

# A Knowledge Graph for the Quality Assessment of Cardiovascular Magnetic Resonance Imaging Data

Valentina Carapella

King's College London

valentina.carapella@kcl.ac.uk

Ernesto Jiménez-Ruiz

The Alan Turing Institute

ejimenez-ruiz@turing.ac.uk

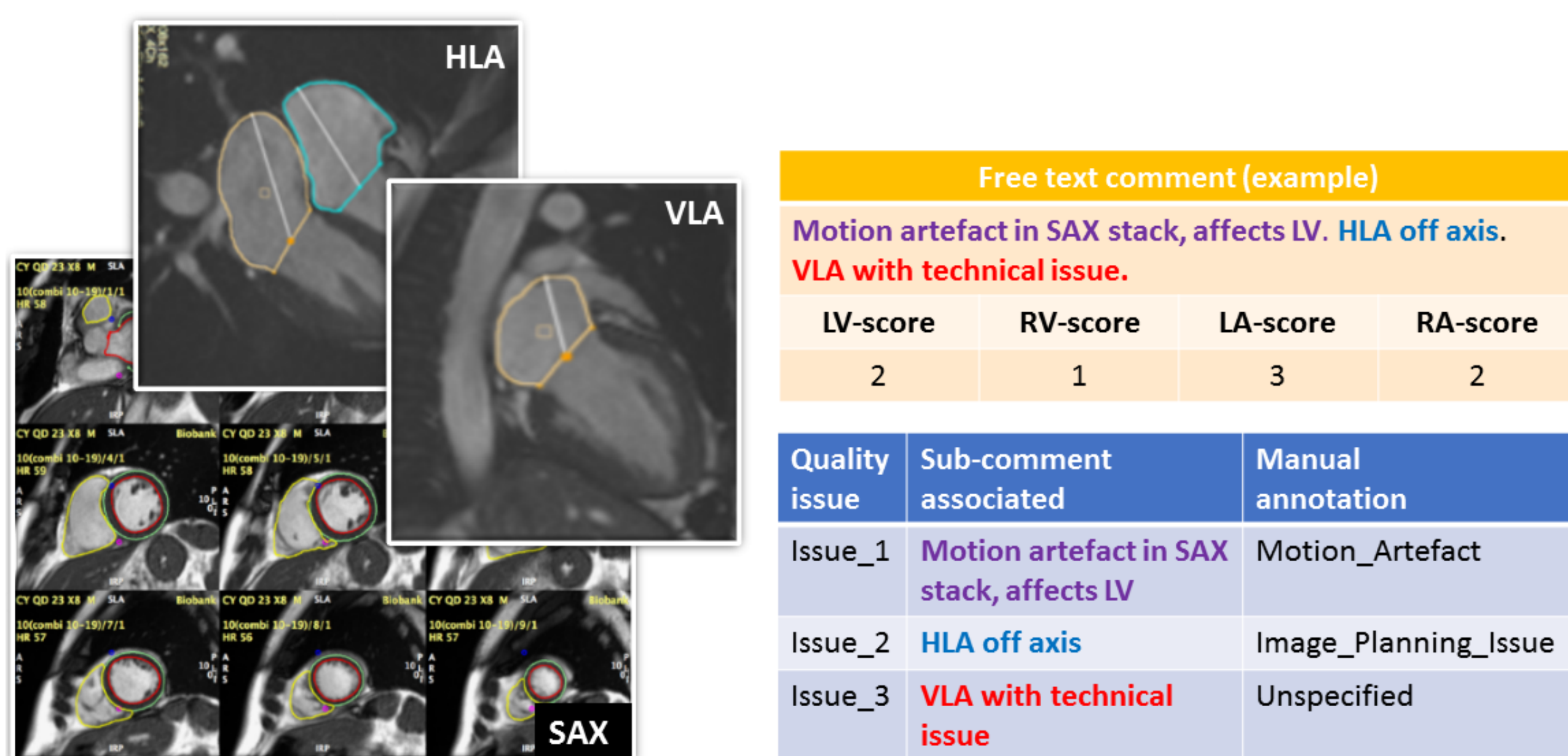
## Motivation

Image quality assessment can greatly benefit from automated approaches to process text-based quality reports. The combination of text mining techniques and knowledge graph representation of QA documentation is a promising approach to:

- Exploit free text information about the quality of images
- Validate manually annotated scans
- Allow efficient querying of large imaging dataset based on image quality features

## Application Scenario: UK Biobank

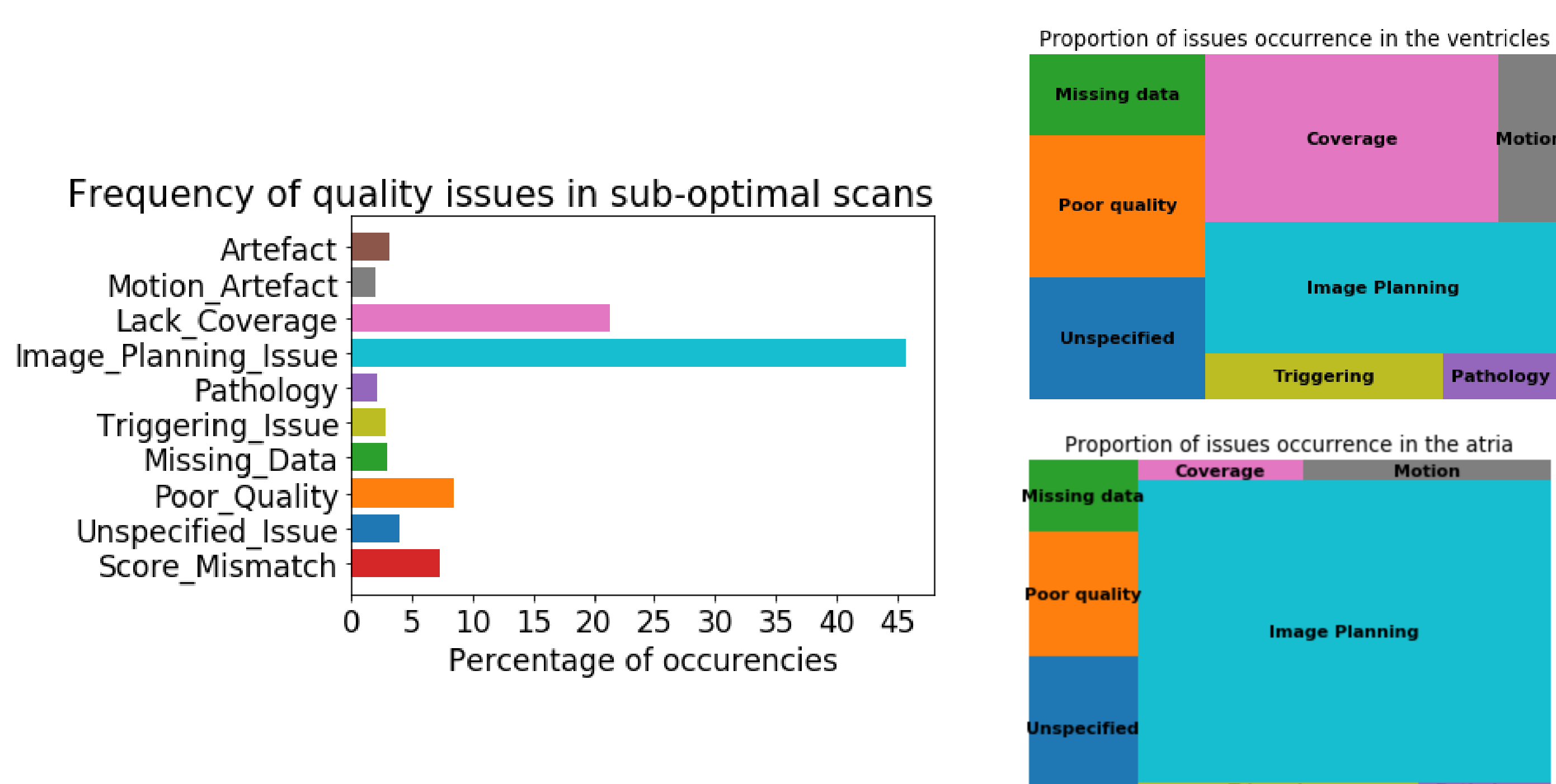
- The UK Biobank CMR pilot study consisted of 5000 scans.
- The cine-MRI part of each scan was contoured to evaluate cardiac function and the image quality assessed by a team of image analysts.
- Contours for left and right ventricles were drawn on short axis (SAX) stack, while left and right atria were contoured on the long axis images (HLA and VLA, respectively).
- The QA consisted of a free text comment about the quality issues encountered in each of the three imaging views (SAX, HLA and VLA). An additional numerical score was given by cardiac chamber, signifying the confidence level of the analyst in the contours drawn in the light of the issues reported in the text.
- Each comment is a composite of individual quality issues, as shown in the figure below.



## Manual annotation

Out of 5000 scans, 1658 (33%) were flagged as sub-optimal or non-analysable. This subset of comments plus scores represents our data. All the comments were manually annotated by decomposing each comment into its constitutive quality issues and separating the issues by cardiac chamber. The aims were:

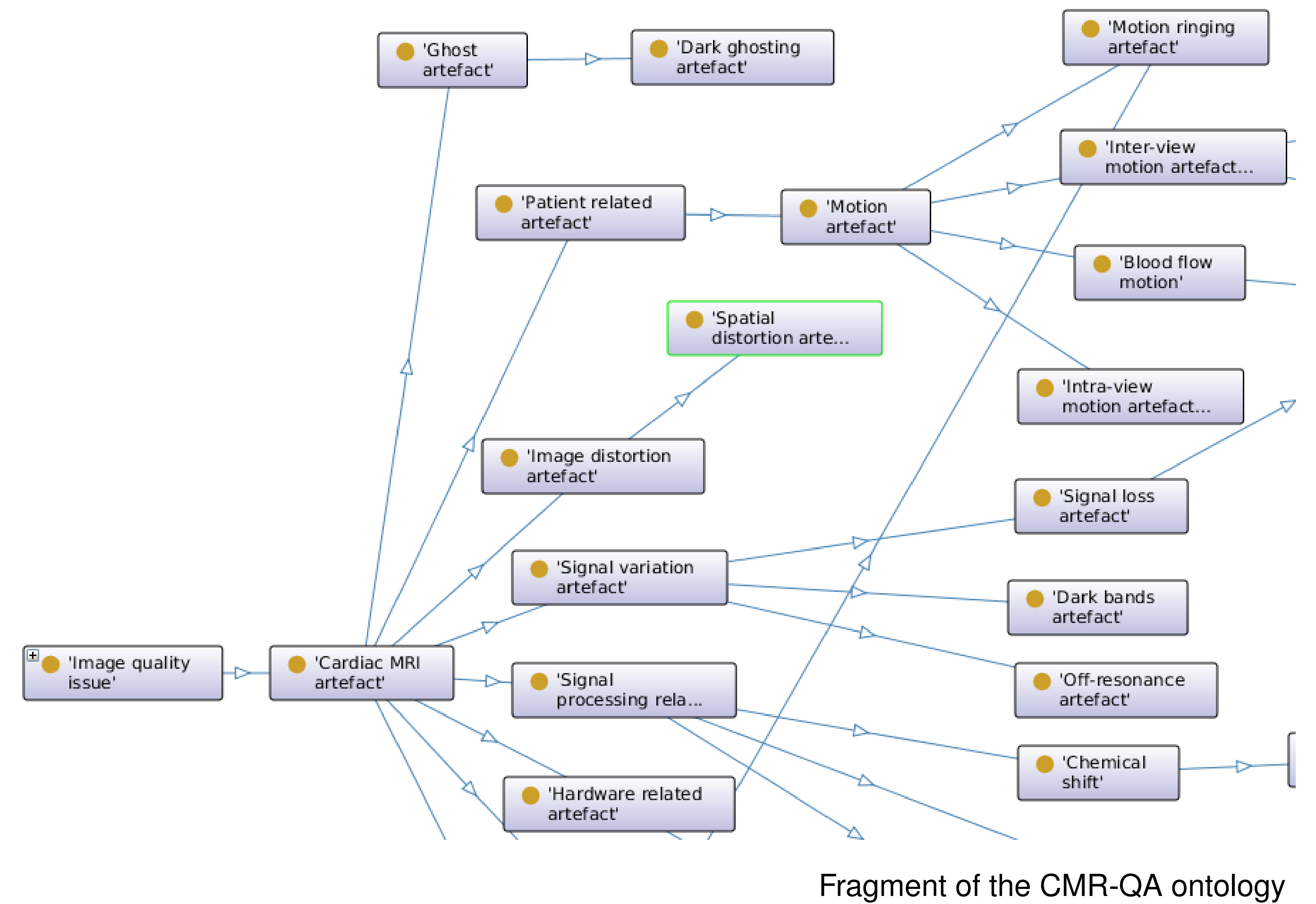
- Dataset curation, by identification of potential mismatches between score and comment
- Validation set for the automatic knowledge graph population



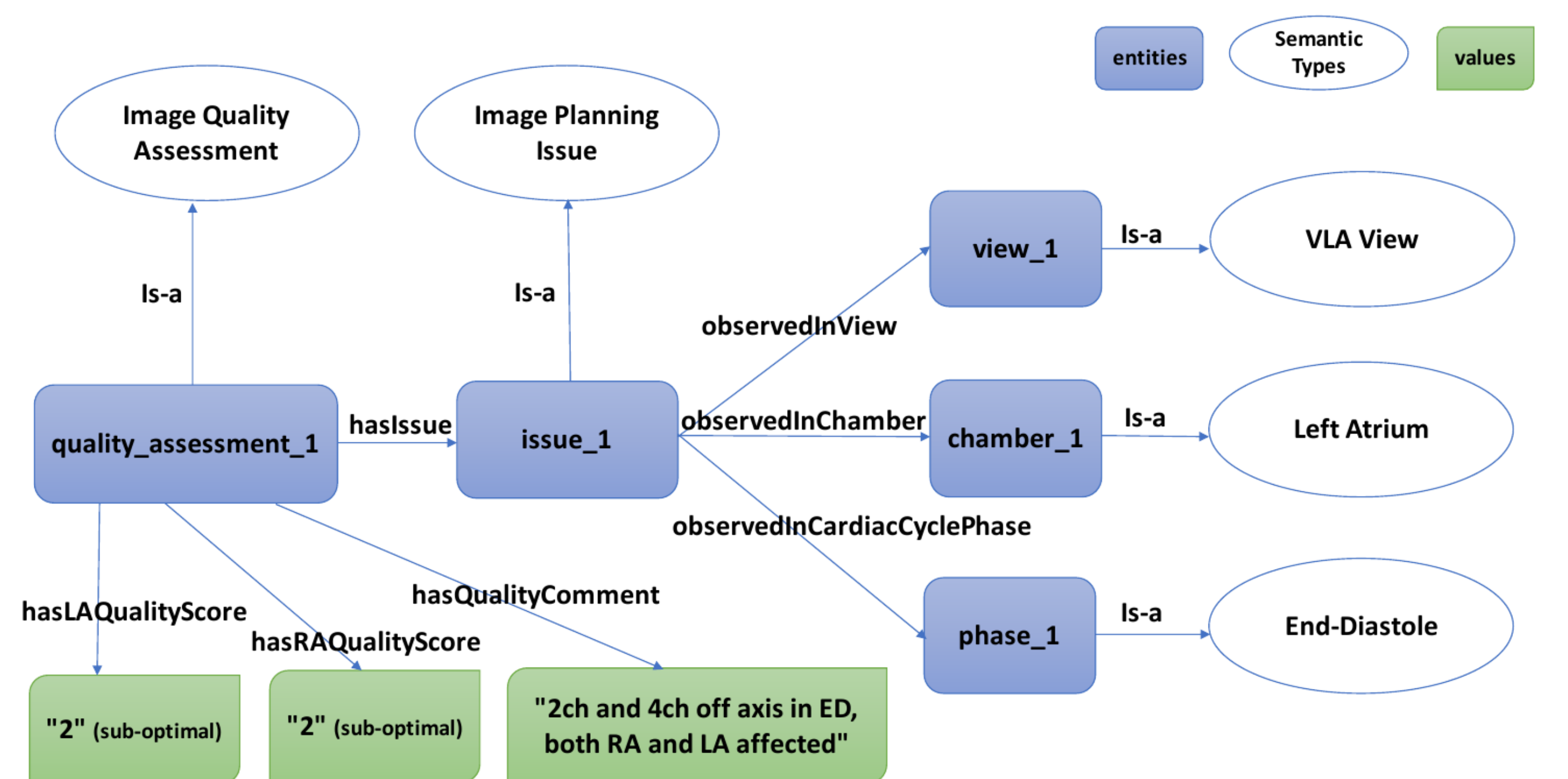
## Knowledge Graph (KG) representation

Definition of KG: "Large network of entities, their semantic types, properties and relationships between entities."

- *CMR-QA Ontology*: ~110 semantic types, ~45 properties
- Available in BioPortal



- *KG population*: from QA assessments (comments and scores) to a network of entities by means of:
  - Automatic extraction from comments
  - Extraction driven by the ontology vocabulary
  - Implicit information extracted via domain specific rules



- *KG-based queries*: Efficient and flexible access of scans with certain characteristics via semantic queries and logical reasoning. For example, it will allow to:
  - Exclude scans where atria are off axis
  - Identify scans with artefacts in SAX
  - Identify cases of incomplete coverage of the ventricular extent, in particular those missing the valve plane.

## Acknowledgements

Work in collaboration with E. Lukaschuk, N. Aung, K. Fung, J. Paiva, M. Sanghvi, S. Neubauer, S. Petersen, Ian Horrocks and S. K. Piechnik. Funded by the British Heart Foundation (BHF) and the National Institute for Health Research (NIHR). Ernesto is also funded by the SIRIUS Centre for Scalable Data Access (Research Council of Norway).



Project



Ontology



Paper