



Fabio Roli

University of Genova and Cagliari, Italy

From known knowns to unknown unknowns in AI: Historical and Technical Issues

AI has been originally developed for *closed-world*, and *noise-free*, problems where the possible states of natures and actions that a rationale agent could implement were perfectly known. One could argue that, at that time, AI dealt with *known knowns*. Since the 1980s, when machine learning became an experimental science, AI researchers started to tackle pattern recognition problems with noisy data, using probability theory to model uncertainty and decision theory to minimize the risk of wrong actions. This was the era of *known unknowns*, characterized by the rise of benchmark data sets, larger and larger year after year, and the belief that real world problems can be solved collecting enough training data. However, prior work has shown that available data sets have often a limited utility when used to train pattern recognition algorithms that will be deployed in the real world. The reason is that modern machine learning has often to face with *unknown unknowns*. When learning systems are deployed in adversarial environments in the *open world*, they can misclassify (with high-confidence) never-before-seen inputs that are largely different from known training data. *Unknown unknowns* are the real threat in many security problems (e.g., zero-day attacks in computer security). In this talk, I give a historical and technical overview of the evolution of AI and machine learning for pattern recognition and discuss how this evolution can be regarded as a transition from *known knowns* to *unknown unknowns*, and the key role that adversarial machine learning can play to make AI safer.

Short bio of the speaker

Fabio Roli is a **Full Professor** of Computer Engineering at the University of Genova and Cagliari, Italy, **Director** of the **Pattern Recognition and Applications laboratory** at the University of Cagliari and **co-director** of the **SmartLab** at the University of Genova. He has been doing research on the design of pattern recognition and machine learning systems for thirty years. He has been appointed **Fellow** of the **IEEE**, **Fellow** of the **International Association for Pattern Recognition**, and **Fellow** of the **Asia-Pacific Artificial Intelligence Association**. He was a member of **NATO advisory panel** for Information and Communications Security, NATO Science for Peace and Security (2008 – 2011). Prof. Roli is the recipient of the 2020 **Pattern Recognition Medal** of the international scientific journal Pattern Recognition, and the 2020 **IAPR Pierre Devijver Award**.