



ARTIFICIAL INTELLIGENCE IN SOFTWARE TESTING

Bachelor's thesis

Information and Communication Technology, Bioeconomy

Spring 2024

Heini Jauhainen

Tieto- ja viestintätekniikka, Biotalous

Tekijä Heini Jauhiainen

Työn nimi Artificial intelligence in software testing

Ohjaaja Ari Hietala (HAMK), Lauri Kuusisto (Etteplan Finland Oy)

Tiivistelmä

Vuosi 2024

Opinnäytetyön tarkoituksena oli selvittää tekoälyä hyödyntäviä ohjelmistoja, jotka soveltuvat ohjelmistotestaukseen. Tutkimuskysymyksinä oli, minkälaisia työkaluja markkinoilta tällä hetkellä löytyy, miten ne toimivat ja millaisia kustannuksia niiden käytöstä syntyy.

Työn toimeksiantajana toimi Etteplan Finland Oy ja osana opinnäytetyötä teetettiin kysely kyseisen yrityksen ohjelmistotestaajille Suomessa, Puolassa ja Ruotsissa. Kyselyyn tuli yhteensä 26 vastausta, joista suurin osa tuli Suomesta. Ruotsista ei tullut kyselyyn yhtään vastausta. Kyselyn tulosten perusteella, tekoälyä ei juurikaan tietävästi hyödynnetä ohjelmistotestauksessa. Suurimpina haasteina tekoälyn käyttöönottoon nähdään tekoälyyn ja sen käyttöön liittyvä tiedon- ja taidonpuute, ymmärrys sen hyödyistä sekä se, ettei tiedetä mitä tekoälyn työkaluja voisi käyttää. Muutama vastaaja kertoi käyttävänsä Microsoft Copilotia, mutta muita työkaluja ei vastauksista noussut esille.

Vastausten perusteella tekoälyn toivottiin tuovan apua testitapausten, testiskriptin ja testidatan luomiseen. Vastausten pohjalta kerättiin tietoa muutamasta tekoälyn sovelluksesta, joita tällä hetkellä käytetään ohjelmistotestauksessa ja joiden ominaisuuksissa otettiin huomioon kyselyn vastausten mukaisia ominaisuuksia. Kyselyn vastauksista nousi myös esille tarve lisätä toimeksiantaja yrityksen ohjelmistotestaajien tieto/taito -pohjaa tekoälyn perusteiden osalta.

Avainsanat Ohjelmistokehitys, ohjelmistotestaus, tekoäly

Sivut 37 sivua ja liitteitä 1 sivu

Information and Communication Technology, Bioeconomy
Author Heini Jauhiainen
Subject Artificial Intelligence in Software Testing
Supervisors Ari Hietala (HAMK), Lauri Kuusisto (Etteplan Finland Oy)

Abstract
Year 2024

The purpose of the thesis was to find out software testing software's that utilizes artificial intelligence (AI). The research questions were: what kind of tools are currently available on the market, how they work and what kind of costs arise from using them?

The commissioner of the thesis was Etteplan Finland Oy and as part of the thesis, a survey was carried out to the software testers of that company in Finland, Poland, and Sweden. The survey received a total of 26 responses, most of them came from Finland. There were no responses to the survey from Sweden. Based on the results of the survey, artificial intelligence is not widely utilized in software testing. The biggest challenges to the implementation of artificial intelligence are seen as the lack of knowledge and skills related to artificial intelligence and its use, understanding of its benefits, and it is not known which artificial intelligence tools could be used. A few respondents said they use Copilot, but no other tools emerged from the answers.

Based on the answers, it is hoped that artificial intelligence would help in creating test cases, test scripts and test data. Based on the answers, information was collected about a few artificial intelligence applications currently used in software testing and whose properties were considered according to the answers to the survey. The responses to the survey also highlighted the need to increase the knowledge base of the commissioning company's software testers regarding the basics of artificial intelligence.

Keywords Artificial intelligence, software development, software testing
Pages 37 pages and appendices 1 page

Content

1	Introduction	1
2	Software testing	2
2.1	Manual and Automated Testing	2
2.2	Levels of testing	3
2.2.1	Unit Testing	4
2.2.2	Integration Testing	4
2.2.3	End-to-End Testing	4
3	Artificial intelligence (AI)	5
3.1	History of AI	6
3.2	Benefits of using AI	10
3.3	Disadvantages of using AI	12
3.4	Legislation	14
3.5	AI in software testing	16
3.6	AI tools for software testing	17
3.6.1	Testim	17
3.6.2	Applitools	18
3.6.3	Testsigma	18
3.6.4	Aqua	19
4	Use of AI in software testing at Etteplan Finland Oy	20
4.1	Schedule of the thesis	20
4.2	Survey	21
4.3	Results	22
5	Conclusions, reflection, and future development	27
	References	32

List of Figures

Figure 1, Mike Cohn's testing pyramid on the left and the Google's testing pyramid on the right. (Cohn, 2009, p. 311 ; Wacker, 2015)	3
Figure 2 Testsigma plans (Testsigma, n.d.-b).	19
Figure 3 Schedule of thesis.....	21
Figure 4 Respondents office locations.	22
Figure 5 How much AI is being used in software testing.	23
Figure 6 Efficiency impact of AI tools.	23
Figure 7 Number of respondents using AI in testing.....	24
Figure 8 How AI is used in testing.....	24
Figure 9 Respondents wished AI to help or solve.	25
Figure 10 Reasons why AI is not being used in testing.	26
Figure 11 Respondents' general feelings using AI.	27
Figure 12 Interest in using AI in their work in the future.....	27

Appendices

Appendix 1. The survey questions

1 Introduction

Artificial intelligence (AI) is currently the existing technological innovation, capturing both industry and academic sides with its changing potential. From enhancing everyday applications, such as virtual assistants and recommendation systems to pioneering advancements in autonomous vehicles and healthcare diagnostics, AI's impact is far-reaching. The growing computational power and enormous amount of data available today, have propelled AI from theoretical research into practical, widespread use.

One area to benefit from AI is software testing. Software testing methods are often time-consuming and prone to human error. AI can automate and enhance this process, providing tools for more efficient bug detection, regression testing, and code analysis. Machine learning algorithms can predict potential problem areas in codebases, while natural language processing can improve the understanding and testing of user requirements. By integrating AI into software testing, higher quality software, faster release cycles, and reduced costs can be achieved, leading the way for more reliable and innovative software solutions.

(Draniceanu, n.d.)

This thesis aims to find out what kind of software or applications that are suitable for software testing and utilize AI, are currently on the market. The purpose is also to find out the cost structure of these applications. These also form the research questions. The thesis is carried out for Etteplan Finland Oy, which is a company in the field of technology. As a part of the thesis, a survey will be made to the company's software testers to explore how AI is currently used in their work as software testers and what are the testing-related topics that they wish AI could help with.

From the Etteplan point of view, it is important to keep up with the technology development to meet the customer's needs. When value is created for the customers by reducing costs, speeding up processes, or allocating human resources better, it will also bring value to Etteplan. This results in higher customer satisfaction and builds trust in the software development process. Customers are more likely to return for future projects and recommend the service to others, fostering long-term business relationships which is important for the increase of the company's income.

2 Software testing

When there is a need to check and have confidence that the product, such as software, application, or device is working as it should and that it has as minimum number of dysfunctionalities or failures (bugs) as possible, testing of the product before it ends up into a production is a good way to do it. Software testing also aims to improve product performance. (Leloudas, 2023)

Software testing is a part of software production. Software testers are not just testing, but they can also code and write documentation. Software tester's job description can be different depending on the product to be tested. The testing target defines what kind of testing is needed. For example, a software tester in the gaming industry has a different priority than a software tester who works for the tax administration. (Kasurinen, 2013)

2.1 Manual and Automated Testing

In manual testing, the software is tested as how the end user would use the software. The tester is confirming that the software is working before it is released. Human tester plays the role of an end user and no automation is used in the testing. In test automation, some specific software is used to execute test cases and comparing the actual results to what was expected to happen.

Test automation can save resources such as money and time because when testing is automated, it can be done outside of human working hours. Automation takes time in the planning phase when test cases need to be designed and the automation process is built, but after those are done, automation executes tests faster than human. (Leloudas, 2023)

In some cases, test automation is easier to execute than manual testing. Load, performance, and stress tests are good examples of those cases. These types of testing are usually done to see how system works under pressure. An example of an such case is when hundreds of people are trying to buy tickets from online store. It is much easier to simulate such behaviour in automated environment, than having hundreds of real people trying to check can the online store manage or not under that load. (Srinivasan & Gopalaswamy, 2007)

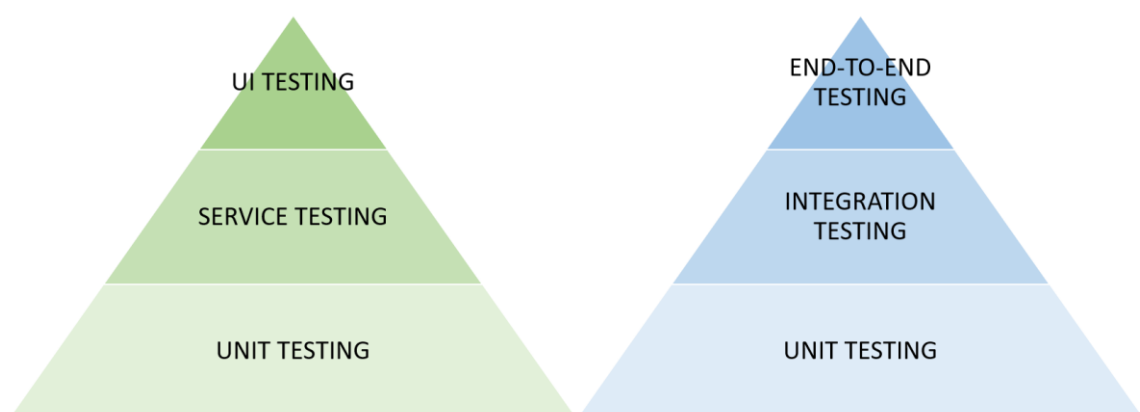
With test automation, those repetitive test cases can be executed faster since machines are not getting disinterested. This also makes it possible for the human tester to concentrate on

other work tasks, such as explorative testing, and finding errors in that way. Test automation is not capable of functions that are not specific or specified. With Manual testing it is also possible to execute ad-hoc testing, which test automation cannot do. Manual testers can also evaluate the user experience and perform usability testing. (Srinivasan & Gopaldaswamy, 2007)

2.2 Levels of testing

In 2009 Mike Cohn presented the three levels of testing in his book called *Succeeding with Agile*. Those levels form the test automation pyramid. On the bottom of that pyramid is unit testing, the middle layer is service testing, and the top layer is UI testing (Cohn, 2009, p. 311). In 2015 Google released their version of that pyramid, where unit testing is at the bottom, integration testing is on the middle and end-to-end testing is the top layer, as depicted in Figure 1. What is common for both testing pyramids, is that most testing should happen on the bottom layer. That is the place where the software testers should have the highest amount of written test cases because that creates the foundation of your testing. After the foundation is ready, the middle layer includes fewer amount of testing than the bottom layer and on the top layer, there should be only a small amount of testing left (ISTQB, 2023, p.49; Wacker, 2015).

Figure 1, Mike Cohn's testing pyramid on the left and the Google's testing pyramid on the right. (Cohn, 2009, p. 311 ; Wacker, 2015)



2.2.1 Unit Testing

In the unit testing phase, the small functional section of the software component is being tested. That is mainly done by software developers since it contains operating with source code. The purpose of unit testing is to have evidence so that every part of the software is working right. As the development is in its early phase, most bugs and failures should be spotted here. Since the amount of code is lower in this point it is also easier to make corrections, than it might be in later phases of development. Finding bugs as early as possible also saves time and money. Unit testing also might minimize poorly written code, because unit testing is easier when the quality of the code is as best as possible (Leloudas, 2023).

The downside of the unit testing is that it might be time-consuming. Every time changes to the code are made, unit testing needs to be redone to be sure that old functionalities are still working. Version control needs to be in good shape so that it is easier to go back if new changes are not working and to be sure what was the change that caused the failure. Lack of time, resources, or knowledge can also prevent the unit testing phase to be less successful. It is also important to be aware, that unit testing alone is not enough and other kind of methods are needed. (VALA, 2023)

2.2.2 Integration Testing

When there is enough confidence and evidence after unit testing phase, that those small functional sections of software components are working as they should, the next phase of testing, is integration testing. In this step, those software components are combined into a larger entity. Combining can cause failures such as data transferring errors and data losses. Integration testing validates, that all components work together (Leloudas, 2023).

The idea behind integration testing is, to start from the assumption where sections are added to a product that has already been proven to work as expected and desired, and the entire product is still functioning as expected. In integration testing the whole product is not ready, but it has a larger number of functionalities that are working together. (Kasurinen, 2013)

2.2.3 End-to-End Testing

End-to-End testing comes into the picture when integrations are tested. In this phase, the software performance and how the software is used in the actual environment is tested. End-

to-end testing mimics how a software user interacts with software within a fully operational application environment. This testing ensures that different user activities function correctly, from simple tasks, such as opening a webpage or logging in, to more intricate processes, such as checking email notifications and processing online payments. (Pittet, n.d.)

While end-to-end tests are highly beneficial, they come with high costs and can be challenging to maintain when automated. Therefore, it is advisable to focus on a selected number of crucial end-to-end tests and predominantly use lower-level testing methods, such as unit and integration tests to detect issues. (Pittet, n.d.)

Quality assurance (QA) teams usually conduct end-to-end testing in specialized test environments, typically after completing functional and system tests. This testing approach begins from the user's point of view, simulating common tasks the application performs. (Gillis, n.d.)

3 Artificial intelligence (AI)

Artificial intelligence (AI) is a part of computer science and it aims to create systems that are capable of performing tasks that have needed human intelligence before. These tasks include learning, reasoning, problem-solving, perception, understanding natural language, and even exhibiting creativity. AI systems use various techniques to mimic or replicate human cognitive functions, and they can be broadly categorized into two types: narrow AI and general AI. (China, 2024)

A narrow type of AI is designed to perform a specific function or a set of functions that has a relation. Examples include virtual assistants such as Siri and Alexa, recommendation systems on platforms, such as Netflix and Amazon, and autonomous vehicles. Narrow AI systems are highly specialized and can outperform humans in their designated tasks but lack generalization beyond those specific functions. General AI refers to systems that can understand, learn and apply knowledge such a human being. These systems would have the capacity for general cognitive abilities and could perform any intellectual task, that a human can. General AI remains largely theoretical, and it is a major focus of ongoing research and development. (IBM, n.d.-a; IBM,n.d.-b)

Between the 1950s and 1970s AI was about neural networks. The initial aim of the neural network approach was to mimic the problem-solving capabilities of the human brain computationally. However, as time progressed, researchers began directing their attention

toward tailoring neural networks to address particular tasks, departing from a purely biological perspective. (SAS, n.d.-a)

Between the 1980s and 2010s AI development started talking about machine learning, where algorithms could learn patterns from data. Techniques, such as neural networks, support vector machines, and decision trees gained prominence. The rise of the internet and digital data provided vast amounts of data, which were crucial for training machine learning models. (SAS, n.d.-d)

In 2011-2020 deep learning was the high theme in AI. The development of convolutional neural networks (CNNs) and recurrent neural networks (RNNs), led to significant improvements in tasks such as image recognition, speech processing and natural language understanding. The availability of powerful GPUs and distributed computing resources enabled the training of large-scale neural networks. (SAS, n.d.-b)

AI today is generative. The development of models such as GPT-3, its successors and DALL-E, demonstrated AI's capability to generate coherent text and create images from textual descriptions. AI started to assist in scientific research, drug discovery, and creative fields such as art and music. The deployment of AI models on edge devices, such as smartphones and IoT devices, became more prevalent, enabling real-time processing and reducing latency. (SAS, n.d.-c)

3.1 History of AI

The history of AI can be said to begin in 1950 when Alan Turing's book *Computing Machinery and Intelligence* came out. Alan introduced the Turing test, which basic idea is to check if machines can think. In this test, a human asks questions from two other participants and tries to decide which one is a computer, based on their answers. One of those participants is a computer and the other is a control person. The questions are asked by typing on a keyboard and the answer is shown on the screen. The test is performed on a large group of people. If a greater number of human questioners cannot decide which one is the computer and which one is a control person, then the computer is considered to be intelligent and able to think. (Taulli, 2019)

In 1952 Arthur Samuel developed a computer program that played checkers. The program was capable of autonomous learning and improving its gaming process. This has been the starting point for future AI algorithm development. Samuel worked as an engineer at IBM and

the machine he used for checkers program development, was IBM 701. That machine had memory limitations and Samuel used a technique called alpha-beta pruning, which is usually combined with the Minimax algorithm and is used in two-player computer games such as tic-tac-toe and chess. In the Minimax algorithm the computer goes through the best possible moves by evaluating the computer's best moves and worst moves of an opponent. Alpha-beta-pruning ignores those worst scenarios what that Minimax algorithm found out. Samuel also invented the term *machine learning* in 1959, when he was referring to the chess game program, in which a computer would play the game better than the person who programmed it. (IBM, n.d.-a; Jain, 2024)

Three years later in 1955, John McCarthy, Marvin Minsky, Claude Shannon and Nathaniel Rochester had an idea of having a summer conference around the topic of AI. A year later that conference was held at Dartmouth College in Hanover. This meeting has later been referred to as the *birthplace of AI*. At that time, McCarthy was a mathematics college teacher, and he has been called the "father of AI". He also developed the programming language called LISP; a popular programming language used in AI research. Minsky worked in MIT's AI laboratory and had a key role in understanding human and machine intelligence. Rochester was an engineer at IBM, and he also participated in the development of IBM 701. Shannon is nowadays known as the *father of information theory*. He was working in Bell telephone laboratories now known as Nokia Bell laboratory. (Wetzel, 2023). In a Dartmouth seminar, Allen Newell, Herbert Simon, and Cliff Shaw presented their first versions of a computer program using AI. In that time, the program was seen as one of those earlier invented chess-playing programs. Later in 1956, the program was named *Logic Theorist* and it managed to validate mathematical theorems from Whitehead's and Russell's publication called *Principia Mathematica*. (McCorduck, 2003, p.123-124)

In 1964 Daniel Bobrow was a student at MIT and for his PhD dissertation, he created an AI program using *LISP*, an abbreviation of *list processing*. The program solved verbal tasks of algebra and it was called *STUDENT*. Daniel's supervisor at MIT was Marvin Minsky, one of the ideators of the Dartmouth conference. The basic idea, or the description, of *STUDENT*, is to cut down the verbal tasks into a smaller and smaller piece until the task is small enough for a computer to be able to calculate it. As a result of the verbal algebra task, the *STUDENT* gives a numeric answer. (Norvig, 2014)

In 1965 Joseph Weizenbaum developed an interactive program called *ELIZA*. *ELIZA* was able to have a written conversation with a human using natural language, English. The program was named after a character from a play called *Pygmalion*. In this play a professor

of phonetics places a bet, that he can teach a flower-selling street vendor, Eliza Doolittle, to act as a duchess at a garden party (Britannica, n.d.). *ELIZA* was programmed more likely to react to the input that human gives, instead of giving new information. *ELIZA* was a therapeutic chatbot, who tried to help the human to solve their problems by themselves instead of providing direct answers. Weizenbaum developed a new programming language, that he used for *ELIZA*. It was called *SLIP* (Symmetric List Processor), a list-processing programming language. The idea of the program was, that it searched a keyword from human input text and reacted to it in a predetermined manner. (Norvig, 2014)

One year later in 1966, Charles Rosen and his group from Stanford Research Institute developed the first moving robot called *Shakey*. It had a camera, a range finder, and a bump detector. *Shakey* was able to find a route from point A to point B in its rehearsal room and it could move small objects, that were on its way. The robot was named *Shakey* because it shook while moving. The whole development of the robot was done between 1966 and 1972 and in 1970, it was cited as *the first electrical person* by Life magazine. In 1983 *Shakey* was donated to Boston's Computer Museum. (Keay, 2017)

In 1973, US Congress strongly criticized AI and leading mathematician Professor Sir James Lighthill delivered a highly negative assessment of AI's progress in the UK. He argued that machines would never surpass the skill level of an experienced amateur in chess and would always struggle with tasks, that require common sense reasoning or facial recognition. This led to significant funding cuts for AI research, resulting in a period known as the AI winter. (BBC, n.d.)

Between 1970 and 1973, The *WABOT-1* was the world's first full-scale anthropomorphic robot. It featured a limb-control system, a vision system, and a conversation system. The *WABOT-1* could communicate with people in Japanese and measure distances and directions to objects using its artificial ears, eyes, and mouth. It could walk using its lower limbs and grasp and transport objects with its tactile-sensor-equipped hands. The mental capabilities of the *WABOT-1* were estimated to be similar to a one-and-a-half-year-old child. Its components included the WAM-4 artificial hands and the WL-5 artificial legs. *Wabot-2* was built in 1980. It was capable of playing keyboard instruments and reacting when a human played musical instruments. (Waseda University, n.d.)

In 1987 Rollo Carpenter invented a chatbot called *Jabberwacky*. It has ability to learn and entertain. It stores the data, that users are giving to it and uses that data, to perform

answers. Since it uses only the user input data, it is not suitable for business use, because no one qualifies the quality of those inputs. (Carpenter, n.d.)

Almost 10 years after in 1995, Richard Wallace developed a chatbot called *A.L.I.C.E.* (Artificial Linguistic Internet Computer Entity). It was named after a computer that was originally used to run it. *ALICE* was rewritten in Java in 1998. With *ALICE* users can chat either as they would chat with a real person. *ALICE* has never passed the Turing test. (Rouse, 2023)

The mid-90s was the time for robotic toys. In 1997 Dave Hampton and Caleb Chung invented a toy called Furby. It used a variation of the 6502 microprocessor, eight-bit chip, which also powered the Apple II and Commodore 64. Furby was able to use voice and had its language called Furbish. The first versions were able to use 42 words. Furby was able to learn the real language. It was also programmed to say "Me scared" after it had been over 30s. upside down. (Marsh, 2019)

A few years later, Sony presented a robot dog called AIBO (Artificial Intelligence Robot). It was a limited edition and it cost USD\$2500. AIBO was able to walk, use its mouth, tail, and head. It had an eight MB memory stick, which was used to insert software. AIBO had a touch sensor, a color CCD camera, a range finder, LED lamps, a mini stereo microphone, a miniature speaker, an acceleration sensor, and an angular velocity sensor. AIBO was able to recognize users' emotions and it was equipped with adaptive learning, which made it possible for it to develop its own personality. (SONY, 1999)

In 2002 i-Robot published Roomba. It was the first automatic home floor cleaning device available in the USA. When the Roomba begins cleaning, it moves in a spiral pattern until its bumper detects an object or wall. It then follows the wall using its Wall Following Sensor, cleaning edges with its Spinning side brush and moving to a new areas of the room. The Roomba alternates between creating new spirals and making straight runs, using the edge-cleaning routine to navigate. Cliff avoidance sensors prevent it from falling down stairs. (iRobot, 2002)

Apple published Siri in 2011. Siri is a virtual assistant that uses Apple iOS operating system and it came with iPhone 4S. Siri is able to send text messages, make calls and schedule reminders. It can learn about the interests of the one who is using it. Originally it used only English but after a year, it was able to understand Mandarin Chinese, Japanese and

Russian. (Sanchez, 2023). In 2014 Microsoft published Cortana, which is their version of Siri, but to be used with Windows and then Amazon published Alexa (Warren, 2014).

In 2020 Open AI released the GPT-3 LLM model (Large language model). A large amount of internet datasets were used to train it. Its purpose was to be able to answer to questions using natural language. A year later Open AI published DALL-E, which is capable of creating images from a user input. (O'Neil, 2023)

3.2 Benefits of using AI

Like any other application, system, technology, or technique has its advantages and disadvantages, so does AI. TU Delft OPEN Publishing published in 2023 Juliana E. Goncalves's book called *Artificial Intelligence, Real Consequences*. In the publication, Goncalves describes modern-day AI application opportunities and the risks and disadvantages it has.

One of the benefits of AI is, that AI can help humans make decisions. Since AI can offer data and suggestions, that are based on information instead of emotion, it can help make reasonable decisions. One example of this, is different streaming services, such as Netflix and Spotify which have ability to give to the user suggestions based on their user history. AI which is helping make decisions, is called Intelligent Decision Support Systems (IDSS). Also, Google Maps is such a system, because it can give the application user suggestions of the best or the fastest possible routes to their destination. As another example, IDSS can help the government by providing forecasts based on the changes in criminality. (Goncalves, 2023, p.19)

The Covid-19 pandemic forced people to come up with solutions to ensure a functional society and services such as education. AI can offer easier access to information, training and services. All this increased the amount of the world's data production and it is expected, that the amount will keep growing from 33 zettabytes to 175 zettabytes in the year 2025. One zettabyte is one thousand billion gigabytes. (European Parliament, 2023b)

AI can assist in the work environment. Robots can help by performing tasks that are dull, repetitive or dangerous (A3, 2019). This might also make some professions to disappear, but robots are not allowed to operate themselves, so for example technicians, engineers and developers are still needed. (ResumeMent, 2023).

In the healthcare sector, IDSS can suggest treatments or medication based on the patient's history. AI can also be used to analyze X-ray images. In 2018 Stanford Medicine released their study, about an AI algorithm that can evaluate X-ray images and give analysis based on the evaluation. The algorithm is trained to detect 14 pathologies. The algorithm can recognize 10 of those diseases like a real professional, the algorithm did not recognize three of the diseases and for one of those diseases, the algorithm was able to recognize better than the professionals. The algorithm was trained by using 112,000 different X-ray images. At the same time, three radiologists checked 420 X-rays and categorized those into 14 pathologies, which formed the final test for the algorithm. (Armitage, 2018)

In the form of robots, AI can assist in medical work. These so-called carebots can, for example, help serving food or medicines in nursing homes. Carebots can monitor patients' vital functions, remind them to take medicines and assist and help nurses when they need to lift patients and keep them company. Using robots in medical care also protects nursing staff being abused by patients who, for example, have dementia and can be unpredictable. (Goncalves, 2023, p.20)

Natural disasters are becoming more common with climate change and forming a weather forecast can be challenging. AI can help in analyzing environmental data and create possible outputs and scenarios, that can be difficult for humans to predict. AI can also help to maintain biodiversity by analyzing the disappearance of wild species. WildTrack is an AI-based system, that tracks footprints from nature. AI is also used to observe animal populations. AI can assist in analyzing satellite images and monitor the status of deforestation. (Goncalves, 2023, p.21)

AI can help predict environmental issues, but it can also help predict possible power cuts, energy consumption peaks and electricity prices. AI can be used to optimize office and residential buildings' energy consumption. Energy consumption is also one of the largest reasons for carbon emissions being high. It is estimated that AI can help to reduce 1.5-4% of global greenhouse emissions by 2030. (European Parliament, 2023b) Natural energy sources are increasing its popularity and AI can have its role in boosting renewable energy being used even more. With the help of AI, it is easier to predict solar and wind energy intakes optimize and coordinate grid operation and solar panels to prevent power cuts. The current world situation is not as desirable as possible and war actions are ongoing. One of the attack targets are power plants since countries are dependent on energy sources. Cyber-attacks to power plants can cause a lot of harm to many people and business. Preventing those attacks and detect faults in grid, is something where AI could assist more in the future.

Current research show that it possible, but still more development is needed. (Goncalves, 2023, p.22)

In the field of transportation AI can assist in cost reduction by optimizing and planning routes for waste management for example. Autonomous vehicles usually use voice recognition and image recognition as a part of the autopilot function. Different kind of sensors are needed and those produce data for AI to analyse. Besides of route planning, AI can also observe traffics, so called bottlenecks, to predict possible traffic accidents or nature's effects on traffic flow. On Singapore's airport AI is helping to predict flights arrival times and, in that way, help handling people. Those predictions have been almost 95% accurate. (Goncalves, 2023, p.24)

Clean water is a condition for life. United Nations has researched that between 2015 and 2020 the number of populations that have access to safe drinking water increases from 70% to 74%. It is predicted that the need for clean drinking water is increasing and one of the reasons is climate change, which is forcing people to move to another location. With the help of AI, water consumption can be observed and optimized. AI can detect possible leaks from sensor sent data, analyse water quality and soil situation so, that water consumption in agriculture can be adjusted, for example. (Goncalves, 2023, p.25)

In agriculture, different kinds of sensors and satellite images can be used to observe pesticides, weeds, weather conditions and nutrient levels. AI can analyse that data as well. That decreases the need to go into the field which lowers the carbon emissions. (Goncalves, 2023, p.26)

3.3 Disadvantages of using AI

AI has so many good qualities, but it has also disadvantages. Quality of AI depends on the data that is used in its training and the person executing the training. What goes in, comes out. One of the uses for AI, is different voice control systems such as Siri and Alexa. (Goncalves, 2023, p.32) Human gives voice commands to those systems and usually user needs to either speak loudly or repeat what was said. Alexa and Siri are both using female voice and that is most likely on purpose. Studies show that female voices are much nicer to listen to than male voices. This is not the problematic part, but usually user is not polite when commanding or talking to Alexa, Siri or some other voice-operated systems. When combining rude behaviour towards female voice, it makes it problematic. This can lead into a situation where act towards certain gender stereotypes is reinforced. (Kudina, n.d.)

Data used to teach AI systems, may only contain material of certain genders, accents, and races. That can cause situations, where AI voice assistant is not recognizing what a human is saying. Having only voice data from certain aged people can be problematic, since voice assistant recognize speech based on the data that is used in its training. (Goncalves, 2023 p.32) Studies have shown, that people who are bilingual, have more experience on voice assistants being unable to recognize their voice. This can be seen also when comparing voice to speech accuracy between different languages. This means, that the audio material used for teaching AI is too limited. (Tecnologyorg, 2021)

AI operated voice assistant have the ability to listen, so they know what they are asked to do. This can cause data privacy issues because those systems can also listen even when no one is talking to them. One of the Google features is, that they listen users via phone so that they can give you specified ads. This data is saved to their servers and used in voice assistant teaching purposes. ChatGPT also saves the user's data input. That data could be used for training models. User can also disable chat history. That raises a question: even if the data is protected in servers, is it protected sufficiently from cyber-attacks, for example? AI systems are not so transparent and it is a bit mystery how AI performs outputs and what happens behind it. EU data protection regulation (GDPR) is made to protect personal data being given or sold to third parties that user is not aware of. GDPR gives users the right to request information about the data, that companies are collecting from them individually. (Goncalves, 2023, p. 34-38)

AI can be used to specify ads based on users behaviour in internet. This works in other way also. If user is consuming informal data, such as fake news and conspiracy, AI will most likely promote similar material based on the user's previous behaviour. That is harmful, because the user only sees fake information instead of trustworthy information. AI can be used to create fake news. New AI tools make it possible to replace voice, generate images and combine audio to videos. This offers good possibilities for those, who want to cause harm to others. This can also lead into a situation, in which actors' are replaced with AI and/or actors' voices is being used. Use of AI can also make some professions in movie business to disappear. (Goncalves, 2023, p. 34-38)

Environmental impact of using AI can also be enormous. AI can help others to lower their energy costs, but using AI itself, consumes a great amount of energy. Computers cannot be used without electricity and since the internet and AI never sleep, those are continuously using electricity. (Goncalves, 2023, p.38) International Energy Agency's (IEA) study predicts, that datacentres total electricity consumption might be over 1.000 terawatt (TWh) hours in

2026. One TWh is 1000 000 000 kilowatt hours (kWh) (Buckley, 2024). For comparison, a Finnish detached house consumes around 17,000 – 20,000 kWh in a year and about 1400 – 1700 kWh in a month (Sähkövertailu.fi, n.d.). Datacentres also need cooling systems, which are also consuming electricity. Chat GPT consumes electricity 0.0017 -0.0026 kWh per query and ChatGPT receives around 10 000 000 queries per day ((Ludvigsen, 2023; Lammertyn, 2024). That means, it is consuming 17 000 – 26 000 kWh in a day. Chat GPT also uses 500ml water during every five to 50 prompts, because supercomputers need to be cooled (Singh, 2023). A prompt is an instruction, which a user inserts to the AI to get answers. Training AI also consumes electricity. It is estimated, that ChatGPT 3 electricity consumption emitted over 500 metric tons of carbon dioxide (L., 2023). If computers building materials and those carbon emissions are added to that electricity consumption, then we can think is AI digitality as green as we have thought.

3.4 Legislation

One of the European Union's focuses is digitalization. European Parliament is helping to create and modify policies that are making the EU strong in new digital technologies. Supporting peoples digital skills, employees digital know-how and helping to digitalize public services are those, what this policy is trying to achieve. The EU wants to make sure that fundamental rights are considered and respected. Digitalization means that digital technology, for example digital communication such as emails, online stores and electronic appointment booking, are integrated as apart of public and private sector services. Digitalization is also affecting to a society by these digital technologies. (European Parliament, 2021a)

The EU is aiming to be digitally independent and have their own norms and regulations instead of following someone else's. European Commission has presented a digitalization related political program that includes targets for the year 2030. In May 2021, the EU parliament accepted a report on building digital future of Europe. In that report, commission of the EU was suggested to take advantage of internal digital markets, improve usage of AI and support digital innovation and competence. (European Parliament, 2021a)

Members of the EU parliament have highlighted that human-oriented legislation for AI is needed, to be able to take all the possibilities, that AI offers, into use (European Parliament, 2021a). In April in 2021, the EU suggested that AI systems are categorized based on the threads those are causing for the one how is using them (European Commission, 2021). That legislation, called AI Act, has been approved on 13th of March in 2024 and that is the world's

first legislation on AI. 534 members of the European parliament voted for it, 46 members voted against it and 49 votes were empty (European Parliament, 2024d).

This regulation prohibits those AI applications that are seen as a threat for citizen rights. For example, biometric classification systems and those that are collecting data using facial recognition. Applications that are giving social points for humans based on, for example their salary, education or social status or manipulates people's behaviour are also prohibited. There are few exceptions when biometric recognition can be used. One of those exceptions is, to use targeted search on finding missing person or prevent terrorist attacks. In these exceptional cases, court permission is required and criminal activity must be involved before permission can be granted. (European Parliament, 2023c)

Those AI systems that have negative impact on safety and/or fundamental rights, are categorized as a high risk. That category is also divided into two subcategories. First group includes those AI systems that are used in the products that belong under EUs production safety directive, such as toys, aviation, cars, medical equipment, and elevators. Second group includes those AI systems that needs to be registered into a EUs database, such as control and handling of critical infrastructure, general- and vocational education, law enforcement and border control. All AI systems that have high risk, needs to be evaluated before those can be released to the market and those systems needs to have regular checks during their life cycle. Citizens can also make complaints when needed. (European Parliament, 2023c)

General AI systems are not categorized as a high-risk system, but those need to follow certain regulations. For example, ChatGPT needs to tell its user that the content is made by AI. Those kind of systems needs to be developed so that creating illegal content is prevented. The system needs to publish summaries of copyright protected data that is used for educational purposes. Content, such as images and sounds that AI has created, needs to have clear marking that AI has created them. In that way users know when they are using data that AI has made. (European Parliament, 2023c)

EU AI Act has penalties for those who are breaking the rules. Using prohibited AI systems can lead to fines of up to 40 million euros or 7% of worldwide annual turnover depending on which one is higher. If the AI is not following the transparency regulation, then the fine can be up to 20 million euros or 4% of worldwide turnover which one ever is higher. If the AI system does not comply with other requirements, then the fine can be up to 10 million euros or 2% of worldwide turnover. (EU Act, 2024)

3.5 AI in software testing

AI can be helpful in software testing, by performing those repetitive tasks and assisting in automating testing processes. AI algorithms can analyse application requirements, specifications, and historical data to generate test cases automatically. By identifying relevant test scenarios, edge cases, and potential risks, AI-driven test case generation can improve test coverage and reduce manual work. (Onix, 2023)

AI techniques, such as machine learning, can analyse code metrics, defect histories and other factors to predict potential defects or areas of the code, where bugs might be found. By identifying high-risk areas early in the development cycle, AI-driven defect prediction can help prioritize testing efforts and allocate resources effectively. (Onix, 2023) Analysing test results and changes in the code, test execution can be optimized. By focusing on critical test cases and application areas that has high impact, AI-driven test prioritization improves testing efficiency and reduces time consumption. (Onix, 2023)

Natural language processing (NLP) is another key area. It is developed to understand and generate human-like text, which helps in automating the creation of test scripts from requirements and bug reports written in natural language. (Jacinth, 2023) Search-based software engineering (SBSE) uses advanced meta-heuristic search techniques, such as genetic algorithms and simulated annealing, to improve different parts of software testing. SBSE engineering has been used in software testing tasks such as automatically generating test cases, minimizing them and prioritizing them. It has also been applied to regression testing to some extent. (Atlanti International University, n.d.)

Model-based testing (MBT) uses models to represent the desired behavior of the system under test. AI enhances MBT by automating the generation and execution of test cases from these models. Formal methods provide rigorous mathematical frameworks for specifying, developing and verifying software systems. AI assists in automating the application of these methods in testing. (Singla, 2024)

Data mining techniques are used to analyze historical test data and software logs to uncover patterns and insights that inform test planning and defect prediction. Cognitive computing involves simulating human thought processes in a computerized model, used to understand and reason about the testing processes, thus aiding in the intelligent automation of testing activities such as test data generation, test coverage analysis and defect prediction. (Ruwan, 2023)

3.6 AI tools for software testing

There are software testing tools that utilize AI in their operations. Testsigma categorizes the types of those testing tools into categories, which are tools for test case management, test case generation, preparing test data, test execution, coverage measurement, performance testing, bug reposting, project planning and tracking and log analysis (Testsigma, n.d-a).

When choosing best suitable AI testing tools for the company, it is important to define what tool should do understand what the need is and does infrastructure have limitations. Evaluating tools functionalities and scalability is also needed. (VALA, 2023a)

3.6.1 Testim

Tricentis is a software testing company, which was founded in 2007 and its headquarter is in Austin, Texas (Tricentis, n.d.). Tricentis offers products for test automation, test management, performance testing, mobile testing and data and change intelligence.

Testim is a no-code AI powered platform from Tricentis, used to end-to-end testing of mobile- and web applications. It uses AI (NLP) and machine learning to identify changes in the application, diagnosing failed tests and find root causes. Testim is a cloud-based software, so it is always updated, and it can be used with multiple browsers such as Safari and Chrome. Although Testim is codeless, user can still add JavaScript to test when needed. It also has integration to Jira. (Testim, n.d.-b)

From April 2024 Tricentis added Azure OpenAI Copilot to be part of their AI solutions. Tricentis is using GPT 3.5 Turbo AI model. Testim Copilot offers live support, it can identify pixels and simplify debugging. Copilot also explains existing code and creates the code and test steps using generative AI. (Brooks, 2024)

Pricing is based on the ordering company's needs and size. Tricentis does not give any exact information about their prices, so the user needs to contact them directly. There is free 14 days trial version available. After that trial is ended, user can continue using Testim with Community free plan which means that monthly web runs can be executed but limit is one account per organization. (Testim, n.d.-a)

3.6.2 Applitools

Applitools is also a low/no code solution, which runs in cloud, behind a firewall or as a public SaaS. It can be integrated to other platforms, such as Robot Framework, Selenium or Cypress, or AI (NLP) could be used to create tests. Applitools has scalable architecture. Tests can be created with a recorder, which makes it possible to add end-to-end tests without coding. Tests can also be created by using natural language. (Applitools, n.d.-b)

Applitools offer also free 14 -day trial for test automation and during that time user can get automatically created test cases using English, visual AI-scanned user interfaces and self-healing test runs. With applitools eyes, user can test webpages, applications or documents by using visual AI. User can sign up for a free account. Applitools preflight offers self-healing locators, which means that if locator is missing or changed, preflight can correct it in the test by itself. Visual AI crates images and charts on UI, testing can be scheduled, and it can also test your emails so that those have everything needed. Preflight can also test PDFs and it offers a free trial. (Applitools, n.d.-a; Applitools, n.d.-c)

3.6.3 Testsigma

Testsigma offers low-code test automation using English. Users can plan testing, create, execute, debug and report testing. Web, mobile, desktop and API testing are included. It has over 3000 mobile devices and browsers that user can use for mobile and web applications. Testsigma makes it possible to have self-healing tests and it uses NLP. (Testsigma, n.d.-c)

There are two plans available, pro and enterprise which are displayed in Figure 2. Enterprise plan has the same services than Pro and few extra services such as desktop automation. On AI point of view, both versions have an auto-healing function that fixes test steps, automatically creates regression tests, plans for failed test cases and provides correction suggestions for failed tests. (Testsigma, n.d.-c)

Figure 2 Testsigma plans (Testsigma, n.d.-b).

The image shows two pricing plans for Testsigma. The 'Pro' plan is highlighted as 'MOST POPULAR' and is for 'Fast-Growing Teams'. The 'Enterprise' plan is for 'High-Scale Teams' and includes everything in the Pro plan plus additional features.

Plan	Target Audience	Key Features
Pro	For Fast-Growing Teams	<ul style="list-style-type: none"> Supports testing Web, Mobile Web, iOS Apps, Android Apps, REST APIs Unlimited applications and projects Unlimited automated testing minutes 800+ browsers/OS combinations 2000+ real mobile devices Parallel execution 30+ integrations Auto-healing scripts 5 GB Cloud Storage per Parallel 24/5 support
Enterprise	For High-Scale Teams	<ul style="list-style-type: none"> Everything in Pro, Plus SAML 2.0-based Single sign-on (SSO) Public, private or on-prem cloud deployment Desktop apps automation IP Whitelisting & Local Testing 24/7 support

Both plans include buttons for 'Request Pricing' and 'Sign up for Free'.

3.6.4 Aqua

Aqua is a product of a company called Andagon. Andagon offers software testing, test automation and software quality consulting-related services. Its main office is in Köln; Germany and the company was founded in 2001. (Aqua, n.d.-a)

Aqua is a tool for test management and it is designed for developers, testers and management. With Aqua, user can create test steps by using plain words or using Aqua copilot and let the tool create testcases. Andagon says that 42% of those copilots generated testcases, do not need adjustments or fine tuning. If something needs to change, for example in test cases, users can simply type how they want the documentation to be changed. No prompting is needed. Test cases can also be prioritized automatically. Aqua shows the requirement coverage and user can create detailed and shareable overviews. Aqua has over 12 third-party integrations, such as Jira and Jenkins, included. It has user friendly interface and that helps if manual testing and test automation management is combined in Aqua.

Pricing depends on the size of the user's company. Aqua offers flexible monthly, annual, and multi-year subscriptions. (Aqua, n.d.-b)

4 Use of AI in software testing at Etteplan Finland Oy

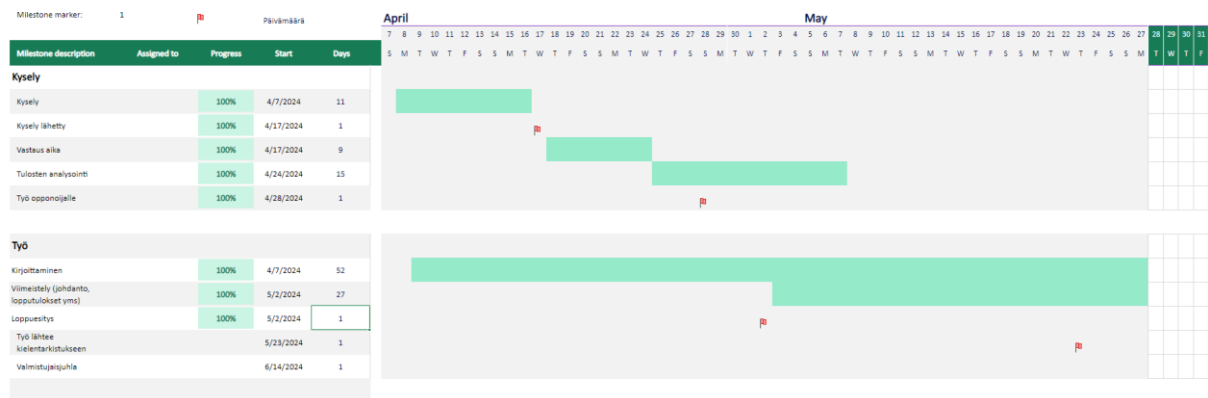
The commissioner of the thesis, Etteplan, is an engineering company founded in 1983 by Ensio Juotasniemi, Tero Elomaa, Tapani Mönkkönen and Esko Poltto. The company name ETTE is formed by the initials of those founders first names. Etteplan has over 4000 employees and it has offices in Sweden, the Netherlands, Germany, Poland, Denmark and China in addition to Finland. One sales office is also located in the USA. The company's headquarter is in Espoo, Finland. Juha Närkki is currently working as a CEO of the company and he started in that position in 2012. The company offers services in mechanical engineering, software and embedded services, technical communication, testing and green transition. Etteplan has created solutions for example in aerospace and defence, energy and power transmissions, medical technology, ICT, and high tech and offshore and marine. The company's customers include example ABB, Valmet, Philips and KONE.

Chapter 4 reviews the schedule, execution, and the results of this thesis. The results of chapter four are analysed in more detailed in chapter five.

4.1 Schedule of the thesis

The topic of the thesis was approved on 12 February 2024 and after that, a kick-off meeting was held with the supervising teacher. As the Figure 3 display, the thesis background research started after the kick-off meeting and the writing begun at the turn of March and April. The plan was to get the survey ready during week 15 so that it could be sent to the respondents the next week. A week was given to answer and after that the answers were analysed. The thesis was sent to the opponent before the final seminar, which was held on 2 May.

Figure 3 Schedule of thesis.



The schedule of the thesis was tight, and the thesis was done entirely alongside the author's own daily work. During the thesis process, weekly meetings were held with the commissioner of the thesis. The progress of the thesis and the keeping of the schedule were monitored in those meetings. If necessary, changes were made to the schedule.

4.2 Survey

During the thesis process a survey was executed to find out, if software testers at Etteplan are currently using AI on their daily work and are those tools offered by their customers or are those tools used purely by Etteplan's software testers' decision. Etteplan has software testers in Finland, Poland, and Sweden. The survey was anonymous to avoid the spread of possible company secrets. Since there are over 90 software testers in the company and the thesis was carried out alongside the author's own work, survey was selected instead of interviews. It was also expected that the survey will collect responses that are more in line with the truth than the answers obtained during the interview. Surveys also offer the respondents freedom to select the best suitable time to response. Anonymous responses were also possible with the survey.

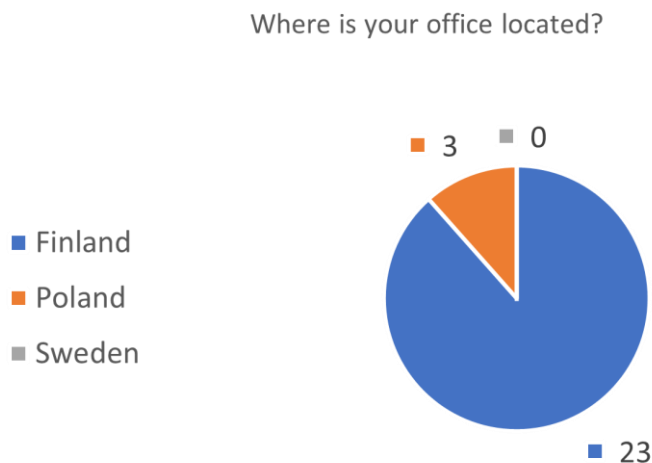
The question number one forced the respondent to select their office location. The location was asked so that, if necessary, the differences in those responses can be examined country by country. Questions number two and three were based on the respondent's own experience or feeling of the subject, rather than how things are. The response scale was from one to six and one was "not at all" and six was "very much". The question number 13 has also the response scale from one to six and respondents could select the value, that described their general about using AI. One was "not happy or do not like at all" and six was "very enthusiastic and curious". Since questions two, three and 13 were based on a feeling,

it was not seen as necessary to open that scale further. The survey questions can be found at the end of this thesis in Appendix 1/1.

4.3 Results

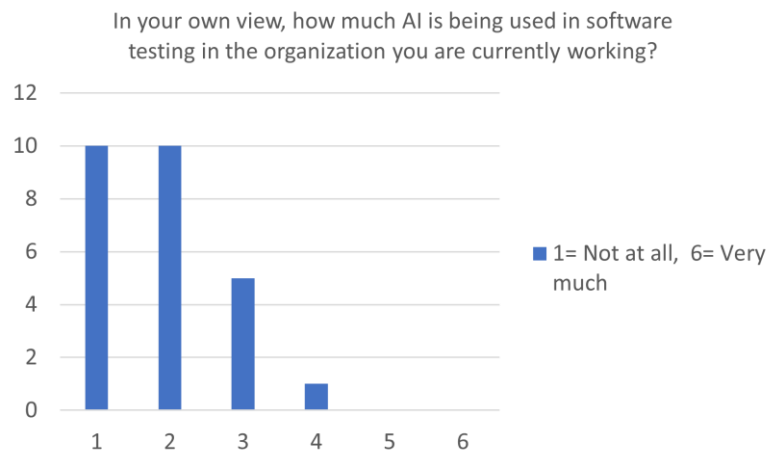
The survey response time was one week after the survey was sent and, in that time, 26 responds were collected. 23 of those responses came from Finland, 3 from Poland and none of the respondents were from Sweden as presented in Figure 4. In total, there are 63 recipients in the software testing-related email list in Finland and 22 in the Poland email list. Based on the total number of people in those lists, Finland's response rate was 36.51% Poland's 13.64%.

Figure 4 Respondents office locations.



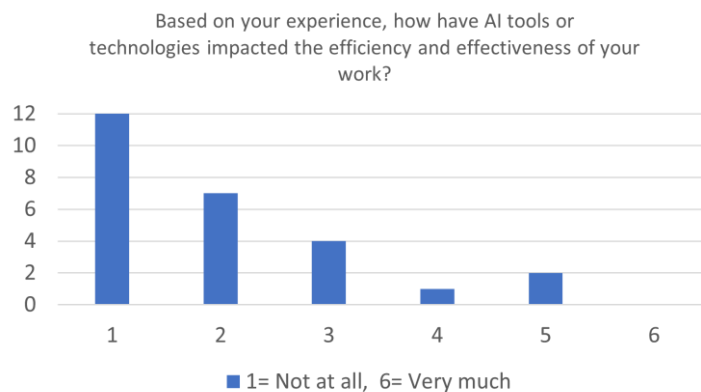
The second question was about respondents own experience on how much AI is being used in the software testing in that company where they are currently working. As displayed in Figure 5, the response scale was from one being not at all to six being very much. Half of the 20 respondents selected value one and the other half selected value two. Five respondents selected value three and only one had selected value four. The average of all responses was 1.88. Based on this, less than half of the respondents are using AI in their work.

Figure 5 How much AI is being used in software testing.



In the third question, the respondent had to estimate how much AI had impacted the efficiency and effectiveness on their work. The response rate was same, then in the previous question, from one to six. Most of the respondents had selected value one and seven had selected value two, which was also the average value of all responses as displayed in Figure 6. The average of all responses was 2.00. This means, that respondents did feel that AI has improved their work efficiency only little.

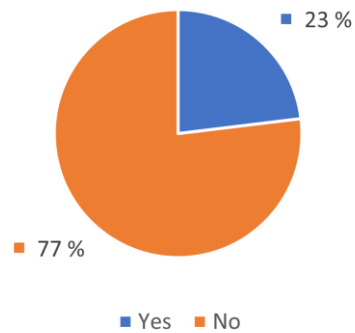
Figure 6 Efficiency impact of AI tools.



20 of the respondents are not currently using any AI or tool in their work as a software tester. Based on the responses, 20 (77%) respondents answered “no” to this question and 90% of those respondents, selected value one or two in question number two. 6 (23%) respondents who answered “yes” to this question are from Finland. (Figure 7) None of the Polish respondents are using any AI tools in their work.

Figure 7 Number of respondents using AI in testing.

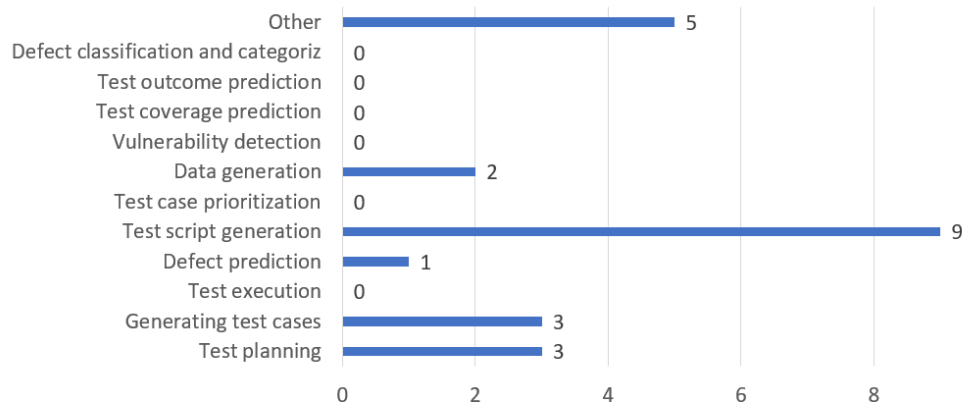
Are you currently using any AI technologies or tools in your work?



Nine respondents are using artificial intelligence in test script generation as displayed in Figure 8. Other use targets were testing planning, generating test cases, data generation and one of the respondents is using it to defect prediction. 11 respondents did not answer to this and five respondents who answered “other” are not using AI to anything currently.

Figure 8 How AI is used in testing.

How are you, or the customer you are working for, using AI tools related to software testing?

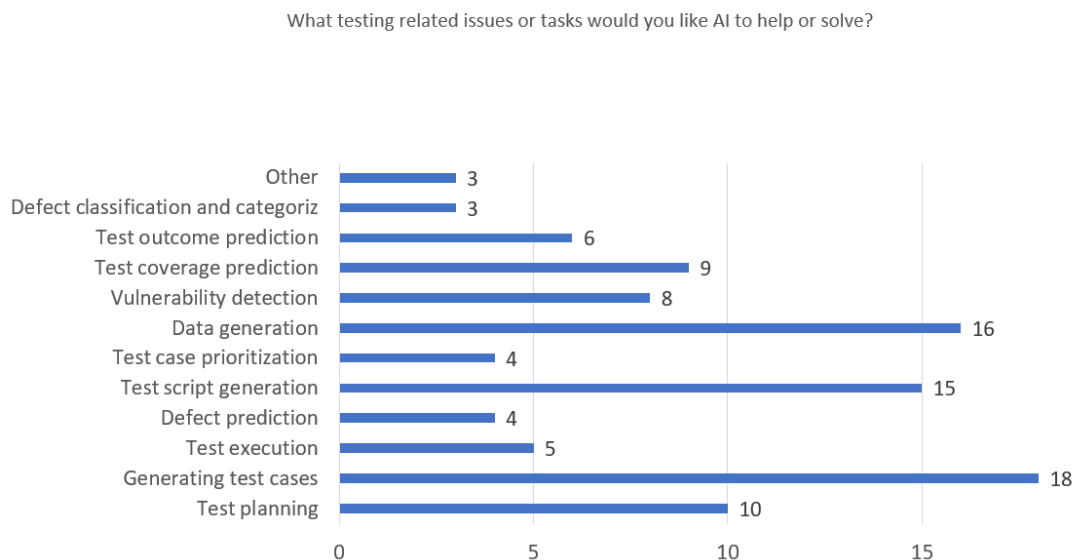


Tools and technologies, which are mainly used, are Microsoft Copilot and OpenAI's ChatGPT. Also, Tricentis and Daisee were mentioned in the list. 19 out of 26 respondents answered to this question and 31% of the respondents answered Copilot. Five respondents answered that they are not using any AI tools or technologies.

Question number seven considered respondents' thoughts about taking any AI tool to use and what tools they might start using. 16 respondents answered this question and two of those answered Copilot. In most of the answers were said that they are not thinking taking any AI tools into a use because they do not know what that tool could be, but they also said that they are interested in using AI tools.

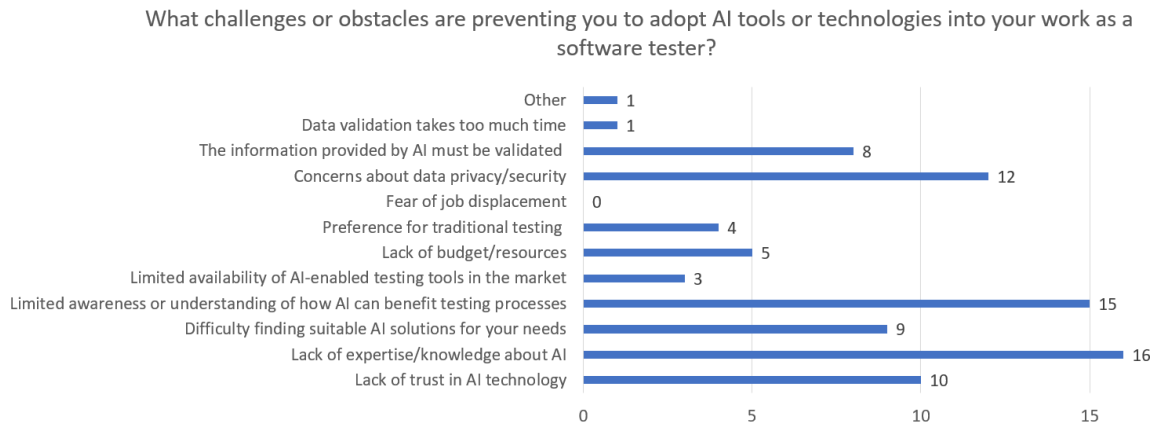
25 respondents answered question eight and 18 of them said that test case generation is the issue that they wish AI could help with or solve totally as displayed in Figure 9. Secondly data generation and on third test script generation. Defect classification, vulnerability detection and test coverage prediction got also around 10 selections.

Figure 9 Respondents wished AI to help or solve.



24 answers were collected to question nine. As displayed in Figure 10, 16 respondents said that lack of expertise or knowledge is preventing them from taking AI tools into a use in their work. 15 said that they do not understand how AI could help in their work and 12 had concerns about data privacy or security. 10 of the respondents have a lack of trust in AI and 9 are having trouble finding suitable AI solutions for their needs.

Figure 10 Reasons why AI is not being used in testing.



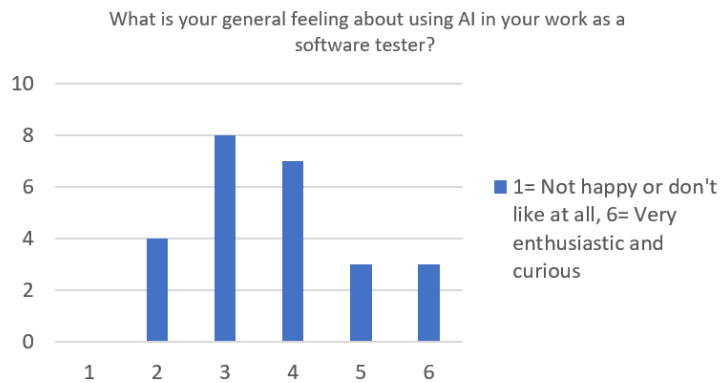
The question number 10 got only few answers. One of the respondents said that creating test cases can be easy, but when it is needed to make changes to that AI created test case, that can be sometimes difficult to do. Also, when AI has been used, user must also fine-tune those outcomes and that takes time. Time consumption was mentioned in a few other answers as well. It takes time to learn how to use new tools or take those into a use.

in the question 11 the respondent was asked to tell, if there are any lessons learned or recommendations that they would like to share with a colleague who is interested in adopting AI in their work. From the answers, a recommendation emerged to check, what answers the AI offers before using them and that you should not be too dependent on AI tools, because of the risks of having the wrong outcome.

In the answers to the question 12, more training related to AI was desired. Respondents also seemed to have an understanding that AI is focused more on in software development, than in software testing and AI tools can be used, but they are not seen as reliable. In total this question got 12 answers.

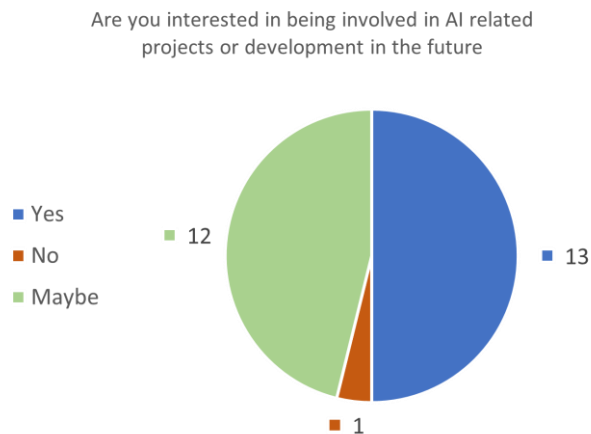
The question 13 was about respondents general feeling about using AI in their work and 24% of respondents gave their score between a value five and six as displayed in Figure 11. This question got 25 answers in total and the average of all responses was 3.72. Most of the respondents had selected values three and four for their answer. Based on average (3.72) respondents are really interested in using AI more.

Figure 11 Respondents' general feelings using AI.



In the last question, respondents were asked if they are interested in being involved in AI-related projects or development in the future. As displayed in Figure 12, 13 respondents had answered “yes” to this question. 12 respondents answered “maybe”. 92% of those respondents who selected “maybe” to this question, had selected “Finland” in question number one. One respondent selected “no” to this question.

Figure 12 Interest in using AI in their work in the future.



5 Conclusions, reflection, and future development

The response rate was only 36.51% in Finland and 13.64% in Poland. Less than half of the potential respondents answered the survey, so the responses that were collected, are not giving the exact truth of the situation. This needs to be taken into consideration, when

forming solutions or the status. Respondent did not have to log in to be able to give their answers and because of that, one person could have answered multiple times with different answers. If that happened, then the results would not be correct either.

Mostly AI is not used in the software testing at the moment. This can also mean that respondents do not know if they are using AI, even though they had a feeling that they are not using it. Many applications and tools have at least some AI integrated in them. For example, Robot Framework has libraries for image recognition and RPA. OpenAI -Library supports OpenAI, which makes it possible for the robot to access the Open AI Microsoft API. Since respondents did not either use, or recognized using, AI, they do not have the experience of AI being helping to improve the efficiency of the work either. The lack of information and knowledge related to AI may also cause it to be distrusted. On the other hand, it is a good feature that not all sensitive information is being shared over the internet. 12 respondents were concerned about data security and it prevents them from using AI.

The respondents did not have much AI-related experience, but most of them were at least a little interested in working with it or being involved in projects that included AI. Since AI is predicted to be increasingly used in the future, it is a positive feature that Etteplan software testers are interested of that subject. Only one of the respondents chose that they are not interested using AI in their work. The survey did not ask the respondents age, so this answer may have come from an older respondent, who is not interested in newer technology. No one had selected the fear of job displacement as the answer to question nine, so that might not be reason. If the Etteplan provides a clear list and instructions on which AI tools are allowed to be used and how those can be used, it might create a sense of security for the software testers knowing that they are using safe tools approved by the company.

Based on the survey results, there is a need for the AI-related basic training such as what AI is, how it works, how it is used and how to use it. It is also beneficial for the Etteplan's customers, when Etteplan software testers are familiar, at least with the basics, so when they start working in a customer projects, it is easier to adapt the customer's AI tools and/or help the customer to take AI into a use in their projects. Willingness to work with AI also support the interest of participating to the training. Providing also examples of how AI could be used in software testing, might courage to learn more.

Participating into a trainings, seminars and conferences that are related to AI, is also a good way to increase knowledge and get to learn from other experts, who may have more experience or at least different experience. Networking in such situations also makes the

company visible. Probably customer representatives also participate in these events. Presenting practical examples and case studies from other companies implementing AI can also help Etteplan software testers to add knowledge on how to tackle obstacles with the help of AI. From time to time, it may also be more cost-effective to properly train a larger group than to send one or two employees to a course and then ask them to tell others the most important main points. In a well-organized course with competent instructors, the lesson is better remembered, than by reading someone else's notes. Especially when different things can be important and/or interesting to different people.

Increasing automation also brings savings. The savings may be small at first because recreating processes takes time and possibly costs are also generated from new equipment and software. By calculating the total benefit, it can still be more efficient to automate repetitive work steps that do not require a human analyse. After automating as much as possible, more AI could be added to the system. Little by little as technology develops, AI could take more responsibility. AI is not suitable for work that requires human analysis of usability. When AI is helping to do the work it is capable of, then there is more time for human to execute UI tests and exploratory testing.

Etteplan is an engineering office and the company's software testers are working in customer projects or that is the typical situation. Therefore, company itself does not necessarily need to take any AI software testing tool into a use, but it would be beneficial for the company's software testers to know at least few of those tools and how to use them. This helps them to stay up to date and bring more value to the customer they are working for. Using AI tools can help optimizing time and resources which is a positive thing for the customer. The knowledge of AI is something that the customers are looking for and if some other company is more advanced in the field of AI, it will increase the risk of losing business to them.

AI chatbots for example could be used to generate test cases and test scripts. With the use of prompts, AI chatbot could give more specific answers. Adding information about prompting might also be helpful for software testers. ChatGPT 4 accepts also visual inputs. This means that users can feed the document to AI and give commands what they want AI to do with it. For example user could ask it to create testcases based on the document. User can also insert image of a test result and ask AI to create documentation or analysis based on it.

Current AI tools used in software testing are mainly for test automation and/ or operate best when the user is testing web-based applications. One of the development ideas for the future is to start creating an own AI model that is suitable also for manual testing and embedded

systems. Data security is also an area, which needs to be well instructed. Otherwise some or all company secrets may be leaked.

Everything cannot be automated and the AI created data needs to be validated. Testing AI itself is a new area in the field of software testing. Training software testers to also test AI brings competitive advantage.

As part of the thesis, the survey was implemented as planned before the start of the thesis. The thesis also brought up four software testing tools currently on the market that have AI functions included. The cost structure of these software's could not be determined, because it would have required sending a request for quotation. Prices are determined by the size of the organization.

The learning process during the thesis has been upward. As getting to know the subject more, I have learned more myself. The biggest insights during the project have been the events in AI's history. AI has seemed to be the top topic of the last few years, but it all started much earlier. Robotics also goes almost hand in hand with the AI and Robotics seems to be the spark for AI. There is no use in having robots that are not doing anything. There were also challenges in project scheduling. According to the original schedule, the thesis was supposed to be finished in the middle of May, but illnesses and the changing work situation also reformed the schedule. This taught me, or reminded me, how important it is to take the changes into account sufficiently in the schedule. It is also important to consider the time income of other people connected to the project in addition to your own. Making surveys is also a challenge. When preparing the questions, it is needed to carefully consider their layout and content. You must think about what you want to ask and what you want to find out with that question. I have not gained very extensive experience in preparing surveys myself but due to the thesis, more valuable experience was learned and gained. During the thesis, it was also necessary to communicate the direction of both, the person ordering the thesis and the supervising teacher. Strengthening communicational skills cannot be practiced too much, so communicational issues also had to be practiced during the project. It is needed to think carefully about what you are communicating, to whom you are communicating and at what time.

The use of AI has effects not only on the development of society and technology, but also on the environment. AI can help to reduce, for example, the carbon footprint and emissions of farming and forestry, when instead of being transported to the field, the crop and pest situation can be detected and analysed remotely. As in chapter 3.3. stated that AI also

requires electricity to keep the computers operational. Electricity consumption is one of the negative effects of AI on the environment. Chat GPT also consumes water during its operations, as the supercomputers and servers it uses, use water for cooling. The aim is to optimize electricity consumption, and development enables the use of computers that consume less electricity. The consumption of electricity to ensure the operation of computers and to guarantee cooling is still high. The use of AI and computers also creates electronic waste, which can be challenging to recycle. It often contains poisonous metals, chemicals and other harmful substances that ending up in nature cause problems. Especially in countries where recycling is still in its infancy, substances that get into nature also have an impact on people's health. In improving the recyclability of electronic waste, knowing the circular economy, and increasing it play a key role.

For the future development, it would be beneficial to retest the company's software testers AI know how level and compare results into those presented in this thesis. In that way could be seen, what kind of development has happened. One of the respondents said, that this survey took too little into account the negative aspects of AI and conveyed the impression that AI is only seen as a positive matter. If a similar survey is carried out in the future, it is good to pay attention to this as well and more clearly offer answer options that better takes this into account or instructions on how to answer. This could also increase the number of responses and leads to a better coverage.

References

- Applitoools. (n.d.-a). *Applitoools Preflight*. Applitoools.
<https://applitoools.com/platform/preflight/>
- Applitoools. (n.d.-b). *How it works*. Applitoools.
<https://applitoools.com/why-applitoools/>
- Applitoools. (n.d.-c). *Sign up for a free account today!* Applitoools.
<https://auth.applitoools.com/users/general-register?app=eyes>
- Aqua. (n.d.-a). *Aqua*. Aqua.
<https://aqua-cloud.io/about-us/>
- Aqua. (n.d.-b). *Speed up your releases x2 with QA management in aqua cloud*. Aqua
<https://aqua-cloud.io/product-tour/>
- Armitage, H. (2018, November 20). *Artificial intelligence rivals radiologists in screening X-rays for certain diseases*. Stanford Medicine
<https://med.stanford.edu/news/all-news/2018/11/ai-outperformed-radiologists-in-screening-x-rays-for-certain-diseases.html>
- Atlanti International University (n.d.). *Atlantic Internationa University*. Atlantic Internationa University.
<https://courses.aiu.edu/INFORMATION%20SYSTEMS%20ENGINEERING/SEC%204/SEC%204.pdf>
- A3. (2019). *How Robots Are Taking on the Dirty, Dangerous, and Dull Jobs*. Association for advancing automation.
<https://www.automate.org/robotics/blogs/how-robots-are-taking-on-the-dirty-dangerous-and-dull-jobs>
- BBC. (n.d.). *15 key moments in the story of artificial intelligence*. BBC.
<https://www.bbc.co.uk/teach/articles/zh77cqt>
- Britannica. (n.d.). *Pygmalion*. Britannica.
<https://www.britannica.com/topic/Pygmalion-play-by-Shaw>
- Brooks, S. (2024). *Tricentis launches testing Copilot*. Enterprise times.
<https://www.enterprisetimes.co.uk/2024/04/18/tricentis-launches-testing-copilot/>
- Buckley, S. (2024). *IEA Study Sees AI, Cryptocurrency Doubling Data Center Energy Consumption by 2026*. Data Center Frontier.
<https://www.datacenterfrontier.com/energy/article/33038469/iea-study-sees-ai-cryptocurrency-doubling-data-center-energy-consumption-by-2026>
- Carpenter, R. (n.d.). *Jabberwacky.com*.
<http://www.jabberwacky.com/j2about>

- China, C. R. (2024). *Breaking down the advantages and disadvantages of artificial intelligence*. IBM.
<https://www.ibm.com/blog/breaking-down-the-advantages-and-disadvantages-of-artificial-intelligence/>
- Cohn, M. (2009). *Succeeding with Agile: Software Development Using Scrum*. Addison-Wesley.
- Draniceau, Andreea (n.d). *AI in Software Testing*.
<https://theqalead.com/ai-ml/ai-in-software-testing/>
- EU Act (2024, March 11). *EU AI Act Compliance Checker*. EU Artificial Intelligence act.
<https://artificialintelligenceact.eu/assessment/eu-ai-act-compliance-checker/>
- European Commission(2021). *Proposal for a Regulation laying down harmonised rules on artificial intelligence*. European Commission.
<https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>
- European Parliament (2021a). *Shaping the digital transformation: EU strategy explained*. European Parliament.
<https://www.europarl.europa.eu/topics/en/article/20210414STO02010/shaping-the-digital-transformation-eu-strategy-explained>
- European Parliament (2023b). *Artificial intelligence: threats and opportunities*. European parliament.
<https://www.europarl.europa.eu/topics/en/article/20200918STO87404/artificial-intelligence-threats-and-opportunities>
- European Parliament (2023c). *EU AI Act: first regulation on artificial intelligence*. European Parliament.
<https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
- European Parliament (2024d). *Artificial Intelligence Act: MEPs adopt landmark law*. European Parliament.
<https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law>
- Gillis, A. (n.d.). *end-to-end testing*. TechTarget.
<https://www.techtarget.com/searchsoftwarequality/definition/End-to-end-testing>
- Goncalves, J. E. (2023). *Artificial Intelligence, Real Consequences:How AI is changing the way we live*. TU Delf OPEN.
- IBM. (n.d.-a). *The games that helped AI evolve*. IBM.
<https://www.ibm.com/history/early-games>

- IBM. (n.d.-b). *What is artificial intelligence(AI)?*. IBM.
<https://www.ibm.com/topics/artificial-intelligence>
- IBM. (2023). *Understanding the different types of artificial intelligence*. IBM.
<https://www.ibm.com/think/topics/artificial-intelligence-types>
- iRobot. (2002). *iRobot Introduces Roomba™ Intelligent FloorVac - The First Automatic Floor Cleaner In The U.S.* iRobot.
<https://media.irobot.com/2002-09-18-iRobot-Introduces-Roomba-Intelligent-FloorVac-The-First-Automatic-Floor-Cleaner-In-The-U-S>
- ISTQB. (2023). *Certified Tester: Foundation Level Syllabus v4.0*. ISTQB. https://istqb-main-web-prod.s3.amazonaws.com/media/documents/ISTQB_CTFL_Syllabus-v4.0.pdf
- Jacynth, P. (2023). *Natural Language Processing (NLP) in Software Testing: Automating Test Case Creation and Documentation*. DZone. <https://dzone.com/articles/natural-language-processing-nlp-in-software-testin>
- Jain, R. (2024,). *Minimax Algorithm with Alpha-beta pruning*. Hackerearth.
<https://www.hackerearth.com/blog/developers/minimax-algorithm-alpha-beta-pruning/>
- Kasurinen, J. P. (2013). *Ohjelmistotestauksen käsikirja*. Docendo.
- Keay, A. (2017). *SVR*. Shakey is first robot to receive IEEE Milestone award.
<https://svrobo.org/shakey-is-first-robot-to-recvie-ieee-milestone-award/>
- Kudina, O. (n.d.). *This is how digital voice assistants influence your life*. TU Delf Stories.
<https://www.tudelft.nl/en/stories/articles/this-is-how-digital-voice-assistants-influence-your-life>
- L., J. (2023). *How Big is the CO2 Footprint of AI Models? ChatGPT's Emissions*. carbon credits.com.
<https://carboncredits.com/how-big-is-the-co2-footprint-of-ai-models-chatgpts-emissions/>
- Lammertyn, M. (2024). *60+ ChatGPT Statistics And Facts You Need to Know in 2024*. Invagate.
<https://blog.invgate.com/chatgpt-statistics>
- Leloudas, P. (2023). *Introduction to Software Testing: A Practical Guide to Testing, Design, Automation, and Execution*. Apress.
- Ludvigsen, K. G. (2023). *ChatGPT's energy use per query*. Towards data Science.
<https://towardsdatascience.com/chatgpts-energy-use-per-query-9383b8654487>
- Marsh, A. (2019). *Coded for Cuteness: How the Furby Conquered Hearts and Minds*. IEEE Spectrum.
<https://spectrum.ieee.org/furby-toy>

McCorduck, P. (2003). *Machines Who Think*. A K Peters, Ltd.

https://monoskop.org/images/1/1e/McCorduck_Pamela_Machines_Who_Think_2nd_ed.pdf

Norvig, P. (2014). *Paradigms of Artificial Intelligence Programming*. Morgan Kaufmann.

O'Neil, S. (2023). *The History of OpenAI*.

<https://www.lxahub.com/stories/the-history-of-openai>

Onix. (2023). *The Role of Artificial Intelligence in Software Testing and Quality Assurance*. LinkedIn.

<https://www.linkedin.com/pulse/role-artificial-intelligence-software-testing-quality-assurance/>

Technology org (2021). *The accent gap problem in minorities and dialect speakers*.

Technology org.

<https://www.technology.org/2021/04/23/the-accent-gap-problem-in-minorities-and-dialect-speakers/>

Pittet, S. (n.d.). *The different types of software testing*. ATlassian.

<https://www.atlassian.com/continuous-delivery/software-testing/types-of-software-testing>

ResumeMent. (2023). *Artificial Intelligence and the Future of Jobs | Unlocking the Future: 20 Exciting Careers in Artificial Intelligence (AI)*. LinkedIn.

<https://www.linkedin.com/pulse/artificial-intelligence-future-jobs-unlocking-20-exciting-careers/>

Rouse, M. (2023). *ALICE (Artificial Linguistic Computer Entity)*. Technopedia.

<https://www.techopedia.com/definition/380/artificial-linguistic-computer-entity-alice>

Ruwan, S. (2023). *Data Mining connects with Software Testing*. LinkedIn.

<https://www.linkedin.com/pulse/data-mining-connects-software-testing-sanjaya-ruwan-istqb-ctfl-/>

Sähkövertailu.fi. (n.d.). *Kodin ja kodinkoneiden sähkönkulutus*. sähkövertailu.fi.

<https://sahkovertailu.fi/kodin-ja-kodinkoneiden-sahkonkulutus>

Sanchez, V. (2023). *The history of Siri and its impact on today's technology*.

<https://blog.routinehub.co/the-history-of-siri-and-its-impact-on-todays-technology/>

SAS. (n.d.-a). *Artificial Neural Networks*. SAS.

https://www.sas.com/en_us/insights/analytics/neural-networks.html

SAS. (n.d.-b). *Deep Learning*. SAS.

https://www.sas.com/en_us/insights/analytics/deep-learning.html

SAS. (n.d.-c). *Generative AI*. SAS.

https://www.sas.com/en_us/insights/analytics/generative-ai.html

SAS. (n.d.-d). *Machine Learning*. SAS.

https://www.sas.com/en_us/insights/analytics/machine-learning.html

Singh, P. (2023). *Every time you talk to ChatGPT it drinks 500ml of water; here's why*.

Business Today. <https://www.businesstoday.in/technology/news/story/microsofts-water-usage-surges-by-thousands-of-gallons-after-the-launch-of-chatgpt-study-397951-2023-09-11>

Singla, R. (2024). *Model based testing in Test Automation*. Testsigma.

<https://testsigma.com/blog/model-based-testing/>

SONY. (1999). *Sony Launches Four-Legged Entertainment Robot*. SONY.

<https://www.sony.com/en/SonyInfo/News/Press/199905/99-046/>

Srinivasan Desikan, G. R. (2007). *Software Testing: Principles and Practices*. Pearson India.

Taulli, T. (2019). *Artificial Intelligence Basics: A Non-Technical Introduction*. Apress.

Testim. (n.d.-a). *Pricing*. Testim.

<https://www.testim.io/pricing/>

Testim. (n.d.-b). *Testim*. Testim.

<https://www.testim.io/developers/>

Testsigma. (n.d.-a). *Top 12 Software Testing Tools List To Look For*. Testsigma.

<https://testsigma.com/tools/software-testing-tools/>

Testsigma. (n.d.-b). *Which Testsigma plan is right for me?* Testsigma.

<https://testsigma.com/pricing>

Testsigma. (n.d.-c). *Why Choose Testsigma?* Testsigma.

<https://testsigma.com/why-choose-testsigma-for-continuous-testing>

Tricentis. (n.d.). *Careers*. Tricentis.

<https://www.tricentis.com/company/careers>

VALA. (2023a). *Miten löytää parhaat tekoälytyökalut ohjelmistotestaukseen?* VALA.

<https://www.valagroup.com/fi/blogi/miten-loytaa-parhaat-tekoalytyokalut-ohjelmistotestaukseen/>

VALA. (2023b). *Yksikkötestaus: Mitä se on ja miksi se on tärkeää?* VALA.

<https://www.valagroup.com/fi/blogi/yksikkotestaus/>

Waseda University (n.d.). *Humanoid Robotics Institute, Waseda University*. Humanoid Robotics Institute, Waseda University.

https://www.humanoid.waseda.ac.jp/booklet/kato_2.html

Wacker, M. (2015). *Just Say No to More End-to-End Tests*. Google Testing Blog.

<https://testing.googleblog.com/2015/04/just-say-no-to-more-end-to-end-tests.html>

Warren, T. (2014). *This is Cortana, Microsoft's answer to Siri*. The Verge.

<https://www.theverge.com/2014/3/3/5465264/microsoft-cortana-windows-phone-screenshots>

Wetzel, A. (2023). *The Dartmouth Conference (1956) and its Lasting Influence on Artificial Intelligence - back to the beginning of AI*. LinkedIn.

<https://www.linkedin.com/pulse/dartmouth-conference-1956-its-lasting-influence-back-arthur-wetzel/>

Appendix 1. The survey questions

1. Where is your office located?
2. In your own view, how much AI is being used in software testing in the organization you are currently working? 1= Not at all, 6= Very much
3. Based on your experience, how have AI tools or technologies impacted the efficiency and effectiveness of your work? 1= Not at all, 6= Very much
4. Are you currently using any AI technologies or tools in your work?
5. How are you, or the customer you are working for, using AI tools related to software testing?
6. List the AI tools or technologies you or your customer is using and how those are being used
7. Are there any AI tools or technologies that you're considering integrating into your work or would like to try to use in your work, as a software tester?
8. What testing related issues or tasks would you like AI to help or solve?
9. What challenges or obstacles are preventing you to adopt AI tools or technologies into your work as a software tester?
10. Can you give any examples of a situation/s where you have faced challenges or obstacles when adopting AI tools or technologies?
11. Are there any lessons learned or recommendations you would like to share with your colleagues who are interested in adopting AI in their work?
12. Is there anything else you would like to add regarding your experience with AI tools or technologies in software testing?
13. What is your general feeling about using AI in your work as a software tester? 1= Not happy or don't like at all, 6= Very enthusiastic and curious
14. Are you interested in being involved in AI related projects or development in the future?