Track #3: Hegemonie verschieben

Report:

From “Mechanical Turk” to “Amazon Mechanical Turk”: The hidden role of Human Intelligence in Artificial Intelligence

Introduction

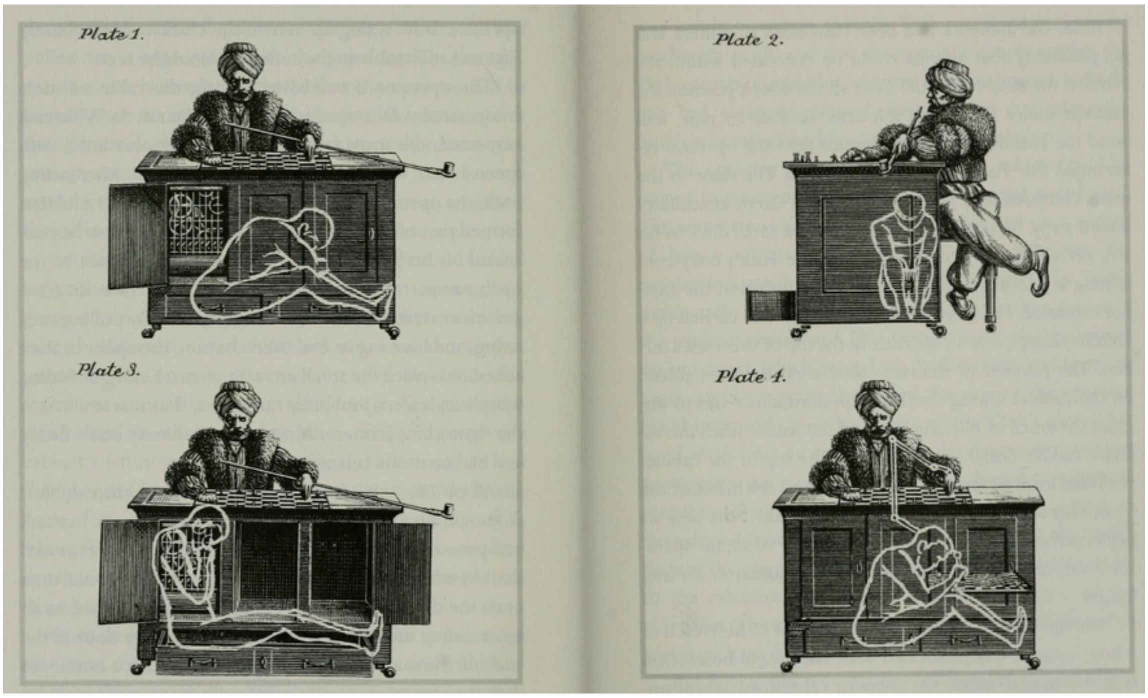
What is the role of human intelligence in artificial intelligence? While machines are often credited for their capabilities, the contributions of humans remain less visible. This report aims to highlight the human element in the machine learning process. We start with the "Mechanical Turk", an 18th-century chess-playing machine that secretly relied on human operation. Moving to the present, we discuss "Amazon Mechanical Turk", a platform by the cloud provider Amazon Webservices(AWS) that employs human workers behind a programming interface to perform tasks a computer can’t solve. We then present a research example of teaching a robot to recognize trash image patterns, emphasizing the human effort in data structuring and labeling. Through historical and current examples, this report underscores the human involvement in automation and the Machine Learning Lifecycle. The base of every application using Artificial Intelligence is collective human intelligence.

Mechanical Turk

The Mechanical Turk, often simply referred to as "The Turk", was a famous chess-playing automaton built in the late 18th century. It was created by the Hungarian inventor Wolfgang von Kempelen (1734-1804), who also developed the "speaking machine", one of the earliest approaches to creating artificial speech (Brackhane et al. 2017: 25).

"The Turk” was a wooden cabinet with a chessboard on top. It contained various machines such as gears and levers to give the impression that the machine was playing chess by itself. A wooden figure dressed in Turkish costume sat behind the chessboard and maneuvered the pieces. Tom Standages' research (2002), chapter 2, describes the exact layout of the machine. Although it appeared to be a functioning chess player, the “Mechanical Turk” was not an actual device capable of playing chess. A skilled chess player would have hidden in a cabinet and manipulated the mechanisms to move the chess pieces around the board.

The interior was constructed to enable the concealed operator to observe the board and execute moves. Standages illustrates (picture 1.1) how the placement of the human operator within the wooden enclosure is the basis for numerous hypotheses to verify that "Mechanical Turk" is not entirely autonomous (Standages 2002: 198-199). Nevertheless, the secret was maintained for 85 years.



Picture 1.1: Four diagrams showing how the Turk’s operator concealed himself by moving back and forth on a sliding seat and opening and closing various folding partitions (Standage 2002:198-199)

Standage dedicates a chapter to the individuals who supplied computational power to the Mechanical Turk. They were predominantly male and accomplished chess players, such as Johann Allgaier, William Lewis, Aaron Alexandre, and Jacques Mouret. Mouret was a relative of the original promoter and exhibitor of the Turk. Some also served as operators and directors, including Johann Maelzel and William Schlumberger. Standage also writes about “a young Frenchwoman" (2002: 207-208). Unfortunately, he does not mention her name.

Based on the design of the „Mechanical Turk”, we can infer about the working conditions of the operators. The operator had to fit into a compact and restricted space within the cabinet of “the Turk”. Limited mobility was required while maneuvering the machinery to move the chess pieces, while staying concealed. Minimal ventilation might have made it challenging to breathe over an extended period due to the stuffiness of the space. Chess games can last for extended periods which can lead to operators being confined in those conditions. Additionally, the operators had to keep the secret of “the Turk's” operation, which may have added stress or pressure to their role.

Amazon Mechanical Turk (AMT)

The “Mechanical Turk” served as the inspiration for the "Mechanical Turk" platform offered by Amazon Web Services (AWS). As one of the largest cloud providers, AWS offers various on-demand services, including compute, storage, databases, analytics, networking, mobile, developer tools, management tools, IoT, and security services. Additionally, the platform provides the option of human compute:

> “Amazon Mechanical Turk is a web service that provides an on-demand, scalable, human workforce to complete jobs that humans can do better

> than computers, for example, recognizing objects in photos." (Amazon Mechanical Turk API Reference)

Common use cases include developing, managing, and assessing machine learning workflows, human-subject research, and application programming interfaces (APIs) that involve humans in applications. Essentially, this means you can integrate humans to perform tasks that computers are incapable of completing and conceal them behind a programming interface. From a developer's standpoint, this is fantastic since it allows for the confluence of human and machine computation through the integration of microservices. The high levels of abstraction found with AWS Services can create the illusion of possessing extraordinary abilities to complete tasks without the need of humans or servers. However, it's just you running a script. You don't see servers and humans around the world executing tasks behind the scenes.

Image 2.2 "Inside the AI Factory", by Josh Dzieza (New York Magazine 2023)

Millions of individuals globally work covertly for Silicon Valley companies, imparting knowledge to algorithms to perceive reality, as Josh Dzieza points out (New York Magazine 2023). This task is expected to broaden as AI advances, ultimately revolutionising the information economy. Among the most significant instances of crowd employment is "Amazon Mechanical Turk" platform (Bergvall-Kåreborn & Howcroft 2014). Bergvall-Kåreborn and Howcroft (2014) depict the monotonous nature of 'clickwork' micro-tasks, which are often laborious and paid inadequately, with remuneration frequently falling below the minimum wage threshold. The working conditions are described as follows:

> "Workers registering on MTurk operate under self-employment regulations and have to agree that any services they perform are deemed 'work made

> for hire' for the benefit of the requester, and in this respect any ownership or intellectual property rights reside with the requester. Amazon also specifies that workers are not entitled to any of the benefits that a requester or MTurk make available to their own employees, which includes holiday pay, sick leave, health insurance, retirement benefits and compensation in the event of injury." (Bergvall-Kåreborn & Howcroft 2014)

From a more general point of view, clickwork resembles care work, as it is the basis and the condition for everything. Drawing bounded boxes around cats or weapons can be as time-consuming, monotonous and sometimes dirty as changing diapers, cleaning the toilet or caring for loved ones. Click work enables artificial intelligence, care work enables society. Both are essential but mostly invisible. Lilly C. Irani and M. Six Silberman have already focused on the invisibility of workers in their paper "Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk," published in 2013. Not too many researchers were doing this 10 years ago. The good news is that more and more researchers and journalists (Williams et al., Glanz 2021, Dzieza 2023) are now asking the question: Who are the exploited workers behind AI? And working conditions are being problematized. Annette Gilbert talks about digital forms of Taylorism and the global precariat of the present:

> "Mit dieser digitalen Form des Taylorismus, die man mit Aude Launay als »Algotaylorismus« bezeichnen könnte, geht einher, dass die Crowd nicht mehr in ihrem Potenzial als Schwarmintelligenz gefragt ist, sondern sich in der Rolle des globalen Prekariats der Gegenwart wiederfindet. Die Crowdworker agieren zumeist in einer rechtlichen Grauzone zu äußerst fragwürdigen Bedingungen: prekär beschäftigt, unterbezahlt, nicht sozialversichert (pseudo-)selbstständig, gewerkschaftlich unorganisiert, ohne Arbeits- und Kündigungsschutz, der ständigen Bewertung ihrer Performance in einem höchst kompetitiven globalen Umfeld sowie der Willkür der Auftraggeber*innen unterworfen, die – damit wirbt Amazon offensiv – die Annahme und somit die Bezahlung der erbrachten Leistung verweigern können und trotzdem das Eigentum daran erlangen, ohne dass die Worker dies anfechten können."(Gilbert 2021)

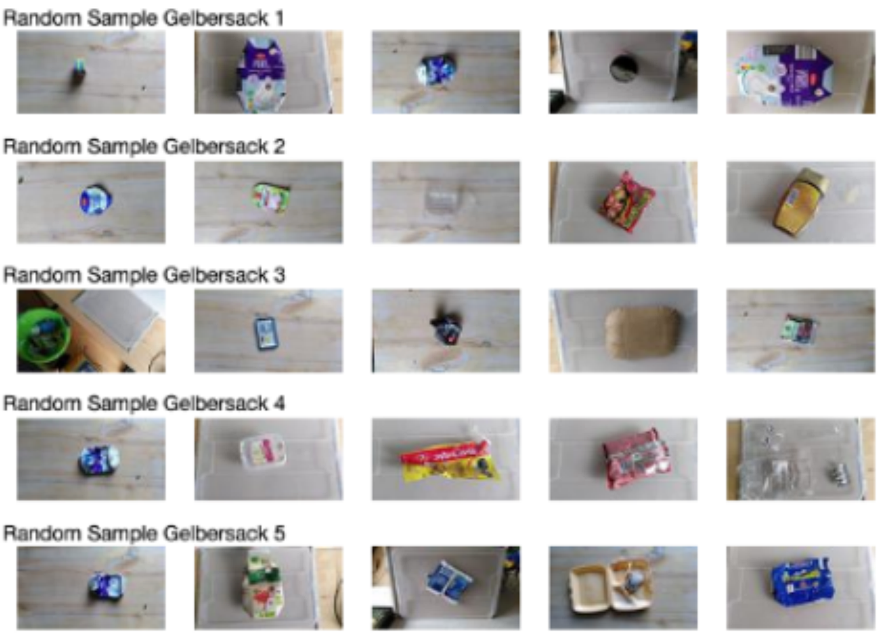
Research example: Teach a robot to recognize patterns in trash images

Big tech companies like AWS know how time-consuming machine learning lifecycles are, that is why they offer services like Amazon Mechanical Turk. In particular, the first two stages (data collection, data annotation) of the machine learning project lifecycle (data collection, data annotation, model training and validation, deployment) require intensive manual labor and are critical to the performance of any machine learning model (Pande 2022:891). Let us take a closer look at the first two stages through a thought experiment. To understand the impact of clickwork on a real-world problem we dive into my area of research:

Imagine that you want to build a robot to sort plastic waste in order to increase the recycling rate.

There are several issues in this experiment, but we will focus on the recognition task for now. We want our robot to recognise post-consumer plastics collected in the German "Gelber Sack". The sensors used will be RGB cameras and we will need about 100,000 images to train our baseline model. A baseline model is like a starting point, it is the simplest model to start with to solve a problem. By comparing the results of more complex models with this baseline model, it is possible to see if the added complexity is advantageous in achieving better results.

The first step is to see if we can find a suitable dataset to train our model on. The machine learning model represents the software side of our robot and we are moving into the field of computer vision. Waste is a popular topic in the machine learning community. There are various so-called "waste" or "trash" datasets. For example, Agnieszka Mikołajczyk provides a comprehensive overview in her repository "Waste dataset review". The programmer classifies the datasets by name, number of categories, number of subcategories, type of annotation (classification, detection, segmentation), type of waste and context, and the platform on which the dataset was published. The 21 datasets all contain RGB images representing waste. However, the majority of the datasets do not focus on post-consumer plastics, and none of the datasets contain images of post-consumer plastic waste from the German "Gelber Sack". The conclusion is that we need to create our own dataset of post-consumer plastics to train our robot. This means we need to scan our rubbish. Let us say we asked all our friends and their friends to scan their rubbish for a month, and now we have 100,000 rubbish images for our baseline model. A random sample of our trash dataset might look like this:



Picture 3.1: Random samples of post-consumer plastics from “Gelber Sack”, a German recycling system for packaging waste, where consumers dispose of lightweight packaging materials in a yellow bag (own picture)

Now we move on to the second stage of our machine learning lifecycle: Data annotation. To annotate data, we need an annotation tool and categories to annotate the given images. In this thought experiment, we choose object detection as our methodology. Object detection refers to the computational technique employed in computer vision that allows for the identification and localization of specific objects within images or videos. Object detection is like giving a computer a pair of magic glasses. When the computer looks at a picture with these glasses, it can also point out and show where things are, like "There is a yoghurt cup ", "That is a Tetra Pak", or "I see a plastic bag here."

Annotation in object detection is foundational for training precise models, and utilizing a well-structured taxonomy is essential to this process. For example, when training a model to recognize various types of plastic waste, a detailed taxonomy would differentiate between "Plastic bottles" such as water bottles, soda bottles, and milk bottles. When annotating images, bounding boxes would be drawn around each item, and they would be labeled according to these specific categories. By employing such a taxonomy during the annotation process, the model learns not just to identify the broader category of plastic waste but also to distinguish between its subcategories like "Tetra Paks" for juice packaging and "Household product packaging" for shampoo bottles. By using such a taxonomy during the annotation process, the model is not only taught what each object looks like and where it's located but also the finer nuances between similar objects. The precision of annotations, combined with a thoughtfully chosen taxonomy, ensures the model's enhanced ability to detect and classify objects with greater accuracy. For our experiment it means, that we need to put bounding boxes to 100.000 trash pictures, before we can start with an actual training process. From my own experience it might take a few seconds to a couple of minutes.

If it takes 45 seconds to annotate one image, it would take 1,250 hours to annotate 100,000 images. This is very time consuming and is the reason why platforms like Amazon Mechanical Turk are successful business concepts used by the scientific community. Who would want to go through 100,000 images of garbage (and that is by far not the worst object to annotate)?

Conclusion

Automation can perform repetitive tasks faster than humans, resulting in increased productivity and throughput. This key promise of automation is currently impossible in the early stages of an artificial intelligence lifecycle. The introduction of Amazon Mechanical Turk as a service acknowledges the limits of automation by promoting human computation for precarious conditions. Humans teach machine learning models to recognize objects, millions of clickworkers are the core of what we call artificial intelligence, and they do not get any credit.

The invisibility and precariousness of crucial work is nothing new in capitalism, but the denial of human intelligence behind any system that uses artificial intelligence is disturbing. The story of the Mechanical Turk, written by Standage (2002), is deeply connected to the history of chess. But this story could also be interpreted as the beginning of a fake hegemony of machines. Fake in the 18th century, because young intelligent people operated the machine. Fake in the 21st century, because (as the garbage example hopefully shows) every AI learns by finding patterns in huge amounts of data, and that data first has to be organised and labelled by humans. The question remains: who has an interest in the hegemony of AI systems, and why? And why did the directors of Mechanical Turk set out to create an illusion?

I put this question to my favourite chatbot and this is the answer:

In essence, the Mechanical Turk was a product of its time, blending the era's fascination with mechanical wonders, the allure of mystery, and the potential for financial and intellectual rewards.

References

Amazon Mechanical Turk API Reference - Amazon Mechanical Turk (no date). Available at: https://docs.aws.amazon.com/AWSMechTurk/latest/AWSMturkAPI/Welcome.html (Accessed: 13 September 2023).

Bergvall-Kåreborn, B. and Howcroft, D. (2014) 'Amazon Mechanical Turk and the commodification of labour', *New Technology, Work and Employment*, 29(3), pp. 213–223. Available at: https://doi.org/10.1111/ntwe.12038.

'Bremer Netzresidenz' (no date). Available at: https://www.beritglanz.de/netzresidenz/ (Accessed: 13 September 2023).

Dzieza, J. (2023) *Inside the AI Factory*, *Intelligencer*. Available at: https://nymag.com/intelligencer/article/ai-artificial-intelligence-humans-technology-business-factory.html (Accessed: 13 September 2023).

Gilbert, A. (2021) 'Kollaterales Schreiben. Digitale Kollaboration im Zeitalter von Crowdworking und Algotaylorismus', in H. Bajohr and A. Gilbert (eds) *Digitale Literatur II*. edition text + kritik im Richard Boorberg Verlag, pp. 62–74. Available at: https://doi.org/10.5771/9783967075496-62.

Introduction to the Theory of Computation, 3rd Edition - 9781133187790 - Cengage (no date). Available at: https://www.cengage.uk/c/introduction-to-the-theory-of-computation-3e-sipser/9781133187790/ (Accessed: 11 September 2023).

Irani, L.C. and Silberman, M.S. (2013) 'Turkopticon'. Available at: https://doi.org/10.1145/2470654.2470742.

Levitt, G.M. (2000) *The Turk, Chess Automaton*. McFarland, Incorporated, Publishers.

Pande, B. *et al.* (2022) 'A Review of Image Annotation Tools for Object Detection', in *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC). 2022 International Conference on Applied Artificial Intelligence and*

Computing (ICAAIC), pp. 976–982. Available at: https://doi.org/10.1109/ICAAIC53929.2022.9792665.

Standage, T. (2002) *The Turk : the life and times of the famous eighteenth-century chess-playing machine*. New York : Walker & Co. Available at: http://archive.org/details/turklifetimesoff00stan (Accessed: 11 September 2023).

Williams, A. (2022) 'The Exploited Labor Behind Artificial Intelligence'. Available at: https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence (Accessed: 11 September 2023).

Context & about the author

My colleagues describe me as a "hybrid". I studied political science and computer science. Before spending the last five years as a devops engineer for a safety company, I spent my days as a trade union organizer. Currently, the "Human in Command" research group at Aalen University has given my various interests a home. I'm starting my PhD. My idea is to create a citizen science project for the creation of post-consumer plastics open datasets, taking into account explicable AI methods in the machine learning lifecycle.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used chatGPT-4 in order to assist the writing process and for general definitions in the machine learning process. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.