Preface

Linked Data have attracted a lot of attention from both developers and researchers in recent years, as the underlying technologies and principles provide new ways, following the Semantic Web standards, to overcome typical data management and consumption issues such as reliability, heterogeneity, provenance or completeness. Many areas of research have adopted these principles both for the management and dissemination of their own data and for the combined reuse of external data sources. However, the way in which Linked Data can be applicable and beneficial to the Knowledge Discovery in Databases (KDD) process is still not completely understood.

The Linked Data 4 Knowledge Discovery workshop (LD4KD), co-located within the ECML/PKDD2014 conference in Nancy (France), explores the benefits of Linked Data for the very well established KDD field. Beyond addressing the traditional data management and consumption KDD issues from a Semantic Web perspective, the workshop aims at revealing new challenges that can emerge from joining the two fields.

In order to create opportunities for communication as well as collaboration channels, the workshop accepted 8 research papers from practitioners of both fields. The first observation one can make from those contributions is that the most obvious scenario for using Linked Data in a Knowledge Discovery process is the representation of the underlying data following Semantic Web standards, as shown in [De Clercq et al.], [Bloem et al.] and [Krompass et al.], with the aim of simplifying the knowledge extraction process. With that said, other contributions targeted other aspects of KDD, such as data pre-processing or pattern interpretation, with the purpose of showing that KDD processes can benefit from including elements of Linked Data. For the purpose of data preparation, [Rabatel et al.] focuses on mining Linked Data sources, while [Zogała-Siudem et al., Ristoski et al.] use Linked Data to enrich and integrate local data. The interpretation step of the KDD process is also addressed, in the work of [Alam et al.] on results interpretation and the one of [Peña et al.] on visualisation. We thank the authors for their submissions and the program committee for their hard work. We sincerely hope that this joint work will provide new ideas for interactions between those two, mostly isolated communities.

Ilaria Tiddi
Mathieu d'Aquin
September 2014 *Nicolas Jay*