

|   |                                      |   |
|---|--------------------------------------|---|
| <b>Institution:</b> University of Hertfordshire   |                                      |   |
| <b>Unit of Assessment:</b> 11 – Computer Science and Informatics  |                                      |   |
| <b>Title of case study:</b> Designing socially intelligent adaptive systems to inform commercial AI development and engage the public in debates over human-robot interaction   |                                      |   |
| <b>Period when the underpinning research was undertaken:</b> 2005 – 2020  |                                      |   |
| <b>Details of staff conducting the underpinning research from the submitting unit:</b>  |                                      |   |
| <b>Name(s):</b>   | <b>Role(s) (e.g. job title):</b>     | <b>Period employed by submitting HEI:</b> |
| Daniel Polani   | Professor of Artificial Intelligence | 2002 – present                            |
| Lola Cañamero   | Reader in Adaptive Systems           | 2001 – present                            |
| Farshid Amirabdollahian   | Professor of Human-Robot Interaction | 2009 – present                            |
| Christoph Salge   | Research Fellow                      | 2011 – present                            |
| <b>Period when the claimed impact occurred:</b> 2014 – December 31, 2020  |                                      |   |
| <b>Is this case study continued from a case study submitted in 2014?</b> N  |                                      |   |
| <b>1. Summary of the impact</b> (indicative maximum 100 words) <p>The Adaptive Systems Research Group (ASRG) at the University of Hertfordshire has developed new concepts and formalisms to design robots that interact with humans and each other in a more socially intelligent, effective way. These have informed: novel artificial intelligence (AI) algorithms by world-leading robotics and technology companies, including Google DeepMind, which described the UH contribution as ‘<i>crucial</i>’ to developing stronger AI algorithms; R&amp;D approaches in the automotive sector; the design of a mobile robot assistant that has generated £2.1m in revenues; the first international standards to regulate the design of personal care robots. The research inspired an annual competition to use AI to create cities in the popular online game Minecraft, which has engaged around one million people, via YouTube and social media site Reddit, in the application of AI to urban planning. It has increased public understanding of what the future relationship between humans and robots might look like, through partnerships with leading museums and arts centres that engaged around 100,000 people and sustained media engagement that reached an audience of millions.</p>  |                                      |   |
| <b>2. Underpinning research</b> (indicative maximum 500 words) <p>The development of sophisticated AI presents society with opportunities, risks and challenges. Autonomous systems must be able to interact in ways that humans find intuitive if cooperation between people and machines is to be meaningful, productive and safe. Research by ASRG has sought to study how animals and humans take flexible decisions to adapt to changes in their environment and harness these principles for the design of more socially competent, emotionally responsive and versatile autonomous systems. The Group has pursued this overarching goal through three interlinked research strands: development of the concept of <b>Empowerment</b>, involving the study of information flow in cognitive agents to understand how to endow robotic systems with biologically plausible, effective and flexible cognitive architectures; <b>assistive robotics</b>, focusing on how machines can interact with, and provide support to, humans in a more empathic and socially acceptable way; <b>affective robotics</b>, developing internal models of emotions for robots so they make ‘better’ decisions and interact with humans more effectively.</p> <p><b>Empowered robotics:</b> In 2005 Polani introduced ‘Empowerment’ [3.1], a quantifiable information-theoretic measure that captures how much an agent is in control of the world it perceives. Inspired by the ability of organisms to select behaviours that help them survive, Empowerment provides agents with intrinsic motivations that might lead to more complex behaviour. It is defined as the amount of information that an agent can send from its actuators to its sensors via the environment, leading to behaviour that enhances its ability to act and affect its surroundings. Polani led work to validate empowerment as a reliable, consistent measure [3.2] and explore ways to mathematically code and compute the concept. In 2017 Polani and Salge proposed Empowerment as a replacement for Asimov’s Three Laws of Robotics, to form a key part of a robot’s ethical behaviour. The idea is that a robot sees the world through the eyes of the human with whom it interacts; keeping the human safe consists of the robot acting altruistically to increase the human’s own Empowerment [3.3]. Building on this work through a</p> |                                      |   |

Marie Curie fellowship [G1], Salge explored how the development of intrinsic motivation formalisms could be used for the automatic generation of engaging computer games. Content generation is a production bottleneck in game design; this is being addressed by procedural content generation, in which evolutionary algorithms modify and adapt content to optimise gameplay. Working with New York University's Game Innovation Lab, Salge evaluated whether these formalisms can capture actual human motivation.

**Assistive robotics:** Under EU-FP7 project CORBYS [G2], UH research focused on robot systems that have a symbiotic relationship with humans, exploring cognitive control architectures that allow machines to anticipate and react to human behaviour, based on biologically inspired information-theoretic principles. This included research by Polani and Salge into maximising Empowerment in a robot to help it avoid collisions with other unpredictable agents (e.g. humans). This led to the first realisation of the CORBYS control architecture for a robot following a human co-worker [3.4]. Having contributed to CORBYS, Amirabdollahian coordinated the EU-FP7 ACCOMPANY project [G3], which researched scenarios in which robot companions were used to facilitate independent living for older people. Studies involved the use of mobile robot assistant Care-O-bot 3 to provide care in a motivating and socially acceptable way, exploring ethical and safety considerations. It found that those in receipt of care were more receptive to the robot than caregivers [3.5].

**Affective robotics:** The EU-FP7 ALIZ-E project [G4] studied the development of interactive robots to support children being treated for diabetes. A consortium used the humanoid Nao to develop the 'Integrated System'. It could play educational games with children, act as a coach and motivator, and was largely operated by humans. Key UH contributions, led by Cañamero, were to develop motion animation modules and a number of expressive elements to convey basic emotions. This built on previous related work under the EU FP-6 project Felix Growing. UH developed its own robot Robin, who was cognitively and motivationally autonomous and moved around the room while it interacted with children [3.6]. Robin 'had' diabetes; it could independently explore its environment and decide what to do based on its needs, affective state and interactions. The use of embodied AI allowed natural and personalised interactions with the children.

### 3. References to the research (indicative maximum of six references)

- 3.1** Klyubin AS, Polani D, Nehaniv CL. All Else Being Equal Be Empowered. Lecture Notes in Computer Science. 2005;2005:744-753. [https://doi.org/10.1007/11553090\\_75](https://doi.org/10.1007/11553090_75)
- 3.2** Klyubin AS, Polani D, Nehaniv CL. Keep Your Options Open: An Information-Based Driving Principle for Sensorimotor Systems. PLoS ONE. 2008;3(12). e4018. <https://doi.org/10.1371/journal.pone.0004018>
- 3.3** Salge C, Polani D. Empowerment As Replacement for the Three Laws of Robotics. Frontiers in Robotics and AI. 2017 Jun 29;4. 25. <https://doi.org/10.3389/frobt.2017.00025>
- 3.4** Leu A, Ristic-Durrant D, Slavnic S, Glackin C, Salge C, Polani D et al. CORBYS cognitive control architecture for robotic follower. In 2013 IEEE/SICE International Symposium on System Integration (SII). IEEE. 2013. p. 394-399 <https://doi.org/10.1109/SII.2013.6776636>
- 3.5** Bedaf S, Gelderblom GJ, Syrdal DS, Lehmann H, Michel H, Hewson D, Amirabdollahian F, Dautenhahn K, de Witte L. Which activities threaten independent living of elderly when becoming problematic: inspiration for meaningful service robot functionality. Disability and Rehabilitation: Assistive Technology. 2014 Nov;9(6):445-452. <https://doi.org/10.3109/17483107.2013.840861>
- 3.6** Canamero L, Lewis M. Making New "New AI" Friends: Designing a Social Robot for Diabetic Children from an Embodied AI Perspective. International Journal of Social Robotics. 2016 Jun 30. <https://doi.org/10.1007/s12369-016-0364-9>

#### Key underpinning grants

- G1** Horizon 2020 INTERCOGAM, 2016-2019, €251,858; all of the grant awarded to UH.
- G2** EU-FP7 CORBYS, 2011-2015, €8.8m; € 780,800 to UH.

**G3** EU-FP7 ACCOMPANY, 2011-2014, €4.8m; €1.2m to UH.

**G4** EU-FP7 ALIZ-E, 2010-2014, €10.7m; €1.0m awarded to UH.

#### 4. Details of the impact (indicative maximum 750 words)

UH research into the design of more socially intelligent adaptive systems enabled new industry approaches to AI algorithmic development and machine learning. Concurrently, a programme of public engagement increased public understanding and awareness of the opportunities and ethical and technical challenges arising from closer human-robot interaction.

##### Enabling commercial development of human-robot interaction and AI algorithms

Polani's concept of Empowerment has influenced the development of novel AI algorithms by technology and robotics companies. A Principal Research Scientist at Google DeepMind has confirmed that: *'Research into the notion of empowerment, as developed by D. Polani and colleagues, is utterly crucial to developing strong(er) AI algorithms, and we can confirm we did a substantial amount of follow-up research and spent a large amount of engineering time based on this work ... We do not (yet) have commercial products based on this, however, we can confirm that we have agents that successfully use the fruits of this line of research as 'auxiliary cost functions'. It is a perfect example of academic research strongly influencing the processes and research/development employed in industry [5.1].'* In a 2015 paper, Google DeepMind researchers presented a new algorithm that *'allows for scalable information maximisation and empowerment-based reasoning directly from pixels to actions' [5.2]*. It cited 3.1, 3.2 and two other Polani papers.

Japan-based Araya, which develops deep learning algorithms, has embedded the Empowerment concept in its product development principles. Araya invited Polani to discuss Empowerment at its headquarters in July 2017 and Salge to talk about it at a Tokyo workshop in January 2020. Araya wrote: *"Research published by ... ASRG has a strong impact on [our] own research efforts towards the development of autonomous agents. A public record of this impact is Araya's publication 'Learning body-affordances to simplify action spaces' ... This is based on the concept of empowerment developed by the group of Professor Daniel Polani. The Autonomous Agent Team at Araya continues to use empowerment as a core principle in its research and development [5.3]."* Empowerment is also being used in the automotive sector. BMW researchers provided an artificial agent with intrinsic motivation using Empowerment for safer self-driving vehicles. Describing their approach as using the *"concept of human empowerment introduced by Salge and Polani [3.3]"*, their 2020 paper said: *"...a robot employing our approach strives for the empowerment of people in its environment, so they are not disturbed by the robot's presence and motion. In our experiments, we show that our approach has a positive influence on humans, as it minimizes its distance to humans and thus decreases human travel time while moving efficiently towards its own goal [5.4]."*

The expressive postures developed by Cañamero for the Nao robot under **G4** benefitted its developer, world-leading robotics company Softbank (formerly Aldebaran Robotics), from 2014 onwards. The company offered customers access to the postures as part of the Nao software. Softbank Robotics Europe said this work was *"the foundation of something very important for our humanoid robot"*. Their Software Innovation Director wrote: *"I've (had) frequent requests to use these animations for research, apps or even ads."* The postures formed the basis of an interactive game app called *Guess The Emotion*, which was published by Aldebaran (before being acquired by Softbank) in 2015 [5.5]. In this game of emotional recognition, the Nao robot cycles through its different emotional states and asks users to correctly classify them. Education technology company Teq wrote on its website that the emotions app, as part of the 'AskNAO platform', was designed to help students within a special education classroom environment, *'learn to recognise emotional cues in themselves and in others'*.

Amirabdollahian's research under **G3**, using Care-O-bot 3, *"informed the design and development of Care-O-bot 4 by providing input in user interaction and user interface design as*

well as usability aspects of the platform” according to the CEO of Mojin Robotics, which has developed the robot commercially since 2015 [5.6]. The Care-O-bot 4 went on to generate more than £2.1m in sales and rental revenue over the impact period, winning a number of industry awards [5.6]. The ACCOMPANY research [G3] highlighted safety aspects of human-robot companions. Based on these observations, Amirabdollahian was invited to join the British Standards Institute (BSI) Technical Committee AMT/10 on Robotics. He contributed to two standards: BS EN ISO 13482:2014, which provided the first robot safety requirements for personal care robots, and BS 8611:2016 on the ethical design and application of robots and robotic systems. For the former, his contributions related to safeguarding against potential hazards caused by companion robots in start-up operations and in motion, or due to the shape of the robot, its embodiment or autonomy [5.7a]; for the latter, his contributions centred on ethical risk assessment and guidelines relating to ethics-based robot design [5.7b].

### **Increasing public awareness and understanding of opportunities and ethical challenges arising from the development of adaptive systems – and their current limitations**

Through partnerships with cultural institutions and proactive media engagement, ASRG has sought to engage the public in what socially interactive robots can and cannot do, applying balance and reason to typically excitable debates on whether robots will ‘take over the world’. Paper 3.3 raised the public profile of the Empowerment concept and prompted considered media analysis around the nature of future relationships between humans and AI. Since July 2017, it has been tweeted 84 times by 59 users, reaching an upper bound of 202,219 followers, and has featured in 20 articles by 19 media outlets [5.8]. It was used in an article in US magazine *Reason* to challenge the talking up of the risks posed by AI by Tesla founder Elon Musk. The piece sparked debate through 110 reader comments [5.9]. An accompanying article by Salge in *The Conversation* had 43,749 reads, was tweeted 84 times and shared 317 times on Facebook and was republished by 18 publications, including *Mail Online*, the *Independent*, *Scientific American*, *Newsweek* and *Yahoo News* [5.9]. Salge was interviewed about Empowerment in *Wired* under the headline: *How to Build a Robot That Won’t Take Over the World*. In 2018 Salge discussed Empowerment in a podcast published by the Burning Man Project’s Philosophical Center [5.9].

Extending this work under G1, in 2018 Salge launched an annual competition that challenges regular gamers and aspiring developers to design AI systems that can create entire settlements in popular online game *Minecraft*. Participants in the Generative Design in Minecraft Competition (GDMC) are asked to write an AI program to generate a functional settlement that is aesthetically adapted to the environment and terrain on a given map. The idea is to explore not just whether AI can produce better games, but also better cities. Entries are judged by game designers, urbanists and architects. The 2018 winning entry was featured on *City Beautiful* a popular YouTube channel on urban planning. The December 2018 episode ‘An urban planner reviews Minecraft cities’ was one of the channel’s most popular, attracting 838,851 views, 21,000 likes and 1,706 comments [5.9]. The video prompted debate around the role of AI in urban planning. One commenter wrote: ‘*I think that AI can replace humans in planning new cities ... But I still think that humans would still be better at changing existing cities.*’ The narrator said: ‘*It’s really interesting that we have the concept that we can generate a city algorithmically as opposed to using the brains of planners. Is this the future? Am I going to be out of a job someday because algorithms can produce plans or development projects far better than us people?*’ The winner published their design on social media platform Reddit. It was ‘upvoted’ 11,100 times, indicating a total audience reach of 22,000-55,000 (based on average ratios for the site), and attracted 167 comments, the majority of which were positive, with many expressing surprise at what had been achieved [5.9]. The 2019 GDMC was featured in a 13-minute video on the ‘AI and Games’ YouTube channel, receiving 53,613 views, 2,800 likes and 143 comments [5.9]. Interest in the competition led to Salge being invited by Minecraft’s owners Microsoft to speak about it at a workshop on how AI and machine learning is transforming the gaming landscape [5.9]. The 2020 GDMC was covered by Chinese media [5.9].



RoboCup is an annual international football competition for robot players. The AI challenges posed by RoboCup are closely tied with Polani's research aims: enabling a machine to understand what to do without painstakingly setting out all possible scenarios. Empowerment research by Polani, President of the RoboCup Federation from 2017 to 2019, fed into the creation of UH's RoboCup team Bold Hearts. Bold Hearts were runners up in the 2014 World Championships and quarter finalists in 2017. Polani used these successes to engage the public in the complexity of designing robots that might one day be able to compete with humans. In 2015, the Bold Hearts team reached a wide audience through broadcast and online media coverage, including BBC News, Sky News, Children's BBC shows Technobabble and Newsround, CNBC and The Guardian. Bold Hearts were demonstrated at the Science Museum in April 2017 as a key part of the two-day *Robots Revealed* event [5.10]. According to the Science Museum, 1,620 visitors engaged with the exhibit and UH researchers, with an additional 4,782 visitors watching. From May to August 2019, Bold Heart robot Ywain participated in the Barbican Centre's exhibition *AI: More than Human* which received 88,811 visitors and was described by organisers as "one of the most successful exhibitions to have been hosted at the Barbican" [5.11]. Ywain was "a key live element of the A.I. timeline created to showcase the essential moments in the ever growing A.I. history" [5.11]. The exhibition's success resulted in it being shown in the Dutch Arts Centre Forum Groningen from December 2019 to August 2020, with plans finalised to take it to Liverpool and Madrid.

Between 2015 and 2018 Robin the diabetes robot received high-profile coverage in publications including the *New York Times*, *Le Monde* and *BBC News* [5.9]. This led to two high-profile live demonstrations of Robin, accompanied by video material of Cañamero's emotion modelling research: a Friday evening event called *Feeling Emotional* at the Wellcome Collection in February 2016 and a two-day interactive exhibit at London Science Museum's Frankenstein Festival in April 2018, reaching thousands of visitors [5.9]. G3 led to a £578,000 EPSRC facilities grant to upgrade UH's Robot House, a residential home adapted into a research facility to explore the role of robots as carers. This received in-depth coverage in a BBC One series on ageing *Holding Back the Years* (2018), BBC Three documentary *Can Robots Love Us?* (2017; also receiving 307,507 YouTube, 2,800 likes, 611 comments), BBC News (2015), The Guardian (2016) and CNet.com (2017) [5.9].

## 5. Sources to corroborate the impact (indicative maximum of 10 references)

- 5.1 Corroborating statement from a Principal Research Scientist at Google DeepMind.
- 5.2 Mohamed S, Rezende DJ (Google DeepMind, London), Variational Information Maximisation for Intrinsically Motivated Reinforcement Learning, Proceedings of the 28th International Conference on Neural Information Processing Systems, Vol 2, pp. 2125-2133, December 2015. <https://dl.acm.org/doi/10.5555/2969442.2969477>
- 5.3 Corroborating statement from the President of Araya Inc, Tokyo.
- 5.4 van der Heiden T (BMW), Mirus F (BMW), van Hoof H. (2020) Social Navigation with Human Empowerment Driven Deep Reinforcement Learning. In: Farkaš I., Masulli P., Wermter S. (eds) Artificial Neural Networks and Machine Learning – ICANN 2020. ICANN 2020. Lecture Notes in Computer Science, vol 12397. Springer, Cham. [https://doi.org/10.1007/978-3-030-61616-8\\_32](https://doi.org/10.1007/978-3-030-61616-8_32)
- 5.5 Corroborating statement from a Software Innovation Director at Softbank Robotics.
- 5.6 Corroborating statement from the CEO of Mojin Robotics, Germany.
- 5.7a BS EN ISO 13482:2014. Robots and robotic devices – Safety requirements for personal care robotics; relevant sections: 5.4, 5.6, 5.10, 5.12. <https://www.iso.org/standard/53820.html>
- 5.7b BS 8611: 2016. Robots and robotic devices – Guide to the ethical design and application of robots and robotics systems; relevant sections: sub sections 4 and 5. <https://tinyurl.com/yya55f2u>
- 5.8 Altmetric details for the article [3.3] <https://www.altmetric.com/details/21403879>
- 5.9 Evaluation report of the media coverage of ASRG research over the impact period.
- 5.10 The Bold Hearts RoboCup team featured in *Robots Revealed* article by the Science Museum: <https://blog.sciencemuseum.org.uk/robots-revealed-meet-the-robots/>
- 5.11 Corroborating statement from an Exhibition Manager at the Barbican Centre.