9.0 ARTIFICIAL INTELLIGENCE

By Valentina Simonetti

ABSTRACT

Artificial Intelligence (AI) can assume multiple and slightly different definitions between different authors, within the spectrum that goes from simple software for control to complex, not yet existing, evolutionary digital intelligences. In general, we can state that the field of AI embraced several different topics that mostly deals with machine high-level task accomplishment, and that it stands on the framework of software development. The concept of AI is strictly bond to the concept of robot when we have an embodiment of the software component, and AI can be considered the "brain" that drives robot's decision and interactions with users.

Considering the division of AI types in: Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI), where ANI are all the AI actually existing nowadays, REELER research efforts should focus mainly on ANI, that is the one driving all the robots worldwide. Only marginal consideration should be given to possible futuristic AGI. Different ethical studies on AI are more focused on discussion over remotely possible futures of human-like intelligent machines and the issues that they could arise, whereas there is less involvement over the many and already existing context-specific applications of ANI. The main reason that drives this trend is the strong psychological effect that humanoid robots generate on humans even tough issues related to AGI and ASI cannot be fully covered due to the fact that they do not exist yet, and maybe they will never; external appearance and made up expedients may let the external viewer think of actual human-like behaviour of some AI, but we are far from an AI able to perform general and extremely complex human-like cognitive tasks, but, furthermore, we are far from an AI that has the consciousness of what it is doing, where consciousness is considered in its human sense.

For the purposes of REELER project we can say that it is impossible to produce a consistent critical analysis on robots design considering hardware and software separately, where AI is considered the software part of the robot. In some cases AI can be a really complex and manifold piece of software, but still a software, with all the issues related to software functioning, limits, and design best practices. Modern optimized approach to software design in low to high complex software development is the Agile approach that finds its concrete application in the scrum methodology that can be considered a mild, undeclared form of collaborative learning, mainly under the aspect of developers-developers and developers-user interaction and information exchanges.

In general we found sources that support the importance of integration an higher form of collaboration between technical developers, users, and other types of professional roles that are not automatically involved in the design process but that could give fruitful insights on the best design choices to achieve better user acceptance and ethical alignment.

9.1 Opening

According to Meeriam-Webster dictionary, Artificial Intelligence (AI) is "branch of computer science dealing with the simulation of intelligent behavior in computers". Thin definition is, however, misleading, due to the ambiguous definition of the concept of "intelligent behavior" that is not fully explained.

A more open definition is given by Birmingham (Birmingham, 2008) which states that AI is a broadly defined discipline involving computer science, engineering, philosophy, psychology, political science, and a host of other disciplines. Birmingham says that this concept is so broad that is hard to succinctly define and, for sake of brevity it uses the handle of "thinking machines," without commitment to

depths of this thinking. Also Aadhityan (Aadhityan, 2015) gives a broad definition of AI as a universal field with "huge variety of subfields, ranging from the general (learning and perception) to the specific, such as playing chess, solving puzzles, proving some mathematical theorems, writing poetry, taking logical decision driving a car on a crowded street, and diagnosing diseases". Gurkaynak et al. (Gurkaynak, Yilmaz, & Haksever, 2016) uses the classification of AI in three subgroups: Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI). ANI is the AI driving task oriented machines and is the only one actually existing in our world, while AGI and ASI are only names given to something that may exist one day. AGI refers to an AI that can be considered human-like, whereas ASI refers to AI that goes beyond human limits

In general, we can state that the field of AI embraced several different topics that mostly deals with machine high-level task accomplishment, and that it stands on the framework of software development. The concept of AI is strictly bond to the concept of robot when we have an embodiment of the software component, and AI can be considered the "brain" that drives robot's decision and interactions with users.

9.2 Methodology

A preliminary literature search has been performed for the concept 'Artificial Intelligence' to determine how this concept has been studied thus far. The SCOPUS database was used for search over single word (Artificial Intelligence) and combined words (Artificial Intelligence + robot / human / learning / collaborative learning / Cyber Physical System / Intelligent Artificial System / ethnography / anthropology). Most of the articles were excluded from first screening on title and abstract, whereas final significant articles were selected after full reading.

9.3 Discussion

The REELER project aim at improving robot design process to embed ethical measures and end user involvement since the very beginning of the robot development, with the outcome of a more ethical and more performing product. In order to address this goal, we focused on the articles that addressed the issues related to ethics, economic impact, design, usability, and collaborative learning which is a core concept to be investigated by the project.

Some authors addressed the important topic of ethical alignment in AI design, showing a pluralism of viewpoints.

Waser (Waser, 2008) says that the development of ethical AI should follow the 5 S's (Simple, Safe, Stable, Self-correcting and Sensitive to current human thinking, intuition, and feelings). It also states that there could be four fundamental reasons why a rational intelligence would not follow a society's dictates regardless of what they were: over-riding self-protection (fear); selfishness (greed); unfairness (error) on society's part; or error on the entity's part. If these four point are covered, the path to safe AI should be clear.

Birmingham (Birmingham, Towards an Understanding of Artificial Intelligence and Its Application to Ethics, 2008) analyzes the present and possible future situation of machine abilities to perform human-like tasks both regarding the aspect of actual technical capabilities and also the aspect of actual machine awareness of the actions performed. It sees a not too distant future that will "allow AI systems to have computational abilities, both in terms of speed of computation and access to vast amounts of information, that far outstrip human performance in at least a great many functions" and the same will eventually hold for sensory input; that is, AI machines will likely have the ability to perceive the

world in a way similar to human perception, including understanding art, music, and literature (Cope, 1991). The main issue that Birmingham rise is whether this clever, extraordinary simulation actually equal humanness, and contend that no matter how sophisticated, a simulation does not become the thing (Searle, 1980) even though it seems that AI is considered, for some reason, as a special exception to this principle that we would never afford to other very clever simulations. Despite this common consideration of AI, it contends that no artificial thing can become the moral equivalent of a human being and, therefore, if we consider AI systems to be like other software systems, then they fall under the same moral concerns as other computation systems. Developers need to ensure they are safe, reliable, secure, legal, protect privacy, do not harm persons or property, and do what they are supposed to do (Tavani, 2003). However, Birmingham foresees a possible futuristic scenario where AI system could leave the realm of "regular software" and that is the remote possibility in which we can create a human-like artificial genome driven by a program able to accurately reproduce the physiological mechanisms that drive genome expression. While it contends that this particular robot, no matter how perfect the system is at simulation, is not a person, such a system raises serious moral issues. The ACM Code of Ethics (1.1)1 asserts that one should "contribute to society and human wellbeing" and the above example would not serve to promote human well-being.

Perri (Perri, 2001) sees that important issues about AI arise in connection with the prospect of robotic and digital agent systems taking socially significant decisions autonomously and it is surprised by the popular anxiety about AI raised during the 1990s, because that decade was one in which developed societies went through panics about many other, and more concrete, technological risks (i.e. nuclear power, alarm about food biotechnology, etc.). Perri says that "the key to understanding these surprising cycles in concern is to recognize that social fears are patently not simply based upon the calculation of probabilities based upon evidence. Rather they reflect the shifting patterns of social organization and the worldviews thrown up by rival forms of social organization, each of which pin their basic concerns on the technologies that are to hand". That means that anxiety comes from the blurring of boundaries between concepts, and, like biotechnology arouses anxiety blurring the boundary between "natural" and "artificial", AI blurs the boundary between "tool" and "agent". Whereas a tool is supposed to be dependent on human action, an agent has the capacity to act and make judgements independently. "Groups within societies grow anxious about such boundaries, not necessarily when particular technologies actually threaten them, but when those boundaries feel most under threat generally". Perri rises the topic of political response to popular risk perception and it lists the different institutional styles of social organization (hierarchy, isolate, individualism, egalitarian) with their correspondent worldview (respectively: bureaucratic, fatalist, entrepreneurial, chiliastic) (see Figure below). Each of these institutional solidarities produces a distinct cognitive bias in how risks, such as those attached to AI, are perceived. The author state that we have little choice but to give some but not excessive recognition to each solidarity, in more or less unstable settlements between them.

In general, Perri considers that the doomsday scenario should not detain us long, but instead consider all the aspects to raise more practical issues of control and accountability of machines in the management of risk, which do deserve some practical attention. In general, it says that, if and when "machine-persons" are produced, or even if they emerge by Lamarckian evolution from AI systems unaided by human intervention after the human production of AI systems that are significantly less

than persons, then there is no reason to suppose that such machine-persons would be malevolent, but rather every reason to suppose that they would be no more and no less benign than human persons, who are capable of great evil but also of co-operation and acceptance of moral responsibilities on an awesome scale. An issue that rises in our present-day society is whether to subject AI systems to the kinds of accountability that are appropriate to a tool or to an agent. Some social pressure should be made on technology communities to think carefully about the institutionalization of their own professional ethics. It is that challenge that must be made to the AI research and development community in the coming decades.

Dautenhahn (Dautenhahn, 2007) takes a critical position against the focus on robots' rights. So far robots are not sentient beings and they will maybe never be. It says that there are many more pressing issues that should dominate the discussions, for example, regarding robots companions: bonding is a bi-directional relationship, we 'get something back' when bonding with a person. What are our rewards regarding our emotional needs when bonding with a robot? The effects that are usually not considered are psychological: how does interaction with machines effect people's sense of self, sense of autonomy and control, sense of belonging, sense of friendship and love etc. These are scientific questions that can be studied experimentally and should be investigated extensively long before actual products are being put on the market

The software nature of AI makes it follow the standard procedures and typical issues of standard software design. When it comes to reach higher level of decision making, AI and robot designers should consider also other aspects in the design process.

Starting from its classification of AI between ANI, AGI, and ASI, Gurkaynak (Gurkaynak, Yilmaz, & Haksever, 2016) talks about the fact that we cannot foreseen the moment in time when the singularity that will lead to AGI will happen, if it will. The author examines the topic of AI regulation and says that there are very few laws or regulations that address the challenges raised by AIs, and no courts appear to have developed standards so far, addressing who is legally responsible if an AI causes harm. National and international laws do not recognize AI as a legal person, therefore, current legal systems cannot hold them liable for the damages they might cause. However, what if an AI was fully autonomous and aware of its actions, causing harm knowingly and willingly? This in the case the singularity occurs. This issue is linked to the debate on consciousness in machines and a conscious AI should naturally be liable for its actions, but we are far from that technological level of artificial creation. The Great AI Panic of 2015' (Sofge, 2015) led an institution called 'Future of Life Institute® (FLI)' to issue an open letter signed by Elon Musk, Stephen Hawking, hundreds of AI researchers in addition to many individuals representing U.S. government (https://futureoflife.org/ai-open-letter/). FLI used statements such as 'AI systems must do what we want them to do', 'We should identify research directions that can maximize societal benefits' and 'AI super-intelligence will not act with human wishes and will threaten humanity' and refers to a specific document "Research Priorities for Robust and Beneficial Artificial Intelligence" (Russell, Dewey, & Tegmark, 2016). The document focuses on both short and long- term research priorities analysing core issues and strategies in the field of economic impact, technological progress, law, and ethics. For the authors, short terms research priorities concerns:

- Optimizing AI's economic impact, in a way that maximizes the economic benefits while mitigating adverse effects, which could include increased inequality and unemployment;

- Law and ethics research, that should consider issues and drawbacks of each specific context that AIs could affects (i.e. Autonomous Vehicles) and should aim at generating policies that minimize risks and maximize human benefits
- Computer science research for robust AI, that means creating systems that robustly behave as intended considering verification, validity, security, and control.

Regarding long-term research priorities, it first discuss the different ideas that researchers have regarding the possibility of developing a system with human-like cognitive abilities and perhaps able to surpass humans (basically AGI and ASI). Even though lot of researchers consider that possibility highly improbable, it says that "to justify a modest investment in this AI robustness research, this probability need not be high, merely nonnegligible, just as a modest investment in home insurance is justified by a nonnegligible probability of the home burning down". In the practice of long term priorities, the authors focus on the improvement of the, already discussed, verification, security, and control with the future perspectives of dealing with systems able to operate on always bigger amount of data, greater autonomy, and more sensible security and control issues.

Perri (Perri, 2001) tries to formalize the decision-making process with standard stages, which are:

- Allocating attention to problems-Diagnosis
- Risk appraisal
- Action set
- Choice considerations
- Preference ranking or decision protocol
- Recommending action and, if appropriate, taking action
- Decision remittal

It says also that there are limits to the soundness of base of any decision-making procedure, as well as to the possibilities for the elimination of judgment. Great attention should be put in considering how far AI are capable higher technical quality decision making than humans. The concept of intelligence itself can have different interpretations: those who see it as general ability for acquiring knowledge, reasoning skills and problem solving, and those who recognize a plural and wider range of capacities including kinetic, spatial, communicative, musical and other context-specific or skill-specific capabilities. Whichever approach is taken the key issue is that pattern recognition, abstract reasoning and other such cognitive abilities will not suffice to deal with the challenges of problem structuring and judgment under conditions of uncertainty. Perri supports the idea of enforcing accountability in AI and robots by embodying some commitments, that can be overridden in certain kinds of emergencies. It states that "the design must serve as the artificial equivalent of the human professional taking an oath binding her or him to the code of professional conduct". In general, it considers the risk management challenge an exacerbation of the eternal problems of technological risk.

An interesting point of view for REELER investigation is the one of Lund (Lund, 2004), that underlines the importance to find support in interaction studies and psychology to obtain a more rewarding human-robot interaction. It states that the focus must be on both creating new possibilities for flexible

robots from the engineering point of view and creating new interactions from the humanistic point of view. Lund report the fact that fields such as psychology, communication, and educational pedagogy are now entering the field providing important methods and tools for the verification of importance of factors involved in the human interaction with robotic systems, bringing as examples [B. Caci, M. Cardaci, and H. H. Lund, "Assessing Educational Robotics by the 'Robot Edutainment Questionnaire'," Maersk Mc-Kinney Moller Inst. Prod. Technol., Univ. Southern Denmark] [B. Caci and A. D'Amico, "Children's cognitive abilities in construction and programming robots," presented at the 11th IEEE Int.Workshop Robot and Human Interactive Communication, Berlin, Germany, 2002.] which developed the Robot Edutainment Grid (REG) as a questionnaire for pupils as a tool for performing quantitative studies of edutainment robotics. The REG can be seen as a tool of investigation of robotic projects not only in order to analyze the educative value of them, but also in order to develop a new methodology based on the principles of the cognitive psychology. The author stress the focus on the important concept of user-guided evolutionary robotics. Below is shown the software structure of the behavior-based control for a specific humanoid robot. The different layers have increasing behavioral complexity from the bottom up.

Barakova and Lourens (Barakova & Lourens, 2013) aim at creating a robot with AI able to interact productively with autistic children, being, at the same time, usable and flexible for therapists. The project faces the "dual user" problem: a domain specialist that uses the robot as a tool to augment his/her practice, and the client or patient that is traditionally served by the domain specialist and who is now (partially) served by a robot. The aims: enabling and empowering the therapists to train the children, as well as engagement of the robot in natural interaction with the children. From the perspective of the therapists the reliability of the technology is the major issue in these applications. This was the reason to provide an option for remote robot control. A remote controller is also needed to fake intelligent behavior when the creation of such is beyond the achievements of the state of the art research.

In the article from Nature "Robotics: Ethics of artificial intelligence" Sabine Hauert, lecturer in robotics for the University of Bristol suggests that AI and robotics stakeholders worldwide should pool a small portion of their budgets (say 0.1%) to bring together these disjointed communications and enable the field to speak more loudly. In the same article, Russ Altman, professor of bioengineering, genetics, medicine and computer science of Stanford University, says that AI researchers who create the infrastructure and technical capabilities for these systems need to engage doctors, nurses, patients and others to understand how they will be used, and used fairly. Manuela Veloso, professor of computer science, Carnegie Mellon University, wishes for an improvement in collaboration between different expert in the field of interest (technical and non-technical), but it also wishes this collaboration to be between robots and humans, with robots ask for help from humans or from the Internet. Robots should know when to ask for help and how to express their inner workings (concept of "symbiotic autonomy").

Also in Barakova and Lourens is said that collaboration between therapists, robot specialists, software developers and human robot interaction researchers is a necessity. The actual implementation requires a careful negotiation of the priorities in the development of real-life applications. The measures used to achieve these goals and develop fruitful interaction between robots and children

with autism were: (a) empowering the therapists with tools to use and reprogram the robot and (b) acquisition of domain specific knowledge from the therapists in order to design robot-child interaction scenarios that accomplish specific learning goals and (c) creation of community of domain specialists and parents that, by the means of social computing methods, will direct the process of scenario creation.

In Yao et al. (Yao, Zhang, & Chen, 2015), a study implementing an AI for library support, a questionnaire survey was conducted, which allowed the researcher to examine if the initial targets and claimed functions were achieved by the system developed, how the new service was accepted by users, and obtain feedback after careful user testing. Semi-structured in-depth interviews were conducted with key members of the library and enthusiastic patrons, exploring their attitudes, thinking, feeling and deep needs, and the potential application of the system, among others.

Considering the software nature of AI it is interesting a paper from Abdullah et al. (Abdullah, Sharp, & Honiden, 2011) that analyzes how artefacts influence the construction of communications and contexts during collaboration in an agile software development team. Agile development can be considered a particular type of collaborative learning where team working is necessary and fruitful.

Social and economic impact of AI are tackled by different authors.

In the aforementioned article from Nature “Robotics: Ethics of artificial intelligence” professor Russ Altman states that “AI technologies could exacerbate existing health-care disparities and create new ones unless they are implemented in a way that allows all patients to benefit, responsibility of the government and those who develop the technology and support the research to ensure that AI technologies are distributed equally”.

Perri states that every time human beings have developed sufficiently reliable technology for specific purposes and, once they have developed an institutional framework to resolve the conflicts that it generates, they usually vacate those economic arenas, despite the occasionally serious Luddite backlash. Autonomous intelligent machines are not necessarily generalists, nor do they necessarily bring any fundamental change in the economic that make some specialism. It says, then, that there is every reason to think that human beings should be able to face changes brought by AI, as they have always done, and shift the range of ecological niches that we ourselves occupy so that we are not put out of business by these technologies, any more than we were by previous less intelligent and less autonomous technologies.

Gurkaynak point of view is liberal: it says that the economic efficiency potentials of AI should be set entirely free in this moment of the history, allowing AI researchers to actively and aggressively research appropriate goals for them, which would not result in the extinction of humankind.

9.4 Conclusion

From the analysis of bibliographic sources it is clear that the concept of AI can assume multiple and slightly different definitions between different authors, within the spectrum that goes from simple software for control to complex, not yet existing, evolutionary digital intelligences. For the purposes of the REELER project we should consider the approach followed also by (Gurkaynak, Yilmaz, & Haksever, 2016) that, as we already discussed about, divides the field of possible AI in: ANI, AGI, and

ASI, where ANI are all the AI actually existing nowadays, whereas AGI and ASI are, more or less futuristic, views of possible evolution in the field. Following the guidelines given by "Research Priorities for Robust and Beneficial Artificial Intelligence" (Russell, Dewey, & Tegmark, 2016), REELER research efforts should focus mainly on ANI, that is the one driving all the robots worldwide. Always according to the guidelines, being the possibility of developing AGI something that lot of researchers consider highly improbable, but still potentially possible, we should give marginal consideration also to that. It is interesting to report that the topics on AI ethics that are more tackled are, however, mainly focused on discussion over remotely possible futures of human-like intelligent machines and the issues that they could arise, whereas there is less involvement over the many context-specific applications of ANI. The main reason that drives this trend is the strong psychological effect that humanoid robots generate on humans; an interesting discussion on this is made by (Geraci, 2007), that considers the robot as a "numinous experience" and discuss also the effects of and on science fiction. The "AGI Sputnik Moment", as coined by Ben Goertzel is something that is still far from reality, so far. Interesting approaches in the direction of the development of human-like intelligence support the exploitation of evolutionary algorithms, as it is said by (Adami,2015). In any case these discussion over AGI and ASI concerns topics and issues that cannot be fully covered due to the fact that they do not exist yet, and maybe they will never; external appearance and made up expedients may let the external viewer think of actual human-like behaviour, but we are far from an AI able to perform general and extremely complex human-like cognitive tasks, but, furthermore, we are far from an AI that has the consciousness of what it is doing, where consciousness is considered in its human sense.

For the purposes of REELER project we can say that it is impossible to produce a consistent critical analysis on robots design considering hardware and software separately. These two parts extremely interconnected and one impose limitations and constrains over the other, deeply influencing the functionalities and also the appearance of the final product. In some cases AI can be a really complex and manifold piece of software, but still a software, with all the issues related to software functioning, limits, and design best practices. Modern optimized approach to software design in low to high complex software development is the Agile approach, which is regulated by the Agile Manifesto (http://agilemanifesto.org/m) that finds its concrete application in the scrum methodology, for details (Schwaber, 1997). The agile manifesto principles give priority to "Individuals and interactions over processes and tools, working software over comprehensive documentation, customer collaboration over contract negotiation, responding to change over following a plan", and its exploitation in the scrum process can be considered a mild, undeclared form of collaborative learning, mainly under the aspect of developers-developers and developers-user interaction and information exchanges.

In general we found different sources that support the importance of integration an higher form of collaboration between technical developers, users, and other types of professional roles that are not automatically involved in the design process but that could give fruitful insights on the best design choices to achieve better user acceptance and ethical alignment (Lund, 2004).

References

Aadhityan, A. (2015). A novel method for implementing Artificial Intelligence, Cloud and Internet of Things in Robots. ICIIECS 2015 - 2015 IEEE International Conference on Innovations in Information, Embedded and Communication Systems, 2-s2.0-84956984387.

Abdullah, N. N., Sharp, H., & Honiden, S. (2011). How Artefacts Influence the Construction of Communications and Contexts during Collaboration in an Agile Software Development Team. FLAIRS Conference.

Adami, C. (2015, #may#). Artificial intelligence: Robots with instincts. Nature, 521(7553), 426-427. Tratto da http://dx.doi.org/10.1038/521426a

Barakova, E., & Lourens, T. (2013). Interplay between Natural and Artificial Intelligence in Training Autistic Children with Robots. In J. M. Ferr{\'a}ndez Vicente, J. R. {\'A}lvarez S{\'a}nchez, F. de la Paz L{\'o}pez, & F. J. Toledo Moreo (A cura di), Natural and Artificial Models in Computation and Biology: 5th International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2013, Mallorca, Spain, June 10-14, 2013. Proceedings, Part I (p. 161-170). Berlin, Heidelberg: Springer Berlin Heidelberg. Tratto da http://dx.doi.org/10.1007/978-3-642-38637-4_17

Birmingham, W. (2008). TOWARDS AN UNDERSTANDING OF ARTIFICIAL INTELLIGENCE AND ITS APPLICATION TO ETHICS. American Society for Engineering Education.

Cope, D. (1991). Computers and Musical Style. Madison, WI, USA: A-R Editions, Inc.

Dautenhahn, K. (2007). A Paradigm Shift in Artificial Intelligence: Why Social Intelligence Matters in the Design and Development of Robots with Human-Like Intelligence. In M. Lungarella, F. Iida, J. Bongard, & R. Pfeifer (A cura di), 50 Years of Artificial Intelligence: Essays Dedicated to the 50th Anniversary of Artificial Intelligence (p. 288-302). Berlin, Heidelberg: Springer Berlin Heidelberg. Tratto da http://dx.doi.org/10.1007/978-3-540-77296-5_26

Geraci, R. M. (2007). Robots and the Sacred in Science and Science Fiction: Theological Implications of Artificial Intelligence. Zygon®, 42(4), 961-980. Tratto da http://dx.doi.org/10.1111/j.1467-9744.2007.00883.x

Gurkaynak, G., Yilmaz, I., & Haksever, G. (2016). Stifling artificial intelligence: Human perils. Computer Law \& Security Review, 32(5), 749-758. Tratto da http://www.sciencedirect.com/science/article/pii/S0267364916300814

http://agilemanifesto.org/. (s.d.). http://agilemanifesto.org/.

https://www.acm.org/about-acm/acm-code-of-ethics-and-professional-conduct. (s.d.). https://www.acm.org/about-acm/acm-code-of-ethics-and-professional-conduct.

Lund, H. H. (2004, Nov). Modern artificial intelligence for human-robot interaction. Proceedings of the IEEE, 92(11), 1821-1838.

Perri, 6. (2001). ETHICS, REGULATION AND THE NEW ARTIFICIAL INTELLIGENCE, PART I: ACCOUNTABILITY AND POWER. Information, Communication \& Society, 4(2), 199-229. Tratto da http://www.tandfonline.com/doi/abs/10.1080/713768525

Russell, S. J., Dewey, D., & Tegmark, M. (2016). Research Priorities for Robust and Beneficial Artificial Intelligence. CoRR, abs/1602.03506. Tratto da http://arxiv.org/abs/1602.03506

Schwaber, K. (1997). SCRUM Development Process. In J. Sutherland, C. Casanave, J. Miller, P. Patel, & G. Hollowell (A cura di), Business Object Design and Implementation: OOPSLA '95 Workshop Proceedings 16 October 1995, Austin, Texas (p. 117-134). London: Springer London. Tratto da http://dx.doi.org/10.1007/978-1-4471-0947-1_11

Searle, J. R. (1980). Minds, Brains and Programs. Behavioral and Brain Sciences, 3(3), 417-57.

Sofge, E. (2015). An Open Letter To Everyone Tricked Into Fearing Artificial Intelligence. An Open Letter To Everyone Tricked Into Fearing Artificial Intelligence.

Tavani, H. T. (2003). Ethics and Technology: Ethical Issues in an Age of Information and Communication Technology. Acm Sigcas Computers and Society, 33(3), 1.

Waser, M. R. (2008). Discovering the foundations of a universal system of ethics as a road to safe artificial intelligence. Books International.

Yao, F., Zhang, C., & Chen, W. (2015). Smart talking robot Xiaotu: participatory library service based on artificial intelligence. Library Hi Tech, 33(2), 245-260. Tratto da https://doi.org/10.1108/LHT-02-2015-0010.