Insight Titanium Research

An Interdisciplinary Approach to Artificial Intelligence Testing
Development of an Artificial Intelligent Quotient (AIQ)

by Marisa Tschopp and Marc Ruef
November 2017

)SCIP(

While more people than ever now are confronted with *artificial intelligence* (intentionally or not), it is still hard to find a common ground of understanding of the concept. The complexity lies within the term itself by adopting the word *intelligence* as technical, procedural capabilities of machines. It has inherited a myriad of challenges from the long history of psychological intelligence research.

Consequently, to reach out for such common ground, artificial intelligence research inevitably requires an *interdisciplinary approach*, primarily focusing on *psychology* and *computer science*. This certainly is a challenging undertaking per se in any field, but specifically as those two disciplines are very disparate by nature.

Human and Artificial Intelligence in Context

Many researchers and professionals (and also non-professionals, laypersons and computer novices) have undertaken more or less serious projects to *measure and compare artificial intelligence*, predominantly with digital assistants like for example Siri (as a part of Apple's operating system) or Google Now (developed by Google). Such tests have various perspectives: between digital assistants, trying to answer the question: who is the smartest digital assistant, or between Humans and digital assistants trying to find out, whether the AI can «outsmart» the human.

However, these test results must be read with high caution and skepticism. Often there is a massive lack of profound conceptualization or operationalization of intelligence. For example, many *do not* disclose the methodology of their testing procedures, which is essential to make a valid judgement as a reader. If you look at scientifically valid human intelligence tests, they are always adapted to specific target groups depending on age, culture, language or disabilities. Hence, it is obviously non-sensical to just use any valid IQ test (such as the *Wechsler Intelligence Scale* or the *Stanford-Binet Intelligence Scale*) for testing artificial intelligence abilities.

Rethinking the Measurement of AI Abilities

To avoid shallow and dauntless statements such as «Siri's IQ is only 30» or «Siri has the IQ of a 5-year-old», it is necessary to rethink the way in which artificial intelligence is measured and situated into appropriate context. Human Intelligence and intelligence testing has been a very controversial and

multi-faceted topic and has undergone dramatic changes in history and therefore is a topic not to be taken lightly in any way.

To understand artificial intelligence from an interdisciplinary perspective within the current environment the huge corpus of psychological academic literature concerning human intelligence must be confronted. Prerequisite and therefore the first step for the development of an *AIQ test*, is to derive a concept of intelligence, which is understood as a system of abilities specifically adapted to the AI environment. This intelligence model integrates various categories, such as numerical reasoning, working memory or explicit knowledge.

As the next step of the research project an intelligence test (or scale) is developed, that mirrors classic intelligence testing categories and procedures from academically constructed IQ tests. One important difference in this research is the evaluation through answer categories as well, for example: knowledge versus understanding. All questions as well as all answers are individually weighted. Several pilot tests are constructed and executed with commonly known, well-used digital assistants, namely Google Now (by Google), Siri (by Apple) and Alexa (by Amazon) to customize the IQ test appropriately. The overall goal is to state a valid conclusion about an *AI's Intelligence Quotient*, hereby referred to as *AIQ*, as a reliable measurement of general ability as well as its progressive development over time.

Interdisciplinarity is the Key

This research is conducted to raise key questions and answers about the understanding, measurement and comparison of human and artificial intelligence. Noticeable, peculiar incidents with reference to test construction and performance will be presented and discussed. Furthermore, implications for practice are considered, such as forecasting the development of abilities of digital assistants or how companies or institutions can make reasonable, consumer-sensitive use of artificial intelligence. It is of vital importance to also evaluate the test critically, repeatedly over time and make necessary changes if relevant. Hence strengths and weaknesses of the *AIQ test* are assessed and implications for future research will be discussed as well.

This research project aims to trigger interdisciplinary exchange and conversation as well as deeper understanding based on solid academic research focusing on practical, applicable solutions.

References

[1] Zimbardo, P., Gerrig, R., & Graf, R. (2008). Psychologie. München: Pearson Education.

[2] Kaufman, J., & Sternberg, R. (2010). The Cambridge handbook of creativity. New York: Cambridge University Press.

[3] Stone, P. et al. (2016). Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA

About the Authors
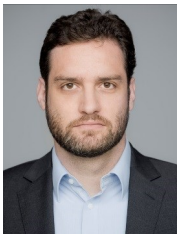
Marc Ruef is Head of Research and Member of the Board at scip ag. His focus is on conducting research in the broad field of cybersecurity, including computer and network security, vulnerability assessment and penetration testing. As an entrepreneur and expert in information technology he has extensive expertise in the development, application and capabilities of artificial intelligence (AI) solutions with specific interest on language and reasoning capabilities, taking security related, business and societal implications into account. Marc has published various best-selling books, more than 400 white papers and media articles, is teaching at several higher education institutions and is a frequent speaker at conferences and workshops in the public and private sector worldwide.

Marisa Tschopp is a researcher at scip ag. Her focus is on conducting research about AI-based systems from a humanities perspective, with a wide range of questions related to psychological phenomena, governance and ethical implications. As an organizational psychologist she has experience in social and educational institutions with specific passion for digital teaching-learning trends. She published various papers and conference contributions on the topics of leadership, creativity and innovation and has been teaching in several higher education institutions in Germany and Switzerland. Marisa holds a Master's degree in Psychology of Excellence in Business and Education from the Ludwig-Maximilians-University of Munich, Germany as well as a BA business degree, focusing on market and consumer psychology.

Mission: Know the Future

The Titanium Research Team is the highly independent for-profit research department which explores essential issues and innovations in the field of emerging technologies. We pursue an interdisciplinary approach comprising technical and non-technical considerations to guarantee the preparation of our customers in the best possible way.

Our research endeavors include a range of projects tapping into various activities such as knowledge building, education, training, consulting and technical & non-technical development. Topics include a broad range of fields like blockchain technologies, drones, smart weapons, robotics and artificial intelligence, autonomous vehicles, and many more, all examined through an interdisciplinary lens to measure social-psychological impact, ethical implications, and additionally, to forecast future development for the public good. We want to build knowledge, strengthen collaboration between research, industry, policy makers and users in order to explore and comprehend the nature of digitalization and emerging technologies.

scip ag is a Swiss company providing consulting services in the field of cybersecurity. Security is our business. https://www.scip.ch/en/