Foreword*

In June 2018, the European Commission has appointed a "AI High Level Expert Group" (AI-HLEG) to support the implementation of the European Strategy on Artificial Intelligence. One of the first results of the AI-HLEG has been to deliver ethics guidelines on Artificial Intelligence.[1] These guidelines put forward a human-centered approach to AI, and list seven key requirements that human-centered, trustworthy AI systems should meet, summarized by the following headers:

1. Human agency and oversight
2. Technical robustness and safety
3. Privacy and data governance
4. Transparency
5. Diversity, non-discrimination and fairness
6. Societal and environmental wellbeing
7. Accountability

Many of today's most popular AI methods, however, fail to meet these guidelines: making them compliant is a scientific endeavor that is as crucial as it is challenging and stimulating. As an example, systems based on deep learning point often provide impressive results, but their ability to explain these results to the user is limited, thus challenging requirements 4 and 7; we lack general ways to formally verify their correctness and assess their boundary conditions, thus challenging requirement 2; and we don't yet have methods to allow humans to collaboratively influence or question their decisions, thus challenging requirement 1. Similar criticalities are present in many other popular AI methods.

The question addressed by this workshop is: what are the scientific and technological gaps that we have to fill in order to make AI systems *human-centered* in terms of the above guidelines? To answer this question, this workshop leverages three keynote talks presenting the ethics guidelines and the scientific and technological approach by two major European projects; the presentation of fifteen contributed papers; and a joint work based on a brain-walk exercise.

These proceedings collect the contributed papers. For the outcomes from the discussions, see the workshop web page `http://nehuai2020.aass.oru.se/`.

Alessandro Saffiotti, Luciano Serafini, Paul Lukowicz
Workshop chairs

[1] `https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines`