On Humans, Uncertainty and System Engineering

Lesandro Ponciano[1]
System Engineering, TU Dresden

August 10th 2022

Talk Description

Since the first electronic computer systems, human beings have been considered an *inspiration* or a *metaphor* for system engineering and theory. The work of human computers in the first half of the twentieth century provided Alan Turing with insights into conceptualising some of the first debates about the feasibility of machine intelligence [1]. In the second half of that century, Hebert Simon, laureated with the Nobel prize and Turing award, drew on the parallel between human cognition and computer architecture to illustrate his Bounded Rationality theory [2], addressing the reasoning process under limited data, storage and computing resources. Such a theory has proved essential to understanding unconscious human heuristic decision-making and some foundations of Behavioural Economics. It sheds light on human and computer imperfect decisions and systematic bias.

Both computers and humans can reach a final wrong result to a computing task even when no malicious behaviour is involved. Computer scientists developed the *dependability* area to address the faults, errors, and failures threats in machine computation, achieving fault prevention, tolerance, removal, and forecasting [3]. Similarly, experimental psychologists developed the *Human Error* theory that draws on the mistakes, lapses and slips in human processing to improve human efficacy and efficiency [4]. Elements from both efforts improved overall performance when a human (named "*user*") interacts with a computer interface to perform a task. Beyond the system's functional requirements of providing correct outcomes, the system's non-functional requirements associated with human satisfaction started to be demanded in such a user-centred design. "User" performance and experience became relevant to the system's adoption. Distributed systems mechanisms also benefited from understanding human behaviour, e.g., user habits in accessing websites [5] or using computing resources [6] have led to caching and energy-saving mechanisms. But increasingly, the human position started to go beyond the "user" role.

The first decade of this century witnessed the reCAPTCHA tools [7] provide some insights into the superior efficacy of humans compared to machine computers when performing some types of tasks. The areas of Human Computation and Collective Intelligence developed further on it, focusing on computations in which humans (named "*computers*" or "*workers*") showed more significant efficacy than machine-based computational systems [8]. Advancements allowed the design of large distributed human computing systems that reached innovations in many areas, including biology [9] and astronomy [10, 11]. *Machine-in-the-loop* or *Human-in-the-loop* are articulating both human cognition and machine embedded with artificial intelligence to perform their *computing affinity* task. Humans become much more than inspirations, metaphors or users. They become an essential part of the computation. As a comparatively new phenomenon, there is still little knowledge on how system engineers can *cope with uncertainty coming from machines and humans* in this context.

Todays's context of the socio-technical system leads us to ask much more complex questions about the effectiveness of the system's internal mechanisms and strategies. The systems are no more based only on the concept of *correctness*. Many types of uncertainty from probabilistic and artificial intelligence algorithms and human decision-making heuristics are involved [8]. Fault prevention, tolerance, removal, and forecasting are no more enough to guarantee the absence of catastrophic consequences for humans and the environment. Static and dynamic verification that emerged as essential quality assurance techniques in

[1] Contact: lesandro.ponciano@tu-dresden.de | lesandro.ponciano@gmail.com | 0000-0002-5724-0094

traditional software engineering seems insufficient in the face of uncertainty. Credibility/believability approaches arose to assess human outputs, and accountability emerged as a requirement to assess algorithmic decisions [12]. Automatic explanations can be essential in both assessments [12,13], but sometimes they conflict with other non-functional requirements such as privacy and confidentiality.

In *non-factual tasks* that are not judged by correctness*, how can the assessment of uncertainty be done as part of the system?* Consensus reached on redundant work (usually obtained by active replication mechanism), its combination with inference rules and stochastic models may inform some decisions [14], but they can also be biased. Let's consider, for instance, privacy beyond legislation definition. Designers must cope with the human non-factual perception of privacy if they are concerned with system adoption. In this case, the one-size-fits-all approach does not suit [15]. The design space can become significantly large and even conflicting by seeking individualisation, flexibility and compliance with general rules. Not shying away from a complex debate, one can say that part of the solution is that the humans who use the system or are part of it must also design it, considering their perspectives and requirements. This notion of a “*system created in use*” is an analogy to Noam Chomsky's perception of “*language created in use*” [16], which significantly changed the scientific study of human language at the end of the twentieth century. Indeed, language balances general rules, individualisation and flexibility, but how to engineer a *system* like human engineer *language*?

All this debate shows an *outside-in pattern*. Since humans become “users”, systems have been re-engineered over time to take advantage of human cognition and address non-functional requirements associated with human subjectivity. Each step informs changes incrementally closer to the core logic of the system. It started with the usability of the interface, then the algorithmic design and explanation capabilities. In re-engineering the system layers (or stack) to meet new requirements, one provoking question is: *how deep will it lead us?* Like the linguistic papers used to do, theoretical computer science papers still do not talk about humans [17]. Are the basic levels of our computer systems robust enough for brain-computer interaction and the solution to new non-functional requirements to be successfully implemented? Or, *will we need more fundamental changes in the inner system mechanism? What about the security mechanism?* Answers to these questions can have implications on the theoretical understanding of what means computing systems be “natural’ for humans in all their potential. This understanding can inform research and development proposals toward their engineering and adoption.

References

[1] Turing, A.M., 1950. Computing machinery and intelligence. *Mind* 59(236), pp. 433-460.
[2] Simon, H.A., 1990. Invariants of human behavior. *Annual review of Psychology*, 41(1), pp.1-20.
[3] Avizienis, A., Laprie, J.C., Randell, B. and Landwehr, C., 2004. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, 1(1), pp.11-33.
[4] Reason, J., 1990. Human Error. Cambridge University Press.
[5] Arlitt, M.F. and Williamson, C.L., 1996. Web server workload characterization: The search for invariants. *SIGMETRICS Perform. Eval. Rev.* 24(1), pp.126-137.
[6] Ponciano, L. and Brasileiro, F., 2010, October. On the impact of energy-saving strategies in opportunistic grids. In E2CG2 2010 11th IEEE/ACM International Conference on Grid Computing (pp. 282-289).
[7] von Ahn, L., Maurer, B., McMillen, C., Abraham, D. and Blum, M., 2008. reCAPTCHA: Human-based character recognition via web security measures. *Science*, 321(5895), pp.1465-1468.
[8] Ponciano, L., Brasileiro, F., Simpson, R. and Smith, A., 2014. Volunteers' engagement in human computation for astronomy projects. *Computing in Science & Engineering*, 16(6), pp. 52-59.
[9] Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D. and Popović, Z., 2010. Predicting protein structures with a multiplayer online game. Nature, 466(7307), pp.756-760.
[10] Lintott, C.J. et al. 2008. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. Monthly Notices of the Royal Astronomical Society, 389(3), pp.1179-1189.
[11] Ponciano, L., Brasileiro, F., Andrade, N. and Sampaio, L., 2014. Considering human aspects on strategies for designing and managing distributed human computation. *Journal of Internet Services and Applications*, 5(1), pp.1-15.
[12] Nunes, I., Jannach, D., 2017. A systematic review and taxonomy of explanations in decision support and recommender systems. *User Modeling and User-Adapted Interaction* 27(3), pp. 393-444.
[13] Ramos, H., Fonseca, M., Ponciano, L. 2021. Modeling and Evaluating Personas with Software Explainability Requirements. Human-Computer Interaction. HCI-COLLAB 2021. Communications in Computer and Information Science, vol 1478. Springer, Cham.
[14] Ponciano, L. and Brasileiro, F., 2018. Agreement-based credibility assessment and task replication in human computation systems. *Future Generation Computer Systems*, 87, pp.159-170.
[15] Ponciano, L., Barbosa, P., Brasileiro, F., Brito, A. and Andrade, N., 2017. Designing for pragmatists and fundamentalists: Privacy concerns and attitudes on the internet of things. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems* (pp. 1-10).
[16] Chomsky, N., 1986. Knowledge of language: Its nature, origin, and use. Greenwood Publishing Group.
[17] de Souza, Clarisse Sieckenius. 2018. A pragmatic turn in computer science. *Interactions* 25, 3, pp.20-21.