Semantic annotation of videos with structured data and knowledge organization in the collaborative web environment.

Denise Gomes Silva Morais Cavalcante

Programa de Pós-Graduação em Ciência da Informação/ECA/USP

INTRODUCTION

The research addresses **semantic annotation** with **structured data** and **moving image indexing** in the context of the **semantic web**. Audiovisual information is considered unstructured data because it does not have a predefined standard data model and requires treatment to be processed. Furthermore, audiovisual metadata for indexing moving images is mostly limited to descriptive metadata and often has little information about the concepts and subjects represented in the image content. Annotation tools generate structured data for regions of interest within the image and aggregate linked concept data from various linked open data (LOD) vocabularies.

MATERIALS AND METHODS

BIBLIOGRAPHIC RESEARCH

Deep image annotation (DIA) is becoming an accepted new concept in treating images for semantic enrichment, aiming to improve the efficiency of image retrieval and access, enhance users' understanding of an image, and support automatic image resources integration and knowledge Discovery (ZENG, 2019, p. 24).

Image annotation with the process and methods of Knowledge Organization (KO) (CLARKE, 2015):

✓ Image-Level Metadata

✓ Annotation-Level Metadata

The **Web Annotation Data Model** specification describes a structured model and format to enable annotations to be shared and reused across different hardware and software platforms (SANDERSON; CICCARESE; YOUNG, 2017) .

Source: (SANDERSON; CICCARESE; YOUNG, 2017)
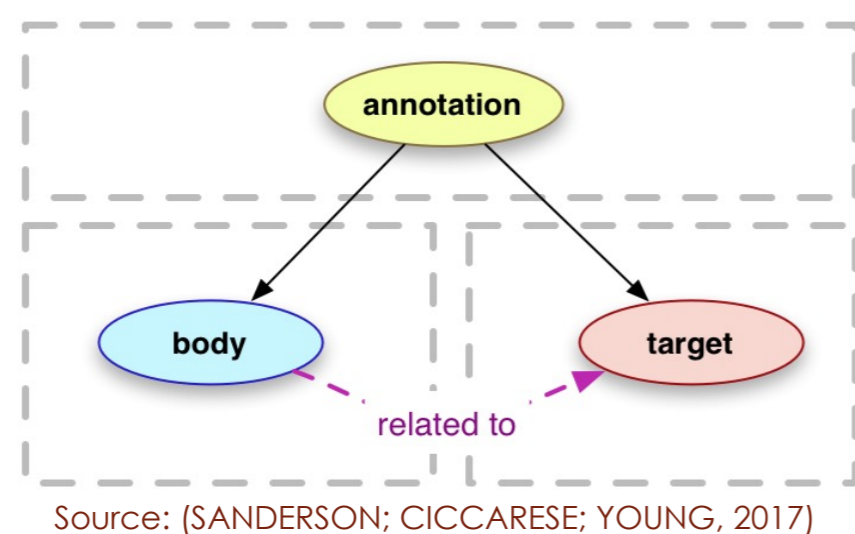
APPLIED RESEARCH

The research structured a **concept schema modeled in SKOS**, for semantic annotation of the subjects present in the moving images. The following steps are in progress:

1. Categorizing and defining the concepts.

2. Vocabulary construction using the SKOS data model:

a) Definition of the relationships and SKOS properties of each concept.

b) Entity Reconciliation and connect the concepts with value and property LOD vocabularies (DBpedia, Wikidata, AAT, etc).

c) Adding labels in other languages.

d) SKOS modeling with turtle (N3) RDF serialization format.

3. Visualizing the vocabulary in a web open-source tool and publishing the project repository with the concept schema on gitpages.

REFERENCES

CALDERA-SERRANO, J.; CARO-CASTRO, C. **Limitaciones y problemas de la lectura de imagen en movimiento**. [s.l: s.n.]. Acesso em: 21 out. 2020.

CASTRO, C. C. **Vocabularios estructurados, Web Semántica y Linked Data: oportunidades y retos para los profesionales de la documentación** 1. II Seminário de Estudos da Informação: Arquivologia, Biblioteconomia e Ciência de Informação: Identidades, Contrastes e Perspectivas de Interlocução: Niterói, Universidade Federal Fluminense, 26- 28 setembro, 2011.

CLARKE, D. **Deep Image Annotation: Making a Difference in Knowledge Organization**. n. July, p. 1–21, 2015.

SANDERSON, R.; CICCARESE, P.; YOUNG, B. **Web Annotation Data Model. W3C**, n. February, p. 1–56, 2017.

ZENG, M. L. **Semantic enrichment for enhancing LAM data and supporting digital humanities**. Review article. Profesional de la Informacion, v. 28, n. 1, p. 1–36, 2019.

RESULTS
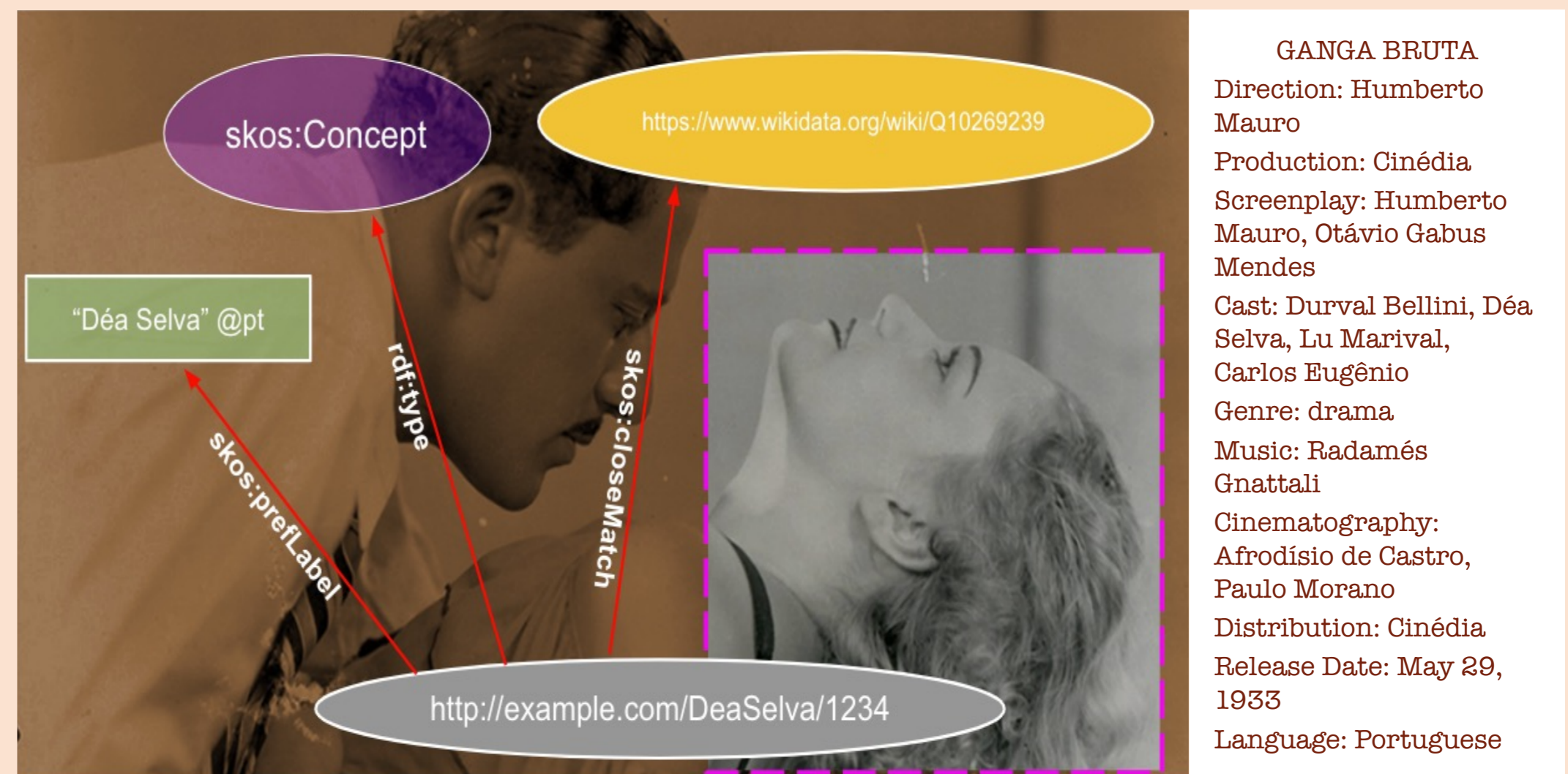
GANGA BRUTA
Direction: Humberto Mauro
Production: Cinédia
Screenplay: Humberto Mauro, Otávio Gabus Mendes
Cast: Durval Bellini, Déa Selva, Lu Marival, Carlos Eugênio
Genre: drama
Music: Radamés Gnattali
Cinematography: Afrodísio de Castro, Paulo Morano
Distribution: Cinédia
Release Date: May 29, 1933
Language: Portuguese

Wikidata item page (Q10269239)
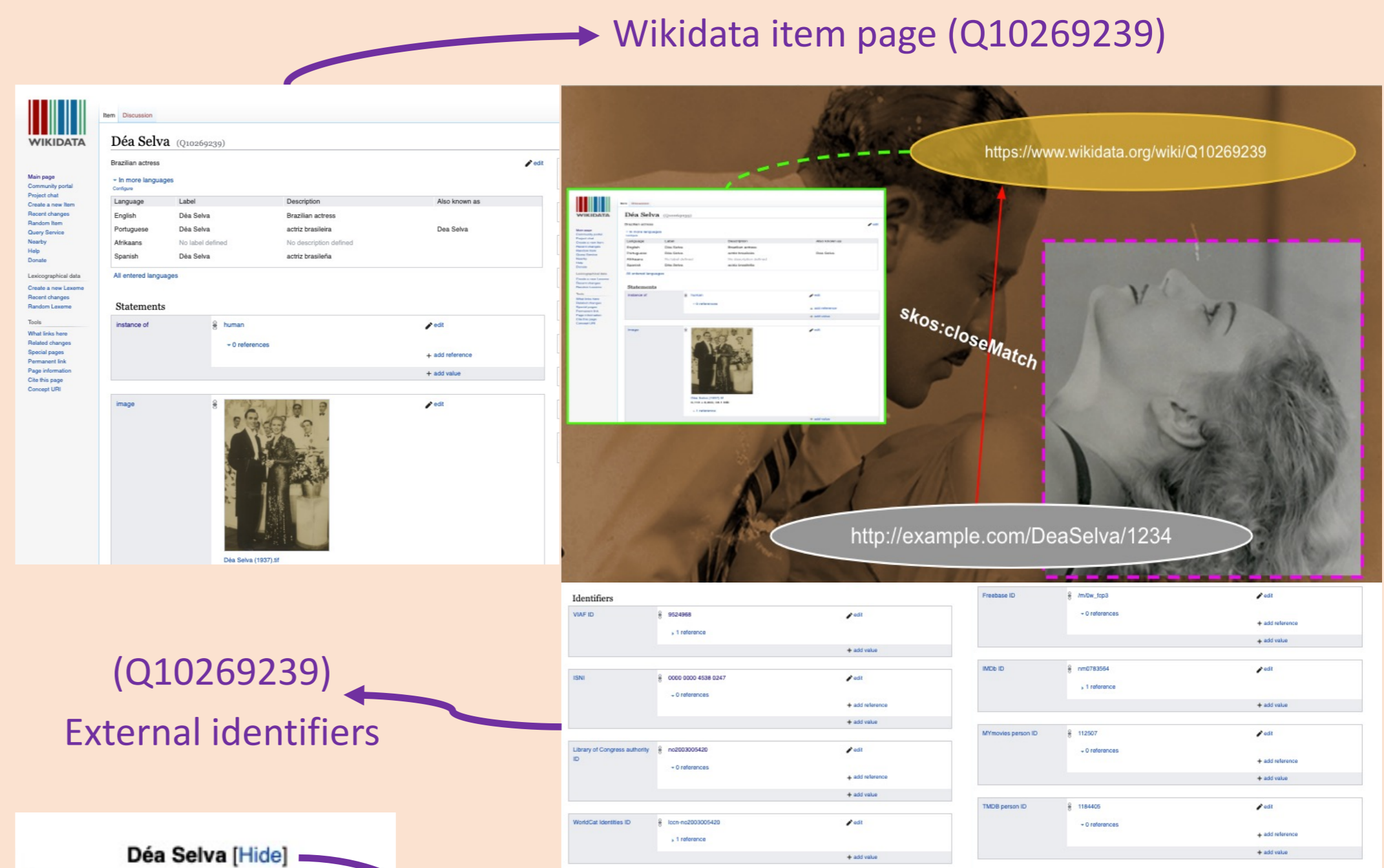
(Q10269239)
External identifiers

Wikimedia Commons Déa Selva

EXAMPLES OF AUTHORITY CONTROL FILE PAGES
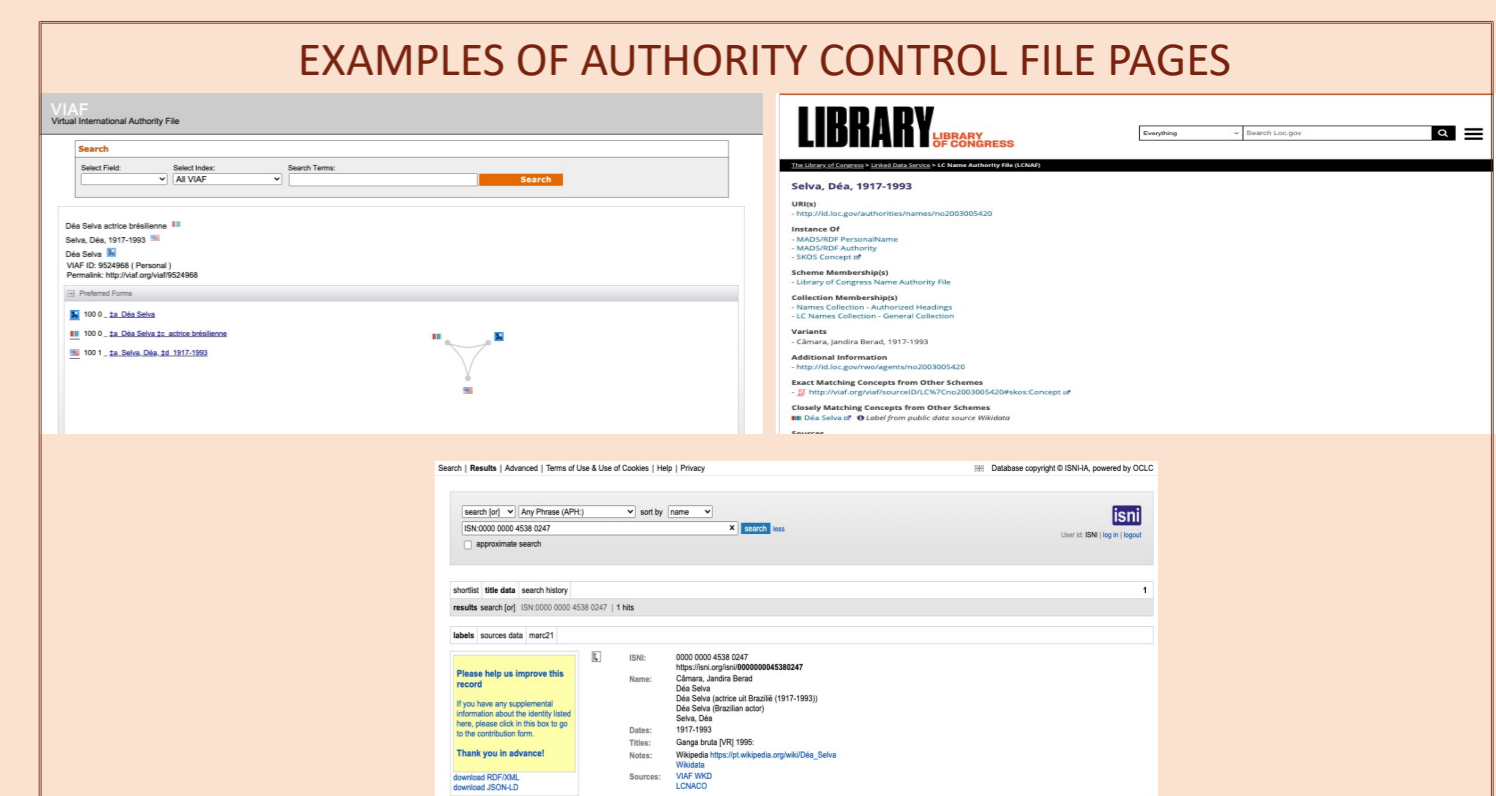
- ISNI 0000 0000 4538 0247
- VIAF ID 9524968
- Library of Congress authority ID no2003005420
- WIKIDATA Item Q10269239

CONCLUSIONS

Semantic enrichment based on image annotation can be enhanced by using KOS to structure machine-readable data from moving image resources and for cultural heritage knowledge organization, and collaboration between humans and machines is required, with humans making manual annotations to assist in semantic enrichment of the information.

CONTACT

denisegsmc@usp.br