



ISO/IEC JTC 1/SC 29/WG 1
(ITU-T SG16)

Coding of Still Pictures

JBIG

Joint Bi-level Image
Experts Group

JPEG

Joint Photographic
Experts Group

TITLE: Call for Technologies for a JPEG ontology for still image description (JPOnto)

SOURCE: JPEG Metadata Subgroup

Ruben Tous (Universitat Politècnica de Catalunya, UPC) rtous@ac.upc.edu

Jaime Delgado (Universitat Politècnica de Catalunya, UPC) jaime.delgado@ac.upc.edu

Kyoungro Yoon (Konkuk University) yoonk@konkuk.ac.kr

PROJECT: JPSearch

STATUS:

REQUEST

ACTIONS:

DISTRIBUTION: WG1 Distribution List

1 Background and rationale

The Joint Photographic Experts Group (JPEG) is a working group of ISO/IEC, the International Organization for Standardization / International Electrotechnical Commission, (ISO/IEC JTC1/SC29/WG1) and of the International Telecommunication Union (ITU-T SG16), responsible for the popular JPEG, JBIG, JPEG-LS, and JPEG 2000 family of imaging standards. The JPEG committee developed the baseline JPEG Standard (ISO/IEC 10918-1 | ITU-T T.81) nearly 20 years ago, and this has been one of the most successful international standards ever produced and the dominant standard for imagery on the Internet, and in digital cameras.

Among the latest series of standards from the JPEG committee, ISO/IEC 24800 (Still Image Search, known as JPSearch) is a project that aims to provide a standard for interoperability for image search and retrieval systems. Recently, within the context of JPSearch, the WG1 group has initiated a new activity on future generation image metadata. The main goal of this activity is to provide a simple and uniform way of annotating JPEG images with metadata compliant to the Linked Data principles. Linked Data is a method of publishing metadata so that they can be interlinked among them, to public Linked Data databases (e.g. DBpedia) and to formal vocabularies, i.e. ontologies. An ontology is a specification of what terms to use for a certain domain, and how these terms are defined. Modern ontologies have transitioned from shared natural-language definitions to shared machine-processable representations, usually under the shared formal model provided by RDF.

The WG1 group has also launched a systematic review and consolidation of its file formats and codestream syntax with a view to consolidating the benefits of the JPEG system of standards and enabling new functionalities. Ontology technologies are expected to play an important role in this effort.

This document is a call for contributions of an ontology for still image description. Eligible candidates must enable the description of digital still images with technical, administrative and semantic metadata compliant with the RDF standard and the principles of Linked Data. WG1 is interested in ontologies that help circumventing the current proliferation of vocabularies for image description by providing a core set of terms and relationships, thus facilitating the cross-community data integration of information related to digital images.

The call puts an special focus on ontologies providing a rich set of constructs for semantic visual content description, including, but not limited to, a uniform description of identities, features, aspect, relationships, actions and emotional information of people appearing in the images, as well as description of events, locations and objects.

Candidate ontologies shall fulfill the requirements contained in this document (Annex 1)

2 Requested action

The individuals or organizations with expertise in multimedia metadata engineering, computer vision, ontology engineering and/or knowledge representation are encouraged to provide inputs to the WG1 committee. Proposals into this effort (including partial proposals and reports of experimental investigations) are strongly encouraged and will be evaluated and formalized through the work of WG1. With this purpose, interested parties and stakeholders are also encouraged to participate in the 60th JPEG (ISO/IEC JTC1 SC29/WG1 meeting in Incheon, Korea (22nd - 26th April 2013).

Proposals should be sent by 22nd March 2013 by email to Prof. Dr. Kyoungro Yoon (yoonk@konkuk.ac.kr) and Dr. Ruben Tous (rtous@ac.upc.edu).

Participants who wish to attend this meeting are requested to contact the WG1 Convener, their National Standards Body, or an organization in liaison with WG1. Further information about WG1 and this proposal can be obtained from Dr. Ruben Tous (rtous@ac.upc.edu) or Prof. Dr. Kyoungro Yoon (yoonk@konkuk.ac.kr).

Annex 1

This annex presents a set of requirements for the JPEG ontology for still image description (JPonto).

1 Use a common representation language

The ontology shall be represented with RDF and/or other Linked Data related technologies (e.g. RDFS, OWL, etc.).

2 Use a common textual syntax

It is recommended to represent the ontology using the W3C's Turtle language.

3 Consider classes and properties already defined in first draft JPonto-core (see Annex 2)

It is recommended that the ontology takes into account core classes and properties already defined within the first draft JPonto-core ontology (Annex 2).

4 Being able to describe the image data characteristics

The ontology shall provide resources (properties, classes, etc.) for describing the technical details of the digital representation of the image. For instance, it shall be possible to describe the image width, image resolution, etc.

5 Being able to describe details about the image creation process

The ontology shall provide resources for describing the people, organizations, dates, locations, tools and processes involved in the different creation steps (capturing, editing, etc.).

5.1 Being able to describe, in various ways, the date and time in which the image was taken

The ontology shall provide resources for describing when the image was taken in an unambiguous way (e.g. UTC time). The ontology shall also provide resources for describing the time with nominal values such as "night", "day", "winter", "the 80s", etc.

5.2 Being able to describe, in various ways, the location from which the image was taken

The ontology shall provide resources for describing the geographic location from which the image was taken in various forms. It shall be possible to specify the longitude and latitude but also nominal values of static locations such as "San Jose" or dynamic locations such as "from a car".

6 Being able to describe regions of interest

It shall be possible to relate metadata to regions in the image.

7 Being able to describe absolute and relative positions of things in the shot (at the bottom, X near Y, X at the background, ...)

8 Being able to describe the quality of the image (in focus, selective focus, out of focus, motion blur, noisy/blocky, etc.)

9 Being able to describe the style of the image (picture-in-picture, circular warp, gray-color, overlay, etc.)

10 Being able to describe the view of the image (portrait, close-up/macro, indoor, outdoor, etc.)

11 Being able to describe details about the image history (awards, involvement in forensic processes, etc.)

12 Being able to describe images within the image

As images can include other images (paintings, book covers, sculptures, wallpapers, etc.) the ontology shall provide resources for describing the images within the image with the same detail as the image itself.

13 Being able to describe the identities and relationships of people appearing in the image

14 Being able to describe the actions and emotional information of people appearing in the image

15 Being able to describe the type of people appearing in the image (gender, age, race, profession, etc.)

16 Being able to describe the aspect of people appearing in the image (make up, clothes, etc.)

17 Being able to describe some common events and types of events depicted in the image (e.g. Halloween, etc.)

18 Being able to describe locations of interest and types of locations of interest appearing in the image (e.g. a church, the Eiffel tower, etc.)

19 Being able to describe certain objects and object types appearing in the image (car, tree, etc.)

19.1 Being able to describe celestial bodies appearing in the image (sun, moon, stars, etc.)

19.2 Being able to describe water-related concepts appearing in the image (underwater, sea/ocean, lake, river/stream, other, etc.)

19.3 Being able to describe flora appearing in the image (tree, plant, flower, grass, etc.)

19.4 Being able to describe fauna appearing in the image (cat, dog, horse, fish, bird, insect, spider, amphibian/reptile, rodent, etc.)

19.5 Being able to describe combustion phenomena appearing in the image (fire, smoke, fireworks, etc.)

20 Being able to describe the atmospheric conditions (snow, rain, etc.)

21 Being able to describe lighting effects (shadow, reflection, silhouette, lens effects, etc.)

22 Being able to describe the type of the image (city life, party life, home life, sports/recreation, food/drink, etc.)

23 Being able to describe the impression produced by the image (euphoria, happiness, fun, melancholy, fear, calm, etc.)

24 Supporting the provenance information of metadata properties

The ontology shall support provenance information of metadata properties. Metadata is being dealt with by for example producers of metadata (e.g. a photo camera), changers (e.g. a person which modifies initially created metadata) and consumers (e.g. an application which processes metadata to make it accessible for search). If several pieces of metadata, created by machines or people in different roles, are in conflict (e.g. contradictory creation dates), a description of provenance (i.e. roles of the metadata creators) can be useful for conflict resolution (e.g. "metadata produced by the changer has precedence over metadata produced by the creator").

25 Providing support for policy and protection information

The ontology shall provide support for linking policy information related to the image. Specific types of policy information are license, rights and access. If an implementation supports policy information, the set of properties must include a link to policy description, represented using e.g. ODRL, in order to express the binding of the policy information to the media resource being described.

Annex 2

WG1 envisages that the JPEG ontology for still image description (JPOno) will be subdivided into different sub-vocabularies covering different aspects of a digital still image description. One of these sub-vocabularies will be a core ontology, named JPOno-core, acting as a central component to interconnect all the other sub-vocabularies. This annex presents the current status of an initial work on JPOno-core that has already started within WG1. It is recommended that candidate ontologies take into account the classes and properties already defined within the JPOno-core ontology.

Outline

A core vocabulary named JPOno-core is the basic set of entities and properties which serves as a central component to interconnect all the other sub-vocabularies of the the JPSearch Metadata Ontology (JPOno). Figure 1 shows the visual representation of the core vocabulary.

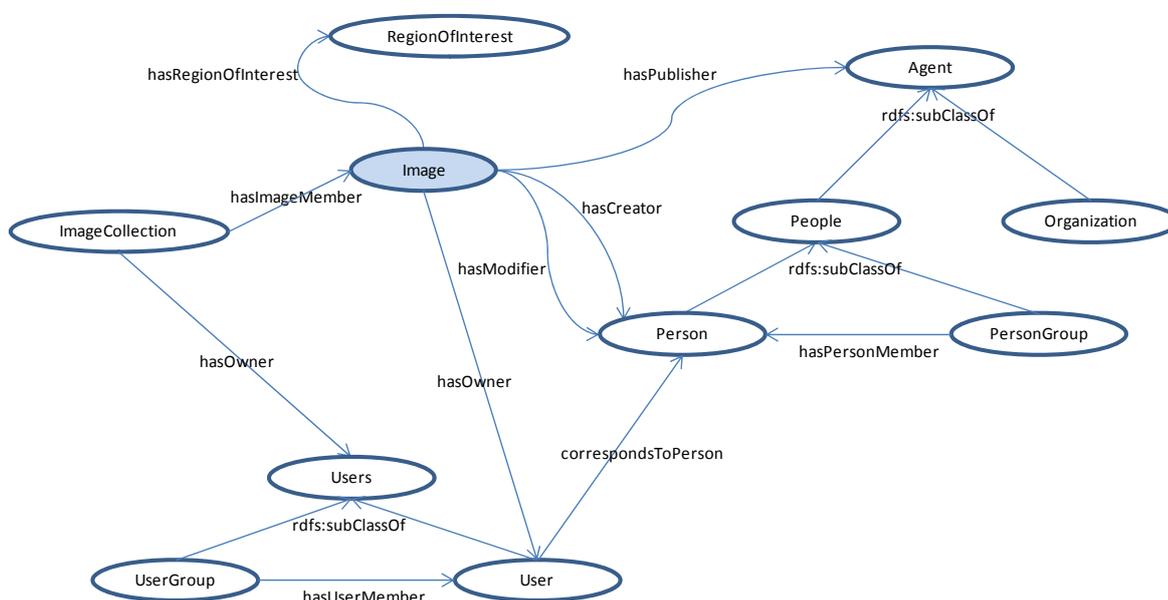


Figure 1. Outline of JPOno-core

Semantics

Class definitions:

Name	Definition
Image	Represents a digital still image.
RegionOfInterest	Describes the information (e.g., content description, keywords, etc.) of a certain region within the image. Note, the whole image itself can also be described by one RegionOfInterest.
Agent	An agent (e.g. person, group, software or physical artifact).
Person	A human being.
People	One or more persons.
PersonGroup	2 or more persons with something in common.
Organization	An organization.
User	User of the system containing the image (e.g. photo-sharing service).
Users	One or more users.
UserGroup	2 or more users with something in common.

Properties with domain *Image*:

Name	Range	Definition
identifier	xsd:anyURI	Describes an identifier of the image in the form of a URI. The identifier must be unique.
hasModifier	Person	Describes who changed the original image resulting in the creation of the image.
hasCreator	Person	Describes who created the image or made contributions in the creation of the image.
hasPublisher	Agent	Describes information about the publishing people or organization of the image
creationDate	xsd:dateTime	Describes the date when the image is created.
modifiedDate	xsd:dateTime	Describes the date when the image is modified.
description	xsd:string	Specifies the content of the image in the form of text.
hasRightsDescription	Not defined	Describes the right related information.
hasStringRightsDescription	xsd:string	Sub-property of hasRightsDescription.
keyword	xsd:string	Describes a list of keywords that characterize the image (optional).
title	xsd:string	Describes the title of the image (optional).
collectionLabel	xsd:string	Describes user provided labels that can be used for the purpose of collection and categorization of images (optional).
preferenceValue	xsd:decimal	Describes the value of the preference of the image in the form of integer value.
ratingValue	xsd:decimal	Describes the rating results that should be one of the corresponding controlled terms. The definition of the terms is provided by JPSearch.
originalImageIdentifier	xsd:anyURI	Describes the identifier of the original image from which the image is created. Moreover, it can be used for the identifiers that can be created by a particular organization or method.
locationAltitude	xsd:decimal	Describes the GPS altitude of the place shown in the image.
locationLatitude	xsd:decimal	Describes the GPS latitude of the place shown in the image.
locationLongitude	xsd:decimal	Describes the GPS longitude of the place shown in the image.
hasRegionOfInterest	RegionOfInterest	Describes the information (e.g., content description, keywords, etc.) of a certain region within the image. Note, the whole image itself can also be described by one RegionOfInterest.
width	xsd:integer	Pixel width of the image.
height	xsd:integer	Pixel height of the image.