Towards the Semantic Enrichment of Free-text Annotation of Image Quality Assessment for UK Biobank Cardiac Cine MRI Scans

Valentina Carapella[1], Ernesto Jiménez-Ruiz[2], Elena Lukaschuk[1],
Nay Aung[3], Kenneth Fung[3], Jose Paiva[3], Mihir Sanghvi[3],
Stefan Neubauer[1] Steffen Petersen[3], Ian Horrocks[2], Stefan Piechnik[1]

[1] Oxford Centre for Clinical Magnetic Resonance Research (OCMR), Radcliffe Department of Medicine, University of Oxford, UK
[2] Information Systems Group, Department of Computer Science, University of Oxford, UK
[3] William Harvey Research Institute, NIHR Cardiovascular Biomedical Research Unit at Barts, Queen Mary University of London, UK

Abstract. Image quality assessment is fundamental as it affects the level of confidence in any output obtained from image analysis. Clinical research imaging scans do not often come with an explicit evaluation of their quality, however reports are written associated to the patient/volunteer scans. This rich free-text documentation has the potential to provide automatic image quality assessment if efficiently processed and structured. This paper aims at showing how the use of Semantic Web technology for structuring free-text documentation can provide means for automatic image quality assessment. We aim to design and implement a semantic layer for a special dataset, the annotations made in the context of the UK Biobank Cardiac Cine MRI pilot study. This semantic layer will be a powerful tool to automatically infer or validate quality scores for clinical images and efficiently query image databases based on quality information extracted from the annotations. In this paper we motivate the need for this semantic layer, present an initial version of our ontology as well as preliminary results. The presented approach has the potential to be extended to broader projects and ultimately employed in the clinical setting.

1 Introduction

UK Biobank is a large scale population study at the national level aimed at improving the understanding, diagnosis and treatment of a wide range of diseases, such as cancer, stroke or cardiac pathologies [1]. In 2006 the recruitment began of 500,000 volunteers aged 40-69 across UK who underwent a number of clinical tests and agreed to have their health followed. UK Biobank is a complex project addressing multiple organs by means of various clinical tests, including different imaging modalities. The outcome of this study will be available to researchers worldwide.

Within UK Biobank, Cardiovascular Magnetic Resonance Imaging (CMR) plays a fundamental role in the assessment of cardiac function. Each volunteer participating to the UK Biobank imaging arm undergoes a series of MRI sequences to image the heart: Cine MRI, tagged MRI, T1-mapping and Phase-contrast imaging [2, 3]. A pilot study of 5,000 CMR scans has been released and is shared for data analysis with affiliated researchers analysing Cine MRI, the most common CMR sequence in clinical practice.

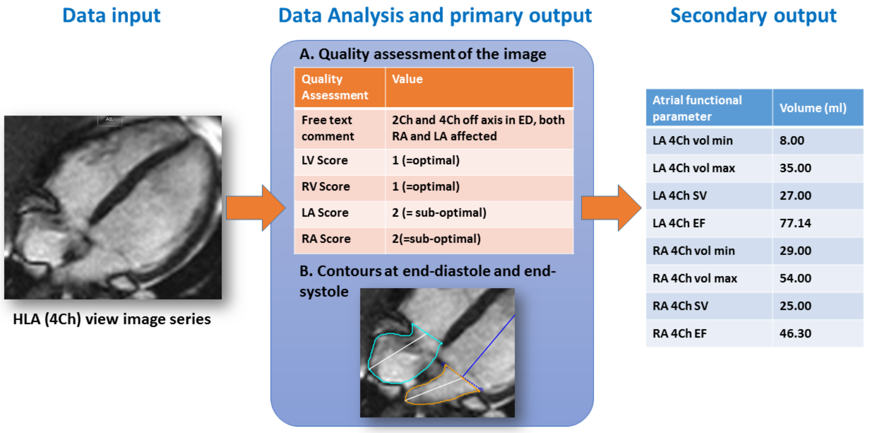
The topic of automatic image quality assessment is of particular importance in relation to the management and post-processing of large scale datasets, such as the UK Biobank pilot study, and more generally in clinical research. Consistent image quality assessment provides means to evaluate how reliable the parameters values obtained from the image analysis are. However, image quality assessment is seldom carried out in an organised and structured way. Clinicians and radiographers might repeat a scan if they detect technical problems in real-time, otherwise, the sub-optimal quality goes undetected and is not recorded. In post-processing, image analysts often discard images with technical problems without further feedback to the image acquisition team. In the context of the UK Biobank CMR pilot study, one of the main aims was to address this lack of cross-talk between the acquisition and post-processing phase. Data analysis of the pilot study was in fact based on two key aspects: quality assessment of the imaging data and manual contour delineation. Figure 1 highlights the two components of the analysis. Quality assessment of the Cine MRI scans was carried out through a combination of free-text comments and numerical quality scores. Manual delineation of contours (also known as segmentation) was carried out for the four chambers of the heart, which then results in the computation of fundamental parameters of cardiac function. As per UK Biobank protocols, all such derived data will be returned to UK Biobank for inclusion in the central database, whence they can be disseminated to other groups with appropriate research approvals.[1]

For the purposes of our work, we focus only on the quality assessment data, which is the combination of free-text annotation and numerical quality scores. The quality scores provide a quick overall classification of the images, for example, for statistical purposes. The free-text annotation is rich in information but cannot be processed in an easy and efficient manner as the numerical scores. A promising efficient solution can be sought in the field of Semantic Web. The semantics of the free-text annotations, which describe the quality of the image analysis, will be defined via a structured vocabulary or ontology, which we are going to call *CMR-QA* (Cardiovascular Magnetic Resonance Quality Assessment). An ontology is an explicit specification of a conceptualisation providing an unambiguous and formal representation of a given domain or field of knowledge [4, 5]. In other words, ontologies provide a controlled vocabulary about the relevant terms in that domain and their relationships. Ontologies are extensively used in biomedicine (e.g., [6]) and in other domains like the energy sector (e.g., [7]). There have also been recent efforts in using ontologies to describe the information within a biobank (e.g., [8–11]).

Generating a semantic layer for the annotations provides the means to structure them in a formal and unambiguous way while retaining all the descriptive power of natural language. This semantic layer will provide machine-readable data and will be a powerful tool for *(i)* fast and efficient processing of the free-text comments; *(ii)* automatic image quality assessment from such comments and generation of quality scores; *(iii)* evaluation of the quality of the free-text comments in terms of information completeness, ambiguity and variability; *(iv)* training purposes (e.g., showing preferred an-

---

[1] `http://www.ukbiobank.ac.uk/register-apply/`



Quality Assessment	Value
Free text comment	2Ch and 4Ch off axis in ED, both RA and LA affected
LV Score	1 (=optimal)
RV Score	1 (=optimal)
LA Score	2 (= sub-optimal)
RA Score	2(=sub-optimal)

Atrial functional parameter	Volume (ml)
LA 4Ch vol min	8.00
LA 4Ch vol max	35.00
LA 4Ch SV	27.00
LA 4Ch EF	77.14
RA 4Ch vol min	29.00
RA 4Ch vol max	54.00
RA 4Ch SV	25.00
RA 4Ch EF	46.30

Fig. 1. Example of analysis pipeline restricted to the case of HLA view. Quality scores are 1 = optimal, 2 = sub-optimal and 3 = unreliable or non-analysable.

notation styles for different types of images); *(v)* efficient semantic access (i.e. database querying) to the images by the UK Biobank target users, such as researchers in the field of automatic segmentation, or clinical researchers who need a specific subset as a control group in their study.

2 Methodological approach

Data analysis for the 5,000 CMR pilot study was carried out by a team of eight *observers* from two clinical research centres. The observers were professionals experienced in this type of analysis but with different backgrounds and expertise. Quality assessment and general data analysis progress was managed through a shared spreadsheet by the team.

2.1 The imaging data

Each individual CMR dataset included in the 5,000 CMR pilot study for the UK Biobank project contains a series of MRI scans of the heart aiming at imaging different aspects of cardiac function and structure. However, for the purposes of the first release of data, only a subset of the images acquired were analysed. Data analysis was initially restricted to the following Cine MRI views: *(i)* Short axis (SA). Left and right ventricle (LV and RV) are contoured in this view at two phases of the cardiac cycle, end-diastole and end-systole. For the left ventricle both endocardial and epicardial contours are drawn. For the right ventricle only the endocardial contour is considered. *(ii)* Horizontal long axis (HLA or 4Ch), also referred to as four chamber view. Left and right atrial (LA and RA) endocardium are contoured in this view at two phases. The first phase is ventricular end-diastole and it provides the minimal atrial volume. The second phase is identified by the

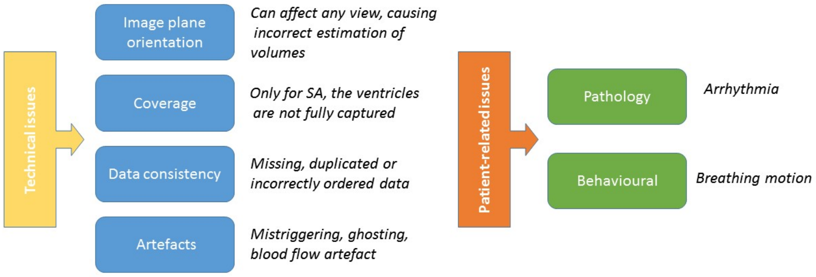


Fig. 2. Quality assessment parameters with examples.

opening of the mitral valve, and it is used to obtain the maximal atrial volume. *(iii)* Vertical long axis (VLA or 2Ch), or two chamber view. Only left atrium is contoured at the same phases of the cardiac cycle as HLA.

2.2 Quality assessment of the data

The individual dataset for each participant in the pilot study thus contained the three aforementioned Cine MRI views. Data analysis of each dataset was subdivided into two phases, quality assessment and manual contouring of anatomical structures. The quality assessment part is the focus of this paper. Quality assessment addressed the three views individually, as they are acquired separately. Therefore, for example, a good quality SA image can be paired with a poor quality HLA or VLA images. Figure 1, shows the pipeline of data analysis and output for the HLA view.

The observers were required to evaluate each subset of images according to certain aspects likely to affect image quality. The level of detail provided also varied with the experience and background of the observer. For example, those with clinical experience were able to suggest the presence of a specific pathology, only when this was considered to affect the quality of the image. Those with knowledge of MRI physics provided more insight on the nature of artefacts. Figure 2 shows a diagram of the different aspects of quality assessment the observer were taking into account, divided into technical issues (left-side panel) and patient-related issues (right-side panel). Examples of possible issues are provided for each sub-area of quality assessment.

For those images flagged as sub-optimal or unreliable, the observers were asked to write a short free-text comment summarising, for the three imaging views, the reasons for such decision. This resulted in a wide variability in the level of detail provided, but also in the vocabulary employed. In combination to comments, the observers were also asked to provide a numerical score (1 = optimal, 2 = sub-optimal, 3= unreliable) to summarise how reliable the contours of each cardiac chamber was in the light of the quality assessment given to the images.

2.3 Definition of the semantic layer

We aim at providing a formal description or *semantic layer* for our *domain*, that is the quality assessment (both free-text annotations and quality scores) of the cardiovascular Cine MRI views as performed in the UK Biobank CMR pilot study. The semantic layer will be composed by: *(i)* an *ontology*, containing concepts relevant to our domain (i.e., vocabulary occurring in the free-text annotations or terms relevant to the image analysis) and relationships among those concepts; *(ii) ontology data* which is the re-formulation of the information content of the free-text annotations according to the ontology; *(iii) ontology rules* to infer knew knowledge in combination with the ontology and the ontology data.

Ontology development. The ontology development is still ongoing and greatly benefits from the close collaboration with the observers involved in the image analysis and quality assessment phase. Given the highly theoretical nature of ontology development, in this section we are only providing some key examples, described in plain text, of the key components of an ontology.

Concepts in biomedical ontologies are typically organised in a hierarchy according to *(i)* a *broader-narrower* relationship, that is classifying concepts from general to more detailed ones, e.g., the heart is an organ; *(ii)* a *part-whole* relationship, e.g., the left atrium is part of the heart. In addition, ontology concepts may also be related to each other by means of other relationships, e.g., arrhythmia affects the heart.

With this framework in mind, we have started developing the *CMR-QA* (Cardiovascular Magnetic Resonance Quality Assessment) ontology to include both general knowledge about the domain and more concrete aspects about the image quality assessment.[2] For example, *CMR-QA* encodes general knowledge about concepts such as *Cine MRI Scan* is a kind of *MRI Scan*; but it also encodes more specific knowledge such as *wrong image plane orientation* is a kind of *technical issue* or *RA off axis* is a specific type of *wrong image plane orientation* (please refer to Figure 2). Although the relationship specifying that *Cine MRI Scan* is a kind of *MRI Scan* may not be seen as an interesting piece of knowledge, this knowledge will not be otherwise known by a computer and thus it should be specified. Furthermore, this top-level knowledge will facilitate the integration with other established domain ontologies that do not contain concepts as fine-grained and specific to our domain as the ones provided in *CMR-QA*.

Concepts and relationships belong to the logical realm. However, non-logical knowledge can also be added to the ontology in the form of lexical information (synonyms, comments, cross-references). For example, *CMR-QA* includes that *LA out of plane* is an alternative label or synonym for *LA off axis*, which captures some of the variability observed in the free-text comments (refer to Table 1 for the complete list of variants).

Ontology data via free-text comment mining. We are developing *named entity recognition* (NER) techniques to transform the free-text comments into semantically rich data according to *CMR-QA* (see [12, 13] for a survey). In other words, each free-text comment is decomposed using text-mining techniques into chunks, which will be then associated to statements or triples in the form of <*subject predicate object*> expressions. For example, suppose we have a composite comment "Basal slice missing, wrong plane RA". The text-mining process identifies two quality issues by breaking the sentence into two parts: "Basal slice missing" and "wrong plane RA". Focussing on the

---

[2] *CMR-QA* ontology and related assets can be downloaded from `https://github.com/ernestojimenezruiz/CMR-QA-Semantic-Layer`

latter, the chunk of text "*wrong plane RA*" in the free-text comment is then associated to the triple: $<$*issue*$_i$ *rdf:type RA off-axis*$>$; this is a computer-friendly representation of the fact that there is a quality issue (the subject), uniquely identified here by *issue*$_i$ where i is a counter. This subject belongs to a certain type (the predicate *rdf:type*), the type being *RA off-axis* (the object).

Ontology rules. Ontology rules are being developed together with the ontology to infer additional knowledge about the data via automatic reasoning (e.g., [14]). The rules are implications between an antecedent and a consequent, that is, whenever the antecedent holds then the consequent must hold as well. For example, if the free-text annotation includes the comment *RA off-axis* (antecedent) then the comment is necessarily referred to the *HLA view* (consequent). Analogously, the free-text comment *basal slice is missing* (antecedent) implies a lack of coverage associated to the *SA view* (consequent). One of the key points of carefully developing such rules in our work is that they can be used to infer numerical quality scores from the comments. For example, the presence of *Lack of coverage* (antecedent) will always lead to a sub-optimal quality score associated to the right and left ventricle (consequent). In addition, we aim to use rules to reveal potential incompleteness or ambiguity. For example, we can classify a comment such as *LA off axis* as incomplete, because neither the imaging view or the cardiac cycle phase are indicated.

3 Preliminary results

The free-text annotations provide a rich source of information for target users of UK Biobank that goes beyond the simple classification provided by the numerical scores. However, free-text is prone to variability and ambiguity which hinders the efficient use of its information content for querying and access of Cine MRI scans according to the quality assessment outcomes. In this section we provide two examples to motivate the need for a formal structuring of quality assessment by means of the design and implementation of a semantic layer for the annotations.

Table 1. Variability for the *LA off axis* example in numbers. The synonyms with asterisk take into account variations (abbreviations and most common typos) associated to them.

Synonyms	**Occurrence**
foreshortened*	2
off axis*	179
off-axis	13
off axis	145
off axes	21
off plane	0
out of axis	26
wrong plane	7

Table 2. Ambiguity in the annotation of HLA (4Ch) when it is not explicitly explained if either left or right atrium or both atria are affected by the off axis technical issue (highlighted in bold in the table).

Off axis subset	**Occurrence**
Total off axis sentences	515
Only referring to LA	173
Only referring to RA	27
Reference to both atria	53
HLA (4Ch) without reference	**226**
Other annotations	36

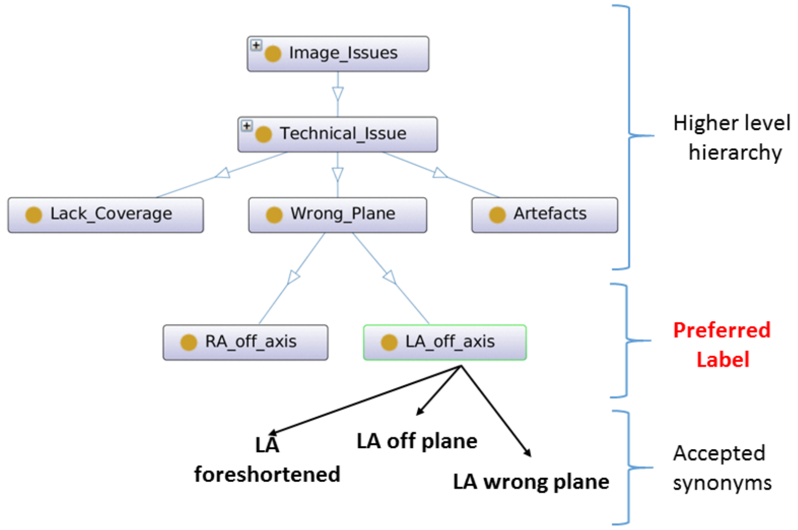


Fig. 3. Example fragment of the *CMR-QA* ontology showing the scope of the concept *LA off axis* and its synonyms.

3.1 Example 1: variability

Variability is due to natural human variability, for example there is difference in the used terminology or different opinions about the quality of the image. This variability can be limited only in part by the use of a standardised analysis protocol. For example, different observers can correctly flag as *LA off axis*, *LA out of plane*, *wrong LA plane*, *2Ch out of plane*, or *LA foreshortened* an image where the plane chosen to acquire a long axis view was not optimally aligned to measure left atrial (LA) volumes. Table 1 shows the occurrence of the set of synonyms used by the observers for the case of *LA off axis* in a total of 214 comments referring to wrong image plane affecting left atrium. In the development of the semantic layer with respect to this specific example, we made a first decision to define as preferred label *LA off axis*, because it is the most commonly used. Then we have defined as accepted synonyms all the other variants in the table. Figure 3 shows a fragment of *CMR-QA* with the preferred label and a subset of the accepted synonyms of the concept *LA off axis*.

3.2 Example 2: ambiguity

Ambiguity is typically due to missing information in the annotations which may lead to an ambiguous interpretation. For example, within the UK Biobank pilot study analysis protocol, the annotation *2Ch out of plane* will always refer to a left atrium that is off-axis since in VLA (2Ch) only the left atrium is contoured. However, *4Ch out of plane* would represent an ambiguous annotation since in HLA (4Ch) both left and the right atria are contoured. Table 2 describes the occurrence of ambiguous HLA (4Ch) related comments, with respect to missing reference about which atrium is affected. On a total of 515 sentences generally relating to off axis issues (could be for any imaging view), 226 address HLA (4Ch) without the explicit reference to LA, RA or atria. Detailed knowledge of the UK Biobank pilot CMR dataset helps addressing such ambiguity, as the observers came to realise that in most of the cases where *4Ch out of plane* applies, the left atrium is most affected. Hence, we have created project-specific ontology rules that *(i)* raise a warning of incompleteness for the comment, and *(ii)* assign *LA off axis* (consequent) as the most likely outcome to such comments where a *wrong plane* is observed in the *HLA view* (antecedent).

4 Discussion

In this paper we have presented the first steps in defining a semantic layer within the quality assessment of UK Biobank Cardiac Cine MRI scans with the main aim of automatically inferring image quality scores from free-text annotations. In this section we comment on the motivating examples shown in the previous section, summarise the future work and emphasise the likely benefits for UK Biobank users and clinical research.

Motivating examples. In Section 3 we have presented two cases in the evaluation of the quality assessment of image analysis where the benefit of introducing a semantic layer can be appreciated. The first example on the lexical variability naturally encountered in free-text comments demonstrates that adding a semantic layer to complement the free-text annotations would be extremely useful to extract commonalities. In particular, the semantic layer will allow us to *(i)* group annotations that describe the same concept under the same preferred label, and *(ii)* find analogies among similar annotations. The second example is about ambiguity due to incompleteness in the comments. The use of semantics will help find patterns in the annotation procedure and define ontology rules to partially resolve such ambiguity. In the context of these two examples we have also shown how the structured vocabulary associated to the quality assessment information is being designed. Preliminary intuition of the development of domain-specific ontology rules has also been provided.

Related Work. Ontologies are extensively used in biomedicine. Prominent examples are BioPortal, a comprehensive repository containing more than 500 biomedical ontologies [6], and SNOMED CT [15], the reference ontology of choice across National Health Service (NHS) information systems.[3] There have also been recent efforts in adding a semantic layer to describe the information within a biobank. Andrade et al. [8] envisaged the benefits of using ontologies for querying and searching the information in a biobank and across biobanks. Muller et al. [11] presents and updated overview of the state of the art and open challenges for the description and interoperability across biobanks where the use of Semantic Web technologies will play a key role. Examples of concrete Semantic Web-based solutions in biobanks can also be found in [9, 10]. Although state-of-the-art ontologies include the description of concepts relevant to our

---

[3] `http://systems.hscic.gov.uk/data/uktc/snomed`

domain, we could not find any ontology meeting all our requirements (e.g., complete description of Cine MRI technical issues) which evidences the necessity of a more specific ontology in this particular domain.

Future Work. As immediate future work, we plan to complete the *CMR-QA* ontology, define the necessary ontology rules and finalise the implementation of the techniques to text mine the comments to extract ontology data. In this way it will be possible to automatically infer numerical quality scores from the annotations. Validation will be carried out by comparing the automatic scores with those manually assigned by the observers as part of their quality assessment. We also aim to design a prototype software for query and retrieval of Cine MRI scans according to certain *semantic* characteristics to be later embedded in the UK Biobank CMR pilot study available tools. Furthermore, we will perform an extensive evaluation to analyse the correctness of our approach. In the long-term, the definition of semantics will enable integration at different levels within UK Biobank and with external vocabularies: *(i)* integration with other parts of UK Biobank where different ontologies and controlled vocabularies may be used; *(ii)* integration with other existing biobank ontologies (e.g., [9, 10]); and *(iii)* integration with medical vocabularies in order to be compliant with state-of-the-art standards, for example SNOMED CT [15]. The integration with other ontologies will allow the interoperability among different group of experts relying on different ontologies and the creation of a broader semantic layer.

Expected benefits for training purposes. An important related application is the development of training material for future observers analysing clinical imaging data. *CMR-QA* will provide a controlled set of preferred quality assessment comments, together with alternative expressions, so that observers can be more systematically trained in image quality assessment. This will result in a significant reduction of variability and improved quality of the information content of the comments provided.

Expected benefits of semantic access for UK Biobank users. The use of a controlled vocabulary provided by the *CMR-QA* ontology will ease the retrieval of Cine MRI scans according to their quality and reliability of the analysis outcomes. For example, a clinical researcher interested in building a control group for a study on pathologies affecting the atria, might want to query only for those images whose atrial volumes have been reliably estimated. Therefore, he/she will exclude scans where atria were off axis or the image quality was sub-optimal. A different example coming from the biomedical engineering world: an expert in development of algorithms for automatic contouring of the left ventricle might be interested in testing a newly developed tool on those Cine MRI scans where it was most difficult to define the basal slice. He/she will specifically query for cases whose image analysis was annotated as having difficult definition of the basal slice for the left ventricle.

List of abbreviations

CMR: Cardiovascular Magnetic Resonance imaging **CMR-QA:** Cardiovascular Magnetic Resonance Quality Assessment **HLA:** Horizontal Long Axis **LA:** Left Atrium

LV: Left Ventricle **MRI:** Magnetic Resonance Imaging **NER:** Named Entity Recognition **NHS:** UK's National Health Service **RA:** Right Atrium **RV:** Right Ventricle **SA:** Short Axis **VLA:** Vertical Long Axis.

Acknowledgements

SEP, SN and SP acknowledge the British Heart Foundation (BHF) for funding the manual analysis to create a cardiovascular magnetic resonance imaging reference standard for the UK Biobank imaging resource in 5,000 CMR scans (PG/14/89/31194, PI Petersen, 6/2015 to 5/2018). SKP, VC and SN were additionally funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre based at The Oxford University Hospitals Trust at the University of Oxford. EJR and IH were funded by the European Commission under FP7 Grant Agreement 318338, "Optique", and the EPSRC projects Score!, ED3 and DBOnto.

References

1. Petersen, S.E., et al.: Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK Biobank - rationale, challenges and approaches. Journal of Cardiovascular Magnetic Resonance **15**(1) (2013) 1–10 (http://www.ukbiobank.ac.uk/).
2. Petersen, S.E., et al.: UK biobank's cardiovascular magnetic resonance protocol. Journal of cardiovascular magnetic resonance **18**(1) (2016) 8+
3. Schulz-Menger, J., et al.: Standardized image interpretation and post processing in cardiovascular magnetic resonance: Society for Cardiovascular Magnetic Resonance (SCMR) board of trustees task force on standardized post processing. Journal of Cardiovascular Magnetic Resonance **15**(1) (2013) 1–19
4. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing? Int. J. Hum.-Comput. Stud. **43**(5-6) (1995) 907–928
5. Guarino, N., Oberle, D., Staab, S.: What is an ontology? In: Handbook on Ontologies. Springer (2009) 1–17
6. Fridman Noy, N., et al.: BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Research **37**(Web-Server-Issue) (2009)
7. Giese, M., et al.: Optique: Zooming in on big data. IEEE Computer **48**(3) (2015) 60–67
8. Andrade, A.Q., Kreuzthaler, M., Hastings, J., Krestyaninova, M., Schulz, S.: Requirements for semantic biobanks. Stud Health Technol Inform. **180** (2012) 569–573
9. Pathak, J., et al.: Applying semantic web technologies for phenome-wide scan using an electronic health record linked biobank. J. Biomedical Semantics **3** (2012) 10
10. Brochhausen, M., et al.: Developing a semantically rich ontology for the biobank-administration domain. J. Biomedical Semantics **4** (2013) 23
11. Müller, H., et al.: State-of-the-art and future challenges in the integration of biobank catalogues. In: Smart Health - Open Problems and Future Challenges. (2015) 261–273
12. Spasic, I., Ananiadou, S., McNaught, J., Kumar, A.: Text mining and ontologies in biomedicine: Making sense of raw text. Briefings in Bioinformatics **6**(3) (2005) 239–251
13. Bodenreider, O.: Lexical, terminological and ontological resources for biological text mining. In: Text mining for biology and biomedicine. Artech House (2006) 43–66
14. Nenov, Y., Piro, R., Motik, B., Horrocks, I., Wu, Z., Banerjee, J.: RDFox: A highly-scalable RDF store. In: 14th International Semantic Web Conference (ISWC). (2015) 3–20
15. Schulz, S., Cornet, R., Spackman, K.A.: Consolidating SNOMED CT's ontological commitment. Applied Ontology **6**(1) (2011) 1–11