and rescue), deliveries and logistics (using drones and other autonomous machines to deliver wares), private security and surveillance and much more. This is not only an issue of safety in terms of having private drones flying around in a region, it also poses unforeseen risks for businesses, government agencies and other actors as drone technologies can be used in very disruptive ways even if not intended. Questions of air traffic safety come into mind a ell as privacy issues or questions on how to deal with the potential social conflicts the widespread use of drones could bring.

Regulating the private use of drone technologies will probably be comparable to the regulation of other "dual use" technologies, meaning that the technology can be used privately as well as commercially or in conflict. This could include approaches from the regulation of the use of radio frequencies but also from gun control. (Brown, B., 2013)

5.6 Government agencies

A remote conflict offers many tactical advantages for the attacker, as opposed to a ground-conflict with real troops. Not only do drones offer 24-hour strike capabilities, they can reach any spot on the planet within a few hours, target individuals and specific targets albeit not with the kind or surgical precision that the US likes to claim. For a democratic society that needs to legitimize its actions to a parliament or other from of representative of civilian society, drone warfare offers the opportunity to minimize war-weariness in the population since losses are not comparable to a conventional conflict. The US drone programme is not headed by the US military, therefore it does not have the same congressional oversight as other operations might receive – instead it is headed by the Central Intelligence Agency, which has more freedom and autonomy in conducting missions which are not subject to public scrutiny. (Mulrine, 2013)

In the case of the United States, several actors are lobbying for more control over the drone programme behind the scenes, although the administration might present a united approach to the public, tensions behind closed doors are running high. Among the actors fighting for control over the drone programme are the Central Intelligence Agency (CIA) and the US department of defence. All the while it is still unclear in what way drones should be applied in the future, opinions ranging from the wide spread use not only by the US itself but also its allies to prevent or counteract militants in controlled regions, others such as the president's counter-terrorism adviser, John O. Brennan have called for more restraint and view drone strikes as last resort measures.

The US is setting a legal precedent by using drones to target and kill their perceived enemies wherever they are located globally. Since 9/11, two administrations have taken the position that the US are at war with Al-Qaeda and therefore are legally allowed to engage them using pre-emptive strikes.

As mentioned before, these strikes are presented as surgical, accurate with minimum civilian casualties. There are "a targeted, focused effort at people who are on a list of active terrorists." according the President Obama. As shown in the Stanford Law Schools report "Living under drones", attacks now regularly include so called "signature strikes" which target certain military "signatures" such as groups of people, vehicles, certain demographics such as young men, certain objects which could be weapons (cone-shaped objects for instance). What constitutes as a military signature that warrants a drone strike is a question of debate and in need or proper definition. It is these attacks which are perceived as random or indiscriminate by the local population which lead to the strongest backlashes towards the US locally but also globally. The killing of signature targets brings into question the policies and technologies which govern the target acquisition and decision to attack. A different case to look at is the usage of drone technology in the European Union. Most notably, the EU is working on a border protection programme under the title of "Perseus" which stands for "Protecting European Seas and Borders Through The Intelligent Use Of Surveillance". This so called "demonstration programme" is part of the European Unions 7th Framework of projects and programmes and is categorized under security/surveillance. The expressed goal of the Perseus programme is to develop and maintain a system of maritime border control – in other words, the EU is developing a fleet of surveillance drones to monitor its sea borders. According to the projects website the test version of the programme is budgeted at 43.7m Euros and limited to a four year test period.

"PERSEUS addresses the call for an integrated European system for maritime border control. Its purpose is to build and demonstrate an EU maritime surveillance system integrating existing

national and communitarian installations and enhancing them with innovative technologies. By means of two large scale demonstrations PERSEUS will prove its feasibility and will set the standards and grounds for the future development of EU maritime surveillance systems. (,,,) The new maritime surveillance system is expected to increase the effectiveness of the current systems by creating a common maritime information sharing environment for the benefit of the network including National Coordination Centres, Frontex and the European Maritime Safety Agency (EMSA). The project also envisages collaboration with non European countries and international agencies such as NATO or the International Maritime Organisation (IMO), among others. This system-of-systems will use all the information provided by the European and national agencies. The data will be integrated and processed for better quality, thus obtaining filtered, reliable and more useful information. In particular, PERSEUS is meant to support the implementation of EUROSUR" (European external border surveillance system (EUROSUR). (2008, March 14). Retrieved November 6, 2014, from http://europa.eu/legislation_summaries/justice_freedom_security/free_movement_of_persons_asylum_immigration/l14579_en.htm)

So far, two campaigns are planned for the Perseus system: the first campaign will take place in 2013 within the Western Mediterranean Basin from the Atlantic approach to Italy and North Western Africa, the second campaign is scheduled for 2014 and will be conducted in the area of the Aegean Sea , potentially including an extension up to the Black Sea.

The Perseus programme is part of the larger EU EUROSUR project set up by the EU to create a state of the art border protection system using innovative and efficient technologies. Set in 3 phases, the project aims to 1) Interconnect and rationalise border surveillance systems at national level 2) Improve the performance of surveillance tools at EU level and 3) Creation of a common monitoring and information-sharing environment for the EU maritime domain.

According to the official project website:

"(...) a European border surveillance system (EUROSUR), focusing initially on the Union's southern and eastern maritime borders, could be developed, and proposes a roadmap for setting up such a "system of systems" over the next few years. It focuses on enhancing border surveillance in

order to:

- *reduce the number of illegal immigrants who enter the European Union undetected;*
- *reduce the number of deaths of illegal immigrants by saving more lives at sea;*
- *increase the internal security of the EU as a whole by contributing to the prevention of cross-border crime.*

A European border surveillance system (EUROSUR) should help the Member States achieve full awareness of the situation at their external borders and enhance the reaction capability of their law enforcement services. "Situational awareness" measures the capability of the authorities to detect cross-border movements and find reasoned grounds for control measures; "reaction capability" measures the lapse of time required to control any cross-border movement and the time and means necessary to react adequately to unusual circumstances. EUROSUR would provide the common technical framework required to rationalise cooperation and 24-hour communication between the Member States' authorities and foster the use of cutting-edge technologies for border surveillance. One essential operational objective must be to create an information-sharing (excluding personal data) environment among national and European systems." (European external border surveillance system (EUROSUR). (2008, March 14). Retrieved November 6, 2014, from http://europa.eu/legislation_summaries/justice_freedom_security/free_movement_of_persons_asylum_immigration/l14579_en.htm)

Although the projects website seems to offer a wide range of information on the programme, more concrete details are harder to find. For instance, there is not much to be found on the actual technology being used in this project. There is however a list of project partners which includes all major players on the aerospace, air, space and arms industry such as Boeing, Airbus the EADS and a whole range of security and technology companies associated with the development of drones and other required components such as sensor technology, communication, navigation and training. (PERSEUS-FP7. (n.d.). Retrieved November 6, 2014, from http://www.perseus-fp7.eu/?page_id=17). Being part of the larger EUROSUR initiative which aims for a stronger protection of the EU borders Perseus would be an important part of the surveillance and sensor infrastructure in place all over Europe. A similar task of border protection in during the cold war would have required an enormous amount of manpower, resources and time – the drone network to be established by Perseus cuts this demand down to a minimum will still increasing the capabilities, coverage and effectiveness of the border agents.

The European Unions tendency to creating a "fortress Europe" against unwanted illegal immigration is subject to much debate and criticism. While EU officials and politicians argue that illegal immigration into the EU mainly from the African continent would unduly tax the European social systems, cause social problems and increase crime human rights activists, NGOs and some political parties and politicians argue that creating an ever expanding security and surveillance infrastructure to combat illegal immigration is not only ineffective, it would also lead to higher death tolls of migrants, increase incentives for organized human trafficking and set a dangerous precedent ion the EU to rely on mainly military and police based solutions. Even though most EU politicians seem to agree that creating a heavily policed border supported by a infrastructure of drones and other means of detection is a viable temporary fix to the issue they also agree that it is not a permanent solution and only improving living conditions in the originating countries would lead to a long lasting effect. One of the claims of the supporters of the Perseus/EUROSUR solution is that with a permanent observation using a drone network boats taking of at the African coast could be detected much earlier and possibly held back by the local authorities. The claim also goes that this would decrease cases of accidents and help with directing rescue boats to any migrant boats in trouble. Critics argue however that search and rescue operations as well as prevention of accidents and deaths on sea would not play a large enough role in the plans. They also voice concerns over some of the wide ranging plans for data exchange to non-EU partners such as NATO and local authorities on he African coast. The incident at the Italian island of Lampedusa (Lampedusa Migrant Shipwreck. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/2013_Lampedusa_migrant_shipwreck) has led to an increase in publicity and debate over the treatment of illegal migrants and the measures taken to combat illegal immigration into the EU. During the accident a boat filled with migrants sunk on the coast of Lampedusa – over 270 people were already found dead amongst them over 20 children while only 155 of the remaining people survived the sinking boat. Such accidents are common for the boats attempting to cross the sea between the African the the Eurasian continent, corpses are washed to the shores of both sides on a regular basis. An accident on this scale however does get the attention of the media which usually does not care much about reporting on the ongoing humanitarian crisis which is illegal immigration using boats unsuitable for sea-crossing. The migrants themselves often get victimised early on during their dangerous voyage across Africa to reach the northern coast, once they reach the coast only few can pay the traffickers and boat captains for the extremely dangerous voyage.

While the Europeans are mainly working on developing a fleet of drones and robots for border protection and surveillance, the US forces are interested in a wider variety of drones and robots to be used for different purposes. US officials are looking for robots which can be used domestically as well as abroad in an armed conflict – for that the robots need to be able to navigate a wide range or environments and have many flexible uses. During a robotic live-fire demonstration at Fort Benning, Ga., army officials got a demonstration of weaponized robots by four different private corporations.

“A weaponized robot acting as a member of a squad of U.S. soldiers fighting on the battlefield is no longer science fiction. They may not be two-legged, humanoid robots yet, but with wheels or tracks they are able to follow troops through a wide range of terrain and back them up in battle. Both the U.S. Army and the U.S. Marines have tested prototypes of weaponized robots on the battlefield. However, armed robots are not currently in the Army's inventory of weapons. Four robotics companies -- Northrop Grumman, HDT Robotics, iRobot Corp. and QinetiQ -- demonstrated their robots' abilities to fire machine guns and take out pop-up targets from a distance of 150 meters during the live-fire demonstration. Phil Coker, director of integrated platform systems at Northrop Grumman, said its robot, the CaMEL (for Carry-all Mechanized Equipment Landrover), can run for 24 hours on three-and-a-half gallons of fuel, and can be equipped with a grenade launcher, an automatic weapon and anti-tank missiles. The CaMEL also can identify targets from three-and-a-half kilometers away, using a daylight telescope or thermal imaging. The robot also can be dropped into a war zone from a helicopter or a plane. The robots can be controlled, tetherless, from a hand-held device that looks much like a gaming control, a laptop computer or a tablet *attached to a vest that a soldier wears. The vest, which weighs about 10 pounds, carries a battery, the hand-held controller and a tablet that flips down from the soldier's chest so he can see what the robot sees. That means the soldier doesn't have to see the live target himself. The robots also can be operated, via satellite radio communications, from hundreds of miles away.”* (Gaudin, 2013)

"It's actually a good thing," said Staff Sgt. Douglas Briggs, Maneuver Battle Lab NCO, stationed at Fort Benning. "It keeps soldiers out of harm's way." Briggs, who has worked with robotic-armed

machines equipped with machine guns in Iraq, said a big part of bringing robots to support active duty soldiers is trust. The soldiers need to trust that their robot will not only work *when needed, but will not harm them.*

"It comes back to old ways and incorporating new stuff," Briggs said. "We need to see if it's going to do what they say it will do. It's like when we started using GPS instead of a compass. I trusted my compass. I had to get used to GPS." While robots may eventually become trusted members of a squad, the military is far from willing to give a robot the autonomy to fire on its own. Tollie Strode Jr., a senior project officer with the Maneuver Battle Lab *at Fort Benning, said, at this time, there always will be a human in the loop when a decision is made to have a robot fire a lethal weapon.*

"The robot may acquire an enemy target, but it will still always ask a human for permission to fire," Strode said. "I think the ability for a robot to acquire and assess a target and ID it as a threat and fire is probably five or 10 years out. However, even if that capability exists... we'll have a human in the process of deciding what to do." (Gaudin, 2013)

5.7 Research institutions

Research and development of military and security technology at US universities and other research institutions is nothing novel, much of the digital technology we use today was initially developed for military purposes. Things like the internet, lasers, microchips or miniaturization of components were developed by or with the help of universities or "science" in general. In this context, university-industry cooperation and ties also plays an important role as they often present the direct tie between military application and research is achieved via proxy by a private corporation using or applying research and developing it into concrete projects, this is sometimes described as the “triple helix" of government-industry-research (Etzkowitza and Leydesdorff, 2000, p. 109-123).
A common term or description of the entanglement of government, industry and the military is the “military-industrial complex” coined by US President Dwight D. Eisenhower in his famous speech given to the American public at the end of his term. In it he warned society of a system that he perceived as a danger to a democratic society in which private corporate interests and the power-interests of governments of institution would form a pact of mutual benefit. Eisenhower was referring to the entanglement of the military with industrial production but also with research and funding, which was at a peak during the time due to the US-Soviet space race and the general arms

race. Eisenhower warned that such a system would eventually become self-sufficient and would integrate itself into broader society, making it more likely for more and more investment in arms and military research and development while applying capitalist market logic to it

"This conjunction of an immense military establishment and a large arms industry is new in the American experience. The total influence -- economic, political, even spiritual -- is felt in every city, every State house, every office of the Federal government. We recognize the imperative need for this development. Yet we must not fail to comprehend its grave implications. Our toil, resources and livelihood are all involved; so is the very structure of our society.

In the councils of government, we must guard against the acquisition of unwarranted influence, whether sought or unsought, by the military industrial complex. The potential for the disastrous rise of misplaced power exists and will persist.

We must never let the weight of this combination endanger our liberties or democratic processes. We should take nothing for granted. Only an alert and knowledgeable citizenry can compel the proper meshing of the huge industrial and military machinery of defence with our peaceful methods and goals, so that security and liberty may prosper together.

Akin to, and largely responsible for the sweeping changes in our industrial-military posture, has been the technological revolution during recent decades.

In this revolution, research has become central; it also becomes more formalized, complex, and costly. A steadily increasing share is conducted for, by, or at the direction of, the Federal government.

Today, the solitary inventor, tinkering in his shop, has been overshadowed by task forces of scientists in laboratories and testing fields. In the same fashion, the free university, historically the fountainhead of free ideas and scientific discovery, has experienced a revolution in the conduct of research. Partly because of the huge costs involved, a government contract becomes virtually a substitute for intellectual curiosity. For every old blackboard there are now hundreds of new electronic computers.

The prospect of domination of the nation's scholars by Federal employment, project allocations, and the power of money is ever present and is gravely to be regarded.

Yet, in holding scientific research and discovery in respect, as we should, we must also be alert to *the equal and opposite danger that public policy could itself become the captive of a scientific technological elite.*

It is the task of statesmanship to mold, to balance, and to integrate these and other forces, new and old, within the principles of our democratic system -- ever aiming toward the supreme goals of our free society. " (Eisenhower, 1961)

University Affiliated Research Centers

One concrete example of the entanglement of university research and the military in the US are the University Affiliated Research Centers (UARC). There are four Institutes located a four different Universities, each specializing in different research and development programmes specifically for military application. At the University of Texas, the Institute for Advanced Technology (IAT) is working in the fields of electrodynamics, pulsed power, and hyper-velocity physics, whose focus has been to develop the fundamental scientific basis for new classes of high velocity kinetic energy weapon systems. The IAT has developed unique capabilities to model and analyse the behaviour of materials under transient conditions. As a result of the Army UARC relationship, the IAT has established unique, dedicated, state-of-the-art experimental and test capabilities for the Army in these critical areas, and has established state-of-the-art vector and parallel classified computational facilities (University Affiliated Research Centers. (2012, November 15). Retrieved November 6, 2014, from http://www.arl.army.mil/www/default.cfm?page=510).

At the Institute for Collaborative Biotechnologies (ICB) located at University of California at Santa Barbara (UCSB) in collaboration with the California Institute of Technology (Caltech) and the Massachusetts Institute of Technology (MIT) and its industrial and Army partners address research in the areas of Biomolecular Sensors, Bio-Inspired Materials, Lightweight Portable Energy, and Flexible Energy-Dispersive Composites, Biodiscovery Tools, Bio-Inspired Network Science, and Cognitive Neuroscience. The third institute, the Institute for Creative Technologies (ICT) at the University of Southern California Focusing on research in to Counter Insurgency (COIN), sustainment operations, tactical intelligence, leadership, decision-making and a wide-range of therapeutic applications, the ICT seeks to redefine the range of skills that Warfighters can obtain from future, dynamic simulation systems. The ultimate goal of the combined research and prototype developmental efforts of the ICT and its partners is to harness the power of artificial intelligence, emerging visuals, immersive simulation technologies, and storytelling to provide America's Army a worldwide technological advantage on the battlefield against terrorism. The ICT conducts basic research, applied research and advanced technological prototype development focused on Virtual Humans, Social Simulations, Emerging Visualizations, Sounds, Graphics, Mixed Reality (MR); and Learning Sciences to advance Army training and analytical capabilities. The focus of the fourth UARC, the Institute for Soldier Nanotechnologies (ISN) located at the MIT focuses its research and development on Lightweight Multifunctional Nanostructured Materials and Hybrid Assemblies, Soldier Medicine: Prevention, Diagnostics and Far-Forward Care, Multiple Blast and Ballistic Threats: Materials Damage, Human Injury Mechanisms and Lightweight Protective Systems, Hazardous Substances Sensing, Recognition and Protection, and Nanosystems Integration for Protected Communications, Diagnostic Sensing and Operational Flexibility in Complex Environments. (University Affiliated Research Centers. (2012, November 15). Retrieved November 6, 2014, from http://www.arl.army.mil/www/default.cfm?page=510).

"A University Affiliated Research Center (UARC) is a strategic United States Department of Defence (DoD) research program that is associated with a university. UARCs were established to ensure that essential engineering and technology capabilities of particular importance to the DoD are maintained. University Affiliated Research Centers (UARCs) are designed to provide critical mass in research areas that meet Army and DoD future needs and anticipated combat requirements. UARCs are university-led collaborations between universities, industry and Army laboratories that conduct basic, applied and technology demonstration research. The universities, considered at the forefront of science and innovation in a specific research area, provide dedicated facilities and share space with Army and industrial participants. The industrial partners provide competence in related technologies, expertise in transitioning technologies from laboratories to markets and cost sharing. The emphasis for each UARC is to conduct research where breakthroughs are likely to enable revolutionary capabilities for our warfighters." (University Affiliated Research Centers. (2012, November 15). Retrieved November 6, 2014, from http://www.arl.army.mil/www/default.cfm?page=510)

6) Conclusion

This small analysis was to show how theoretical approaches to risk and technology assessment can be used to analyse and understand aspects of a new technology, in this case the question of machine autonomy/intelligence and how this could influence factors such as accountability, transparency and acceptance of a technology. In a full analysis, categories of actors (human and non-human) would have to be defined, the development of the algorithm that controls such a system would have to be analysed in regards to pre&inscribed moral and ethical values, modifiers such as machine learning would need to be considered and much more – in the end, an assemblage of technological features, mutually constructed by various actors, values and regulations which arise from a negotiation process in society (involving yet different actors) could be painted.

As any technology, intelligent machines can be used and abused. We now have the luxury of being able to think and theorize about technological progress and the effects technological revolutions can have on our societies and our systems in society. Not only this, but societal change and emerging technologies have become topics for popular debate and democratic decision making, something that was not true during the most part of the 20th century when some of the most dangerous and efficient weapon systems were developed by small power elites without the consultation or involvement of the larger public. When the Soviet Union and the United Stated decided to unleash the power of nuclear weapons there was no involvement of the public in the decision or larger public debate about the effects such devastating weapons could have on society and the way we wage war. A dialogue in society about technological change can be achieved through different means in society – education plays an important role as does the open access to relevant information. The media plays another central role in publishing and popularizing certain framing angles as has been shown in this thesis. NGOs, think tanks, lobby-organisations, manufacturers and researchers, governments and institutions all have their particular interests and viewpoints about an emerging technology such as machine autonomy and therefore subscribe to a certain agenda and framing of the issue.

One way of implementing a long lasting dialogue about the development of intelligent machines which can act in different degrees of autonomy would be the establishment of an international regulatory body similar to the International Atomic Energy Agency or the Organisation for the Prohibition of Chemical Weapons. Such a body, lets call it the Machine Intelligence Agency, could act as a neutral party in evaluating individual research programmes and developments but could also be part of creating international rules and regulations similar to those for ABC weapons in order to avoid certain dangers and risks of the technology. An internationally binding treaty on the use of intelligent machines would be a precedent in human history as it would be the first time when we as a global society could reach an agreement on how we would like to co-exist. Such a treaty could include rules on the degrees of autonomy (remote/semi/fully autonomous) permissible for certain applications. For instance drones for civilian purposes could generally be permitted with a "higher" level of intelligence/autonomous behaviour than military or weaponized drones capable of harming and killing humans. Regulation could also recommend the development of certain aspects of machine autonomy such as the use as an intelligent sensor network or swarm to monitor large scale climate events, ocean currents, animal movements, forests and agricultural areas and so on. Intelligent drones could become part of search and rescue operations or even become transportation and delivery devices for wares – imagine a delivery drone in an urban or rural environment replacing a bicycle courier or postman. An intelligent approach to regulation and a precise set of rules, together with proper monitoring and enforcement could mean that we as a species could finally learn from our mistakes in the past and avoid developing the next generation of weapons of mass destruction. However, time is running out as there are now already over 70 drone programmes in existence most of which focus on the military/security aspect of the technology (Rogers, 2012) . A respected international body consisting of all stakeholders of society could become and important moral advocate for a civilian use of drone technologies for which there is much demand and potential. Intelligent machines not only have the potential to revolutionize warfare, the peaceful use of the technology could also help us create some of the necessary systems to combat such issues as climate change, desertification of arable land or illegal use of natural resources such as illegal gold mining, logging or fishing.

This thesis would like to make the claim that just like the regulation of chemical, biological and nuclear weapons as well as some types of ammunition it is necessary for us as a global society to negotiate a regulatory model dealing with robots/machines, their autonomy and their use for armed conflict. While prior regulatory models for 20th century weapons of mass destruction put a focus on banning access to certain materials, technologies and machinery or limiting the application of certain technologies to only predefined purposes such as a civilian energy programme (Nuclear program of Iran. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Nuclear_program_of_Iran) the regulation of machine autonomy will require novel approaches. The current trend towards developing more and more sophisticated and capable machines is becoming clear when looking at corporate giant Google and its plans for the coming decades: After acquiring eight robotics companies, amongst them Boston Dynamics and Schaft. Googles ambition to dominate the emerging robotics marked shows the importance of this emerging technology for both strategic and commercial purposes. After acquiring and dominating large parts of the economy of the world wide web (Google search, Gmail, Youtube, Adwords, Google Sense etc.) the giants next step is to dominate the internet of things (Internet of Things. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Internet_of_Things), in which robotics and machine autonomy will play a central role. This is just a small example of how an emerging technology such as machine autonomy can be adopted in different ways and for all sorts of purposes, civilian or military. It is up to us as a society if we allow for unmonitored and uncontrolled expansion of machine autonomy with all its potential benefits and risks or if we choose to develop this technology on a conscious route with clear goals and limits in mind. When Google started its search services in the late 1990s, there was very little or no fear of one corporation at one time dominating the kind and amount of information we are able to find and work with online – today, Google controls large parts of the mainstream world wide web and plays an important role in society both as a company and as a cultural entity. A technology as potent as autonomous machines with its associated disciplines such as machine learning or AI involves so many variables, cultural/social/ethical/moral inscriptions, technological pre and descriptions as well as unforeseen factors such as the emergence of any disruptive technology (which then can alter the course of technological development in radical ways) requires our full attention and the best of moral and ethical standards we are willing to subscribe to.

Although it is far to early to construct a model of regulation for the case constructed in this thesis, there are some points which we can already assume or anticipate to play a role.

1) **Dual use:** Unlike with ABC weapons and ammunition types, autonomous machines present a very broad spectrum of uses. While its clear what a missile payload or bomb is supposed to achieve and therefore can be properly regulated or banned a machine is only as dangerous as what it is being used for and what other risk factors it might present. With little customization a machine such as a drone can be outfitted as a civilian as well as an armed/military/police actor. The dual use aspect of regulation in this case presents a large hurdle for regulators, who always have to keep in mind that their prescriptions of the

possible uses of any machine are only as accurate as their imagination allows and the description of such a technological object can lead to very new and unforeseen uses and variations in how a machine is used.

2) **Readily available:** Unlike ABC weapon programmes, autonomous machines do not require any specialised equipment to produce such as a nuclear programme would require. Drones and other robots can be manufactured by any high tech industry available to anyone with the required funds and knowledge which is relatively easy to acquire compared to ABC weapons. Maintaining a fleet of drones does not require scientists with degrees in nuclear physics, biology or chemistry but only relatively easy to train professions.

3) **Politically ambivalent:** The dual use aspect of machines, machine autonomy and other related applications presents a large issue to regulators. With ABC weapons is is easy to determine the purpose of an undertaking and claims of peaceful intents can be relatively easy to dismiss. Since any drone can be outfitted or "weaponized" and therefore "de-weaponized" within just a few hours of simple mechanical maintenance creating a fleet of machines capable of both being used for legitimate civilian purposes as well as for potentially illegitimate armed conflicts or armed missions against any domestic issues is feasible. Any party can therefore easily uphold an image of legitimate use while still retaining the potential for any other use of their existing machine technology.

4) **Wild cards and black boxes:** Autonomy presents a completely novel factor in regulation never before encountered in terms of widespread use and both potential for accident and abuse. Many experts on the field consider machine autonomy the most risky aspect of this technology and the one with the most potential for unforeseen consequences if not properly regulated. It is hard to predict the course technological development will take in the coming years but at the moment the "man-IN-the-loop" and the "man-ON-the-loop" approach is being seen as the most sensible approach to autonomy meaning that any action taken by the machine is still at least being monitored by a human actor at all times and can be interrupted at the press of a button. While there seems to be a certain consensus that full fledged machine autonomy would present too many risks the technological development is certainly progressing in that direction and technological feasibility is only a few more generations of

> machines away. A regulatory model on the use of machines and machine autonomy would obviously mainly have to deal with this issue as it is the defining issue of this technology. Since other factors of the technology such as access to appropriate manufacturing technologies or required base materials cannot be regulated due to their widespread nature and the training to develop and maintain a fleet of machines is also relatively easy to achieve one vector for any real regulation could be the permitted degree of autonomy for any individual system or any class of machines.

Throughout this thesis I have tried to show and argue how different imaginations of a technology and its pre and described values make up a final result, for instance a technological object like the drone or a technological concept like machine autonomy. According to the theoretical approaches presented in this thesis can all be answered differently according to what imagination of autonomy is used. This makes is hard for anyone trying to come up with a model of regulation, since it it unclear on what basis the regulation should take place – if it should focus on risk management or economic freedoms, if it should require a human actor in the loop or permit the development and use of fully autonomous systems. How these questions are answered and thus how they are translated into regulation depends on the understanding of machine autonomy by the involved actors, in addition to moral and ethical considerations towards the issue. By viewing and understanding machine autonomy not just as a technology but also as a concept of delegation and agency it becomes possible to understand why there are different imaginations and ideas about how machines and humans could or should interact. The drone as a real world example of what the technology of machine autonomy can be used for shows us that the development of a technological object always involves a range of actors and stakeholders who shape the development of this technological object – what is expected of it, how certain fears or risks are dealt with, how the technology is used and by whom but also what should not happen with this technology and how to avoid it. The drone as well as the wider idea of machine autonomy are closely linked tu questions of delegation of responsibility, agency and accountability away from human actors to non-human agents. In this sense both Jasanoff and Latour are relevant to the issue, Jasanoffs sociotechnical imaginaries provide a framework of analysing and understanding different viewpoints towards one issue while Latour actor-network approach provides a better understanding of what role agency and delegation plays in our societies and our technological systems and infrastructures.

7) Bibliography

Akrich, M. (1992). The De-Scription of Technical Objects. In W. E. Bijker & J. Law (Eds.), *Shaping Technology / Building Society: Studies in Sociotechnical Change* (pp. 205-224). Cambridge: MIT Press.

Anders, C. (2013, March 26). Obama's drone killing program slowly emerges from the secret state shadows. *The Guardian.* Retrieved from: http://www.guardian.co.uk/commentisfree/2013/mar/26/obama-drone-killing-program-secret-state?CMP=twt_gu (last access: 06.11.2014).

Attrition Warfare. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Attrition_warfare

Anthromorphism. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Anthropomorphism

Beck, U. (1992). *Risk society towards a new modernity.* London: Sage Publications.

Becker, J., & Shane, S. (2012, May 29). Secret 'Kill List' Proves a Test of Obama's Principles and Will. *The New York Times*. Retrieved November 6, 2014, from http://www.nytimes.com/2012/05/29/world/obamas-leadership-in-war-on-al-qaeda.html?pagewanted=1&_r=1

Bijker, W., & Law, J. (1992). *Shaping technology/building society: Studies in sociotechnical change*. Cambridge, Mass.: MIT Press.

Bijker, W. (2003). The Need for Public Intellectuals: A Space for STS: Pre-Presidential Address, Annual Meeting 2001, Cambridge, MA. *Science, Technology, & Human Values,* 443-450.

Biological warfare. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Biological_warfare

Biological weapons convention. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Expanding_bullet#International_law

Bockenheimer, J. (2013, August 25). Ministerium vertuscht Datenschutz-Patzer. *Handelsblatt*. Retrieved November 6, 2014, from http://www.handelsblatt.com/politik/deutschland/euro-hawk-drohne-ministerium-vertuscht-datenschutz-patzer/8690978.html

Broad, W. (2013, March 20). Weapons Experts Raise Doubts About Israel's Antimissile System. *The New York Times*. Retrieved November 6, 2014, from http://www.nytimes.com/2013/03/21/world/middleeast/israels-iron-dome-system-is-at-center-of-debate.html?pagewanted=all&_r=0

Brown, B. (2013, October 8). Drones! Learn how you too can be shut down by the federal government. *Networkworld*. Retrieved November 6, 2014, from http://www.networkworld.com/community/node/84050

Brown, E. (2013, August 16). Linux-based autopilots target commercial UAVs. Retrieved November 6, 2014, from http://linuxgizmos.com/linux-based-autopilots-target-commercial-uavs/

Broekens, J., Heerink, M., & Rosendal, H. (2009). Assistive social robots in elderly care: A review. *Gerontechnology, 8(2)*, 94-103.

Chaboud, A., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2014). Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *The Journal of Finance,* 2045-2084.

Cruise missile. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Cruise_missile

Chemical weapons convention. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Chemical_Weapons_Convention

DPA. (2013, October 10). European Parliament approves Eurosur border surveillance. *Deutsche Welle*. Retrieved November 6, 2014, from http://www.dw.de/european-parliament-approves-eurosur-border-surveillance/a-17149625

Delgado, A., Kjolberg, K., & Wickson, F. (2011). Public engagement coming of age: From theory to practice in STS encounters with nanotechnology. *Public Understanding of Science,* 826-845.

Deri, A. (2012). Costless" War: American and Pakistani Reactions to the U.S. Drone War. *Intersect The Stanford Journal of Science, Technology and Society, 5*. Retrieved November 6, 2014, from http://ojs.stanford.edu/ojs/index.php/intersect/article/view/367/167

Druckman, J. (2001). The Implications of Framing Effects for Citizen Competence. *Political Behavior, 23*(3), 225-256.

ETHICBOTS Emerging Technoethics of Human Interaction with Communication, Bionic and Robotic Systems. (2008). Emerging Technoethics of Human Interaction with Communication, Bionic and Robotic Systems (SAS 6 Nr. 017759). Naples: Ethicbots Consortium, c/o University “Federico II” of Naples

Elish, M. (2012, March 12). Popular (Mis)Conceptions & the Perpetual Rise of the Machines. Retrieved November 6, 2014, from http://blog.castac.org/2013/03/popular-misconceptions-the-perpetual-rise-of-the-machines/

European external border surveillance system (EUROSUR). (2008, March 14). Retrieved November 6, 2014, from http://europa.eu/legislation_summaries/justice_freedom_security/free_movement_of_persons_asylum_immigration/l14579_en.htm

Eurathlon 2013. (n.d.). Retrieved November 6, 2014, from http://www.eurathlon2013.eu/eurathlon_2013.html

Eisenhower, D. (1961, January 17). Farewell Radio and Television Address to the American People. Retrieved November 6, 2014, from http://www.presidency.ucsb.edu/ws/index.php?pid=12061&st=military industrial&st1=

Etzkowitz, H., & Leydesdorff, L. (2000). The Dynamics Of Innovation: From National Systems And "Mode 2" To A Triple Helix Of University–industry–government Relations. *Research Policy,* 109-123.

Expanding bullet. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Expanding_bullet#International_law

Fiorini, P., & Prassler, E. (2000). Cleaning and Household Robots: A Technology Survey. *Autonomous Robots, 9*(3), 227-235. Retrieved November 6, 2014, from http://link.springer.com/article/10.1023/A:1008954632763

Fortun, M., & Bernstein, H. (1998). *Muddling through: Pursuing science and truths in the 21st century.* Washington, D.C.: Counterpoint.

Gaudin, S. (2013, October 11). Machine gun-toting robots may soon back up U.S. soldiers. *Computerworld*. Retrieved November 6, 2014, from http://www.computerworld.com/s/article/9243164/Machine_gun_toting_robots_may_soon_back_up_U.S._soldiers?taxonomyId=128&pageNumber=1

Giddens, A. (1990). *The consequences of modernity.* Stanford, Calif.: Stanford University Press.

Gorman, S., Dreazen, Y., & Cole, A. (2009, December 12). Insurgents Hack U.S. Drones. *The Wall Street Journal*. Retrieved November 6, 2014, from http://online.wsj.com/news/articles/SB126102247889095011?mg=reno64-wsj&url=http://online.wsj.com/article/SB126102247889095011.html

Harvard Kennedy School. (n.d.). Frequently Asked Questions about Sociotechnical Imaginaries. Retrieved November 6, 2014, from http://sts.hks.harvard.edu/research/platforms/imaginaries/imaginaries-faqs/

Harvard Kennedy School. (n.d.). What is STS. Retrieved November 6, 2014, from http://sts.hks.harvard.edu/about/whatissts.html

Harvard Kennedy School. (n.d.). STS Methods. Retrieved November 6, 2014, from http://sts.hks.harvard.edu/research/platforms/imaginaries/ii.methods/methodological-pointers/

Human Rights Watch & International Human Rights Clinic. (2012). *Losing Humanity: The case against killer robots.* United States of America: Docherty, B.

International Red Cross. (2003). *Use of nuclear, biological or chemical weapons: Current international law and policy statements*. Geneva: n.d

International Aerial Robotics Competition. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/International_Aerial_Robotics_Competition

INDECT. (n.d.). FAQ. Retrieved November 6, 2014, from http://www.indect-project.eu/faq#Q1.2

Internet of Things. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Internet_of_Things

Jasanoff, S. (2004). *States of knowledge: The co-production of science and social order*. London: Routledge.

Kasperson, R., Renn, O., Slovic, P., Brown, H., Emel, J., Goble, R., ... Ratick, S. (1988). The Social Amplification Of Risk: A Conceptual Framework. *Risk Analysis, 8*(2), 177-187.

Kato, S., Tsugawa, S., Tokuda, K., Matsui, T., & Fujii, H. (2002). Vehicle control algorithms for cooperative driving with automated vehicles and intervehicle communications. *IEEE Transactions on Intelligent Transportation Systems,* 155-161. Retrieved November 7, 2014, from http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1033758&url=http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber

Knightkimi [Knightkimi]. (2013, August 4th). *IARC 2013 mission 6 completed - Tsinghua University*. Retrieved from https://www.youtube.com/watch?v=B-iZE_Nn52w&feature=player_embedded

Law, J. (2008). Actor-network theory and material semiotics. In: B. S. Turner (Ed.), *The New Blackwell Companion to Social Theory, 3rd Edition.* (pp. 141–158). Oxford: Blackwell

Lee, Y. (1996). 'Technology transfer' and the research university: A search for the boundaries of university-industry collaboration. *Research Policy, 25*(6), 843-863.

Lupton, D. (1999). *Risk (Key Areas)*. Routledge.

Library of Alexandria. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Library_of_Alexandria

Library of Congress Washington DC Congressional Research Service. (2010). *Homeland Security: Unmanned Aerial Vehicles and Border Surveillance* (ADA524297). Washington, DC: Library of Congress

Leiner, B., Cerf, V., Clark, D., Kahn, R., Kleinrock, L., Lynch, D., ... Stephen, W. (n.d.). Origins of the Internet. Retrieved November 6, 2014, from http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet#Origins

Lampedusa Migrant Shipwreck. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/2013_Lampedusa_migrant_shipwreck

Marra, W., & Mckneil, S. (2012). Understanding “the loop”: Regulating the next generation of war machines. *Harvard Journal of Law and Public Policy, 36*(3). Retrieved November 6, 2014, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2043131

Marsiske, H. (2013, September 4). Hoffentlich müssen wir die Roboter nie einsetzen. *Heise*. Retrieved November 6, 2014, from http://www.heise.de/newsticker/meldung/Eurathlon-Hoffentlich-muessen-wir-die-Roboter-nie-einsetzen-1966083.html

Marsiske, H. (2013, September 25). Der Robotereinsatz in Fukushima: Shinji Kawatsuma im Interview. *Heise*. Retrieved November 6, 2014, from http://www.heise.de/newsticker/meldung/Der-Robotereinsatz-in-Fukushima-Shinji-Kawatsuma-im-Interview-1966855.html

Mustard Gas. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Sulfur_mustard

Mulrine, A. (2013, March 20). Does it matter who runs US drone program? Pentagon could supplant CIA. *Christian Science Monitor.* Retrieved November 6, 2014, from http://www.csmonitor.com/USA/Military/2013/0320/Does-it-matter-who-runs-US-drone-program-Pentagon-could-supplant-CIA

Microdrones GmbH. (n.d.). Applications for microdrones Aerial Platforms. Retrieved November 6, 2014, from http://www.microdrones.com/applications/applications.php

National Science Fund. (2007). *Sociotechnical Imaginaries and Science and Technology Policy:A Cross-National Comparison* (NSF Award No. SES-0724133). United States OF America: National Science Fund.

Nuclear non-proliferation treaty. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Nuclear_Non-Proliferation_Treaty

Nuclear program of Iran. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Nuclear_program_of_Iran

Ottawa treaty. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Ottawa_Treaty#Signatories

Overkill. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Overkill_%28term%29#Nuclear_weapons

PERSEUS-FP7. (n.d.). Retrieved November 6, 2014, from http://www.perseus-fp7.eu/?page_id=17 PERSEUS Partners. (n.d.). Retrieved November 6, 2014, from http://www.perseus-fp7.eu/?page_id=325

Rogers, S. (2012, August 3). Drones by country: Who has all the UAVs? *The Guardian*. Retrieved November 6, 2014, from http://www.theguardian.com/news/datablog/2012/aug/03/drone-stocks-by-country

Schröder, A. (2013, May 22). Drohnen-Skandal: De Maizière vertröstet Kritiker auf Juni - SPIEGEL ONLINE. Retrieved November 6, 2014, from http://www.spiegel.de/politik/deutschland/drohnen-skandal-de-maiziere-lehnt-schnelle-konsequenzen-ab-a-901257.html

Slovic, P. (1999). Trust, Emotion, Sex, Politics, And Science: Surveying The Risk-Assessment Battlefield. *Risk Analysis, 19*(4), 689-701.

Samsung Techwin SGR-A1 Sentry Guard Robot. (n.d.). Retrieved November 6, 2014, from h ttp://www.globalsecurity.org/military/world/rok/sgr-a1.htm

Stanford International Human Rights and Conflict Resolution Clinic (IHRCRC) & Global Justice Clinic (GJC) at NYU School of Law. (2012). *Living Under Drones: Death, Injury and Trauma to Civilians from US Drone Practices in Pakistan*.

Self awareness. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Self-awareness#Developmental_stages

Technology - Planetary Resources. (n.d.). Retrieved November 6, 2014, from http://www.planetaryresources.com/technology/

Three laws of robotics. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Three_Laws_of_Robotics

Three laws of robotics. (n.d.). In Wikipedia. Retrieved November 06, 2014, from http://en.wikipedia.org/wiki/Three_Laws_of_Robotics

University Affiliated Research Centers. (2012, November 15). Retrieved November 6, 2014, from http://www.arl.army.mil/www/default.cfm?page=510

Valery, N. (2013, May 14). Automation for the elderly Difference Engine: The caring robot. *The Economist*. Retrieved November 6, 2014, from http://www.economist.com/blogs/babbage/2013/05/automation-elderly

Wall, T., & Monahan, T. (2011). Surveillance And Violence From Afar: The Politics Of Drones And Liminal Security-scapes. *Theoretical Criminology, 15*(3), 239-254.

Wren, A., & Wren, D. (1995). A genetic algorithm for public transport driver scheduling. *Computers & Operations Research, 22*(1), 101-110.

Whitlock, C., & Gellman, B. (2013, September 4). U.S. Documents Detail Al-Qaeda's Efforts to

Fight Back against Drones. *The Washington Post*. Retrieved November 6, 2014, from http://www.highbeam.com/doc/1P2-35093035.html?