

Revista Facultad de Ingeniería

Journal Homepage: <https://revistas.uptc.edu.co/index.php/ingenieria>



Digital Image Representation Model Enriched with Semantic Web Technologies: Visual and Non-Visual Information

Sandra-Milena Roa-Martínez¹

Caio Coneglian²

Silvana Vidotti³

Received: August 9, 2022

Accepted: January 17, 2023

Published: January 31, 2023

Citation: S.-M. Roa-Martínez, C. Coneglian, S. Vidotti, "Digital Image Representation Model Enriched with Semantic Web Technologies: Visual and Non-Visual Information," *Revista Facultad de Ingeniería*, vol. 32 (63), e14815, 2023. <https://doi.org/10.19053/01211129.v32.n63.2023.14815>

Abstract

The types of content of digital images, visual (syntactic and semantic) and non-visual, cause the complexity of their representation. Considering these contents separately hinders digital image retrieval because this creates a gap between the contents of the image and its representation. Therefore, this work aims to present a

¹ Ph. D. Universidad del Cauca (Popayán-Cauca, Colombia). smroa@unicauca.edu.co. ORCID: [0000-0002-2271-6101](https://orcid.org/0000-0002-2271-6101)

² Ph. D. Universidade de Marília (Marília-São Paulo, Brazil). ORCID: [0000-0002-6126-9113](https://orcid.org/0000-0002-6126-9113)

³ Ph. D. Universidade Estadual Paulista (Marília-São Paulo, Brazil). silvana.vidotti@unesp.br. ORCID: [0000-0002-4216-0374](https://orcid.org/0000-0002-4216-0374)



representation model of visual and non-visual information of digital images, with semantic enrichment through the Semantic Web technologies. For that, a qualitative methodology with a bibliographical approach was used. Theoretical subsidies of the topics addressed were sought, and it has an applied focus since it proposes a model and its exemplification. The developed model depicts the representation image process and allows the semantic enrichment of the data. This enrichment facilitates the retrieval in multiple contexts with technologies that favor the use of the data through inferences. Also, a use case with digital medical images is presented, demonstrating the feasibility of the proposal. It is concluded that the representation of visual and non-visual content aims to improve the way images are retrieved in digital information environments. The junction of the content and the context of images should be considered, even though search mechanisms usually treat this separately due to the disaggregation of image representation itself.

Keywords: digital image; image enrichment; image representation; visual and non-visual content; web semantic technologies.

Modelo de representación de imágenes digitales enriquecida con tecnologías de la web semántica: Información visual y no visual

Resumen

La complejidad de la representación de las imágenes digitales está dada por los tipos de contenido —visual (sintáctico y semántico) y no visual— que estas poseen. El considerar estos contenidos separadamente dificulta la recuperación de imágenes digitales porque crea una brecha entre los contenidos de la imagen y su representación. Por lo tanto, este trabajo tiene como objetivo presentar un modelo de representación de información visual y no visual de imágenes digitales, con enriquecimiento semántico mediante el uso de las tecnologías de la Web Semántica. Para ello fue utilizada una metodología cualitativa con enfoque bibliográfico y realizado un caso de aplicación. Se buscaron subsidios teóricos de los temas abordados, además consta de un enfoque aplicado ya que se propuso un modelo y su ejemplificación. El modelo propuesto describe el proceso de representación de la imagen y permite el enriquecimiento semántico de los datos.

Este enriquecimiento facilita la recuperación en múltiples contextos con tecnologías que favorecen el uso de los datos por medio de inferencias. Asimismo, se presenta un caso de aplicación con imágenes médicas digitales, en el que se demuestra la factibilidad de la propuesta. Se concluye que la representación de contenidos visuales y no visuales tiene como objetivo mejorar la forma en que se recuperan las imágenes en ambientes de información digital. La fusión del contenido y el contexto de las imágenes debe ser considerada, aunque los mecanismos de búsqueda suelen tratarlos por separado debido a la desagregación de la representación de la imagen en sí.

Palabras clave: contenido visual y no visual; enriquecimiento de imágenes; imagen digital; representación de imágenes; tecnologías de la web semántica.

Modelo de representação de imagem digital enriquecido com tecnologias da web semântica: informação visual e não visual

Resumo

A complexidade da representação das imagens digitais é dada pelos tipos de conteúdo —visuais (sintáticos e semânticos) e não visuais— que possuem. Considerar esses conteúdos separadamente dificulta a recuperação da imagem digital, pois cria uma lacuna entre o conteúdo da imagem e sua representação. Portanto, este trabalho tem como objetivo apresentar um modelo de representação de informações visuais e não visuais de imagens digitais, com enriquecimento semântico por meio do uso de tecnologias da Web Semântica. Para isso, foi utilizada uma metodologia qualitativa com abordagem bibliográfica e realizado um caso de aplicação. Buscou-se subsídios teóricos dos temas abordados, além disso consiste em uma abordagem aplicada uma vez que foi proposto um modelo e sua exemplificação. O modelo proposto descreve o processo de representação de imagens e permite o enriquecimento semântico dos dados. Esse enriquecimento facilita a recuperação em múltiplos contextos com tecnologias que favorecem o uso de dados por meio de inferências. Da mesma forma, é apresentado um caso de aplicação com imagens médicas digitais, no qual é demonstrada a viabilidade da proposta. Conclui-se que a representação de conteúdos visuais e não visuais visa

melhorar a forma como as imagens são recuperadas em ambientes de informação digital. A fusão de conteúdo e contexto da imagem deve ser considerada, embora os mecanismos de busca muitas vezes os tratem separadamente devido à desagregação da própria representação da imagem.

Palavras-chave: conteúdo visual e não visual; enriquecimento de imagem; imagem digital; representação de imagens; tecnologias da web semântica.

I. INTRODUCTION

The high volume of digital information available and generated by the wide range of digital devices increases the difficulty of finding information resources within a digital information environment. A type of resource with distinctive characteristics within these environments is digital images since they are considered an information source. However, in many cases, image management is not efficient due to the amount of information an image contains. This leads to inefficient results during the retrieval and search processes.

Although there are initiatives to improve image retrieval, there is a gap in the way images are represented. On the one hand, images are represented by associating them with a set of keywords. This set has a subjective character since the description of the content of the images is based on words or texts. On the other hand, it is possible to represent this content through an automatic extraction related to its visual features. In this respect, a complete representation needs to consider aspects of visual content—obtained by automatic extraction and keywords—and non-visual information that is part of an image, such as author data, modification date and format. In this way, image retrieval systems can be improved thanks to a representation that contemplates the content and context of the images.

Based on the formal analysis models of Information Science contents, it is possible to conceive the representation of images covering the different elements mentioned (visual and non-visual information). Based on such models, the semantic representation can be enriched using Semantic Web technologies (SWT), such as using ontologies to structure the data and the connection to databases in the context of the Data Web. Semantic enrichment improves the interoperability of the information made available on the web since it uses established standards to represent the images. Besides, this process supports the construction of retrieval systems that make inferences about the available data. SWT enable the represented data to be understood by computational mechanisms, expanding the capacity to meet the informational needs of the users during the search process, especially in the context of images.

Given the above, this work aims to present a model of representation of visual and non-visual information of digital images, with semantic enrichment through the use of SWT. For this, we used a qualitative methodology with an exploratory bibliographical character, searching for theoretical references of the themes in the literature. Furthermore, the formal analysis models of Information Science content and SWT were reviewed. Also, it has an applied character since we proposed a model and an exemplification of the application of the model to medical digital images. Among the results obtained, it was possible to determine that non-visual information is mainly associated with metadata patterns. In contrast, visual information (syntactic and semantic) can be linked using SWT, such as domain ontologies and Linked Data, which expands the data sources of the content and the context of the images.

The proposed model, together with an application case and the queries made, allow validation of the efficiency and feasibility of the representation and retrieval of digital images by adding a higher level of formal semantics, which will allow the construction of systems based on inferences, with relationships and data enrichment from the information of the images.

Finally, this paper is structured as follows, firstly, the theoretical questions that involve the representation of digital images and how these resources are inserted in the context of the Semantic Web. Then, the results and discussions are presented, introducing and discussing the developed model. The following section focuses on the review of an application case, and finally, there is a section concluding the article.

II. BACKGROUND

A. Digital Images Representation

The process of representation of the information of an image requires the reading of elements in different "layers" or levels according to their meaning and purpose, as it will be described afterward.

The main paradigms for image representation are concept-based and content-based. In the representation by concept or meaning, images are described by words or terms of indexation. These words or terms usually have a meaning. Moreover, existing models of description in Information Science can be used for this task. In most cases,

indexing by terms involves a manual description of the image from the elements analyzed and using a template to select the concepts representing the image. This indexation contains a load of subjectivity due to human participation in the process, besides being highly time-consuming when dealing with large amounts of images.

In reference to the Content-Based Paradigm, indexing is automatic and executed through computational techniques of Digital Image Processing. Indexing is also mainly associated with features such as color, texture, and shape, i.e. at the syntactic level. This syntactic indexation will make the semantic retrieval of images difficult. On the other hand, a large volume of images can be indexed in a short time using this paradigm, and the subjectivity of manual indexing is diminished.

However, the representation of images by concept can be divided into “representation of visual content” and “representation of non-visual content”. The first is related to a correspondence between the elements present in the image (iconographic and intrinsic) and verbal or textual descriptors (words or terms). Non-visual content representation refers to information extrinsic to the image, i.e., not in the image but related to it. This information may be useful during the retrieval process.

Theoretical models for formal analysis of images within Information Science consider types of content and information of digital images. Among them, the first one to think about is the Iconographic Model of Panofsky [1]. This model determines three levels of description: of existing objects in the image from the visual primitives such as color, texture, and shape, without requiring interpretation; themes and concepts next to logical and derived characteristics that need interpretation; and intrinsic meaning of the image, with a high level of semantic inference and knowledge by whom analyses the image or studies the context.

Another revised model was Shatford's [2], which also has three levels of description equivalent to Panofsky's model and includes four facets at each level: who, what, where, and when, and additionally, other attributes with external or non-visual information classified as biographical, exemplified, and related. Moreover, Jörgensen [3] presents a model with three types of attributes in images: perceptive, interpretative, and reactive. Jaimes and Chang's model [4] contains ten levels of description of visual attributes classified into two categories: syntactic-perceptual and

semantic-conceptual, as well as non-visual or externality of the image related to the context and historical aspects of the image of the biographical type. Furthermore, Barreto [5] classifies the visual content as primitive, referring to the visual features, and complex, perceived by humans as sources of meaning. Additionally, it includes content-independent metadata.

The review and presentation of the models above demonstrate that many of them incorporate not only the representation of syntactic and semantic content, but also the representation of non-visual or extrinsic information of the image. For instance, the image format, author, date, location or lighting conditions. We will denominate this information as non-visual information of the image.

Semantic Web technologies can improve the process of image processing and representation. In this context, the next section presents a conceptualization of the Semantic Web, its technologies, and its relation with digital images.

B. Digital Images in the Semantic Web Context

The previous section concludes that the representation of a digital image must consider syntactic and semantic visual content and non-visual information. Representing a digital image from these elements will facilitate the retrieval of images since searching among the high volume of data that circulates through different websites is not an easy task.

In this regard, it is necessary to support the processes of image representation through technologies such as those coming from the Semantic Web. It is worth noting that “The Semantic Web sphere is endowed with tools that allow machines, in the form of computational agents, to identify concepts and process information and possible meanings evoked by the relationships between concepts in the data available on the Internet” [6].

For the Semantic Web, the textual representation of the information contained in a visual document is of interest since the description of the knowledge of the images will be explicit. Therefore, machines understand this knowledge, making it interoperable and available in different formats. Thereby, it solves the lack of syntactic

and semantic interoperability during the image annotation process that results from the different formats and applications used to carry out this process.

According to Splendiani [7], Semantic Web technologies are an effective way of expressing the semantics of an image through a systematic representation and precise definition of the information behind the images. SWT provide elements that allow representing the syntax and semantics of a resource, as well as creating, storing, manipulating, exchanging, and processing non-visual information from an image. Additionally, a wide variety of application areas use these technologies, seeking to be useful when the semantically annotated resource is a static image because: “to textually represent the information conveyed by an image is interpreted as the problem of making the implicit knowledge captured in images more easily processable by machines and, consequently, available in others forms.” [7].

For [8], markup languages capable of representing knowledge from metadata and ontologies are required. These markup languages will explore the semantic web. RDF-based markup languages should contain greater expressiveness and reasoning ability to represent the knowledge and semantics that ontologies include.

In this scenario, the Semantic Web enables the semantic enrichment of information resources. The enrichment can be understood as the process that performs searches in different data sets and ontologies, among other elements, and allows the enrichment of the description of the resources by including a URI that identifies concepts. This URI is inserted into each entity to identify, contextualize, and link with other related entities. There are several concepts and technologies of the Semantic Web used for the semantic enrichment of informational resources. This way, one can establish the purpose and relation of these Semantic Web technologies and concepts. It is worth emphasizing the ontologies as the concept that represents knowledge and semantics —let alone the potentiation of the Semantic Web— when used along with other elements to enrich digital images addressed in this work.

According to [9], the use of ontologies in the area of image retrieval mainly aims to: Unify the description of low-level features; represent the different types of relationships between image characteristics through Ontologies of visual description; model concepts (objects) and their relationships using Knowledge description; help

match the visual level with the semantic level using Semantic mapping. When analyzing these objectives, it was found that the ontologies contribute to an enrichment of the description of the images. Furthermore, the use of ontologies also leads to better retrieval because in the image analysis, the semantic and syntactic concepts will be represented with entities and relations among other elements provided by an ontology. Thereby, a correspondence between the syntactic and the semantic is sought.

Additionally, some ontologies were analyzed: Domain ontology for image semantic retrieval [10] and Fuzzy ontology for image retrieval [11], etc., which reinforce that they present themselves as tools that provide an explicit semantic description and alternative representation of the image. They also reduce the subjectivity of the manual indexing process carried out on images when one wants to describe their semantic content.

III. RESULTS

In this section, we present the results and analyses made from the literature review on content analysis models of images for the representation of visual documents and the inclusion of Semantic Web technologies in the representation enrichment.

Digital Imaging is essential for all subsequent processes in the search for ways to retrieve images with a more significant number of arguments. Regarding the Semantic Web, the Representation of Information helps insert characteristics and metadata capable of inserting higher formal semantics. However, when handling digital images, it is necessary to understand that an image can be represented from two perspectives. The first is the data about the visual content, and the second expresses the non-visual content. Both perspectives have a necessary role in the correct and expressive representation of an image, fundamental in the context of the Semantic Web, mainly for gathering information that seeks to consider everything an image expresses. In this sense, it is necessary to apply different processes for visual and non-visual data to represent the images using the SWT with precision.

Non-visual information is, in short, the metadata that describes digital images. There is a set of metadata standards built with this intent, such as the Digital Imaging and

Communications in Medicine (DICOM) standard, which helps to describe visual medical resources. Although these metadata are fundamental in describing an image presenting its non-visual information, they do not initially follow the Semantic Web guidelines. Therefore, it is necessary to use metadata standards based on semantic technologies or convert data to such technologies.

When analyzing the visual information of an image—for instance, information about the colors and the predominant forms in it—it is verified that the integration with the principles of the Semantic Web becomes more complex due to the difficulty of obtaining and endowing visual information with semantics. These issues hinder computer comprehension of visual information.

The insertion of visual information aggregates metadata that represents image non-visual information. Thus, the representation enlarges and gets a more significant number of data on the image. However, there is an obstacle in the union and the relation of this non-visual information with the visual one. Thereby, it is necessary to carry out processes that capture and structure visual data into metadata. [12] discuss this process, presenting a conceptual framework capable of structuring the visual and context content (non-visual information) of medical digital images. There is a possibility of extrapolating the process to other types of images since the representation of syntactic visual content is provided by feature extraction techniques that apply to all kinds of digital images. Besides, for the representation of some non-visual information, the authors used the DICOM medical scope metadata standard. This standard may be replaced by other standards for describing informational or specific resources following the digital images of interest.

From the union and relation of the visual and non-visual information through a conceptual framework, the next step is to represent the information of an image using Semantic Web technologies.

By structuring information following metadata standards, it is possible to convert this information into RDF structures, a Semantic Web technology capable of representing data with more arguments and more significantly for computers.

For this purpose, firstly, it is necessary for image metadata, with visual and non-visual information, to be converted into machine-readable formats, such as CSV. From this

conversion, it is possible to structure the data in RDF. The structuring process must follow the Semantic Web principles, maintaining the features and image representation. The conversion must occur using ontologies as a basis. Ontologies define the structure in which the data are organized, related, and interconnected. The process includes the definition of vocabularies that will be used to express the properties and links of the resources. Such a process will seek to convert metadata records, which initially maintains a key-value structure, into a more complex structure. In this structure, properties have a more significant semantic value because of the use of vocabularies recognized by the communities and which will be composed entirely of RDF triples. In these RDF triples, resources are linked to other resources through properties and may contain axioms and other characteristics.

Using ontologies to help structure the data composes one of the enrichment processes because of the insertion of RDF semantic elements from the ontology in the description. Thus, after structuring for RDF, one obtains a description of the visual and non-visual features of an image in a structure compatible with the principles of the Semantic Web. However, possible enhancement of this structure in RDF considers interconnecting the information described with other databases. The process of interconnecting with other datasets composes the second stage of enrichment since it will allow data in a domain to relate to other structures consolidated and recognized in specific contexts, such as bases for authority records such as VIAF, or databases of a general field, such as DBpedia. In short, the process will take a simple structure in RDF, and insert other elements that are related to the described resources, such as the creator of an image. The creator may have a registry of authorities, making it possible to connect this information—initially disconnected from the resource information—with a dataset structured according to SWT.

In this work, we will highlight the enrichment process, since the objective is primarily to make a dataset in RDF more significant for machines, especially with the use of domain ontologies, which insert information not initially considered in the converted RDF. It is worth mentioning that the enrichment process can occur with several tools and technologies, having the function of inserting external elements into the data

represented by the RDF. For instance, there are technologies such as ontologies or the relation with databases following the principles of Linked Data. [13] and [14] revise and discuss the process of RDF metadata conversion and enrichment. This process aims to create, from metadata, a database structured in RDF, with the insertion of other related properties, to increase the enhancement of that dataset.

Finally, after the structuring and enrichment process, the last step is the publication of the data. This process will provide a dataset with the data in RDF format, using database systems for structured data, such as Virtuoso. This system offers a SPARQL Endpoint interface, which allows both human and non-human users to look up published data. Figure 1 synthesizes the processes of syntactic representation, semantic and non-visual representation in the context of the Semantic Web and illustrates the different stages of both processes.

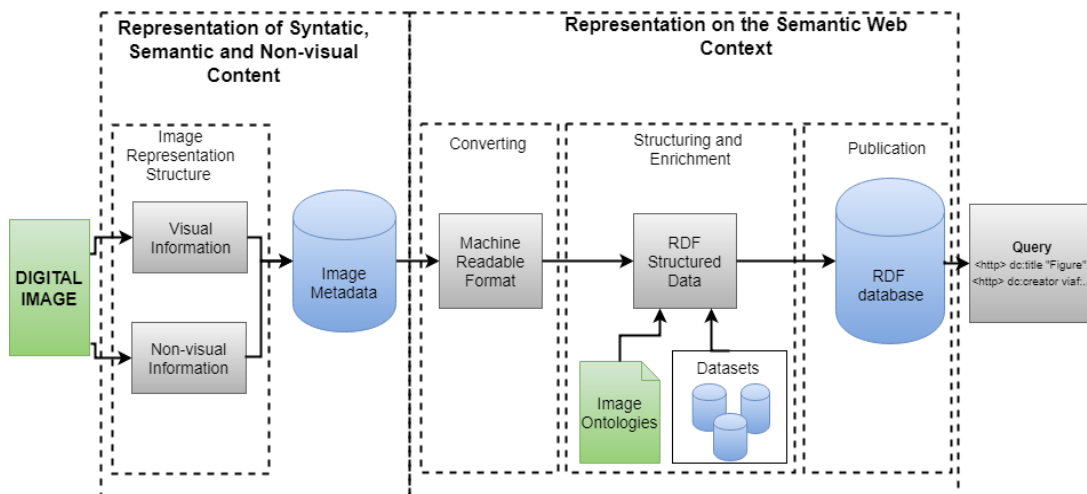


Fig. 1. Digital imaging model enriched with Semantic Web Technologies.

Figure 1 shows how the representation of an image, with both visual and non-visual metadata, can be published in datasets following the principles of Semantic Web and Data Web, performing the enrichment process. For this purpose, it is first necessary to find a way to perform a syntactic, semantic, and non-visual representation that considers both visual content and non-visual information as the conceptual framework for medical images proposed by [12]. This framework demonstrates how this can occur in a conceptual representation of a medical image.

IV. APPLICATION CASE: RDF REPRESENTATION OF A DIGITAL IMAGE

This section presents an example using Semantic Web technologies to represent a digital image. The aim is to demonstrate the application of the previously described model, which allows structuring the data after the process of syntactic, semantic, and non-visual representation, structuring and enrichment.

It is worth mentioning that the presented model applies to any digital image with a representation of both visual and non-visual content. Given that each domain presents its specificities, an analysis is needed to determine elements such as useful domain ontologies, the datasets that can be linked, and the metadata patterns that should be used to describe the features. Thus, when applying the model with the details of each domain, it will be possible to have a representation using SWT and contemplate the visual and non-visual content of the image.

The example presented in this case used medical images as a domain, having specifications to consider. Therefore, the built case used an Annotation and Image Markup (AIM) domain ontology, which presents characteristics related to medical images, such as diseases, algorithms applicable to image treatment, and medical and technical annotations, among other data. Furthermore, expressive datasets were searched in the general scope to relate to the contributors and the creators of the image and in specific bases to link to information that touches on medical issues, such as disease databases. Also, general metadata standards were used to describe non-visual information of images, which are technical metadata such as dates and image creators.

Based on the reported assumptions, the following example shows the representation in RDF of a medical digital image used for disease diagnosis. Algorithms were used to extract some visual information from the image. Thus, the example shown in Figure 2 was created from the ontologies, the metadata patterns, and the datasets, presenting a medical digital image represented with its visual and non-visual information. Figure 2 represents the RDF generated in this example with graphs. To represent the visual (syntactic and semantic) content of a medical image (in the lower center of the figure), an OWL ontology provided by the AIM project was used to

represent entities associated with medical images. This ontology includes low-level features contained in the images and the algorithms for obtaining them (green-labelled syntactic visual content), observations such as lesions pointed out by health professionals from the images (blue-labelled semantic visual content), regions viewed in the images, as well as other metadata (in orange).

In this context, the representation presented in Figure 2 relates AIM ontology entities, elements of the Dublin Core metadata standard, and the datasets used to link to external databases. Consequently, it is possible to represent any medical images following these configurations. The same figure shows, for the reference image, the connection with Linked Data datasets such as DBpedia, VIAF, and GeoNames, as well as the use of the Dublin Core metadata standard for non-visual information representation. When thinking of the Data Web, the related databases allow an improvement in the way the content is represented. This improvement is due to the use of elements from internationally recognized datasets, which improve semantic enrichment, complementing the representation performed in the ontologies.

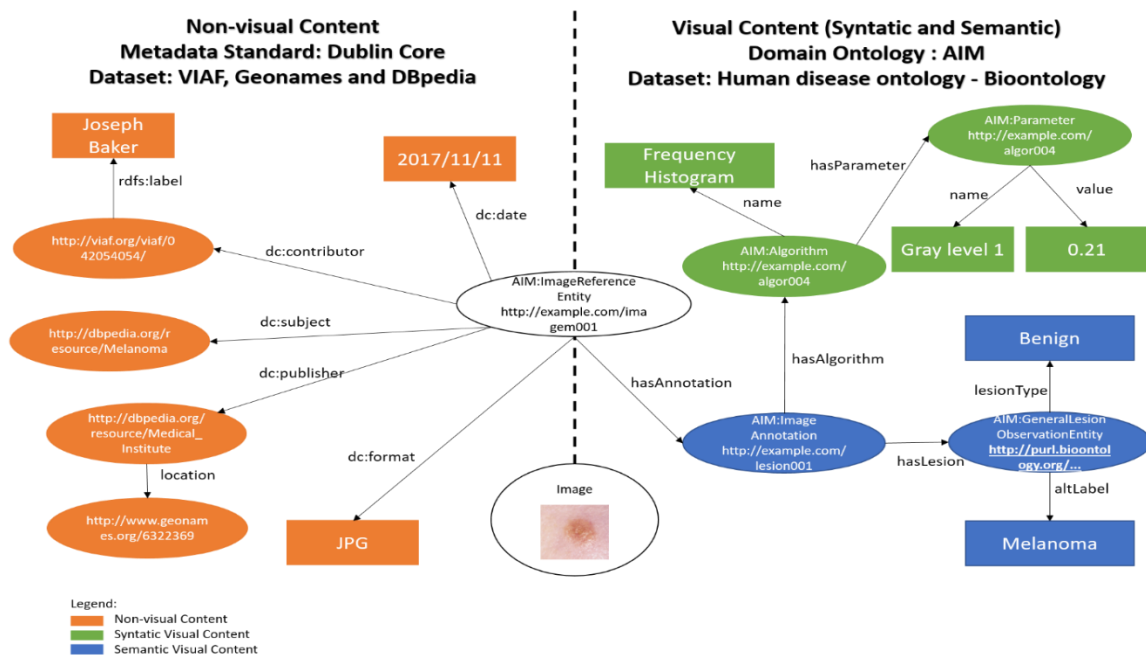


Fig. 2. RDF Example of the proposed representation.

Regarding the specific domain that the image refers to, in the case of medical images, it is necessary to use databases of this domain, such as Human Disease Ontology - bio-ontology. This vocabulary gathers medical information, such as human diseases, and is an essential instrument for referencing such elements in the medical context. Figure 3 shows a sample query in SPARQL and its respective result. This query is a simulation of data that could be queried and retrieved if the model and structure presented were implemented.

The query in Figure 3 presents several relationships made when a professional wishes to obtain a specific type of information. This query is viable through SPARQL and the built RDF structure. In this case, the retrieval refers to images of one day (2021/11/12) with a specific disease (Melanoma) and having a Gray Level 1 parameter of the Frequency Histogram above 0.5. Therefore, the results present only the information of the images, the respective responsible physician, and the value for the parameter meeting the set conditions.

Query	Results																							
<pre> SELECT ?image ?cod_doctor ?val as "Gray Level 1" WHERE { ?image rdfs:type AIM:ImageReferenceEntity; dc:date "2017/11/11"; :hasAnnotation ?annotation; dc:contributor ?cod_doctor. ?annotation :hasLesion ?lesion; :hasAlgorithm ?algorithm. ?lesion :altLabel "Melanoma". ?algorithm :name "Frequency Histogram"; :hasParameter ?parameter. ?parameter :name "Gray level 1"; :value ?val. FILTER(? val >0.5) } </pre>	<table border="1"> <thead> <tr> <th>Image</th> <th>Cod_doctor</th> <th>Gray Level 1</th> </tr> </thead> <tbody> <tr> <td>http://image/0005</td> <td>http://doctor/0005</td> <td>0.6</td> </tr> <tr> <td>http://image/0020</td> <td>http://doctor/0085</td> <td>0.75</td> </tr> <tr> <td>http://image/0301</td> <td>http://doctor/0005</td> <td>0.51</td> </tr> <tr> <td>http://image/0407</td> <td>http://doctor/0050</td> <td>0.64</td> </tr> <tr> <td>http://image/2148</td> <td>http://doctor/0400</td> <td>0.75</td> </tr> <tr> <td>http://image/5210</td> <td>http://doctor/1000</td> <td>0.9</td> </tr> </tbody> </table>	Image	Cod_doctor	Gray Level 1	http://image/0005	http://doctor/0005	0.6	http://image/0020	http://doctor/0085	0.75	http://image/0301	http://doctor/0005	0.51	http://image/0407	http://doctor/0050	0.64	http://image/2148	http://doctor/0400	0.75	http://image/5210	http://doctor/1000	0.9		
Image	Cod_doctor	Gray Level 1																						
http://image/0005	http://doctor/0005	0.6																						
http://image/0020	http://doctor/0085	0.75																						
http://image/0301	http://doctor/0005	0.51																						
http://image/0407	http://doctor/0050	0.64																						
http://image/2148	http://doctor/0400	0.75																						
http://image/5210	http://doctor/1000	0.9																						

Fig. 3. Query in RDF databases of Medical Images.

Through this query, for instance, a physician can identify patients who need specific care from the algorithms applied to the images and represented in the built data model. In this way, the use of SPARQL and environments that use Semantic Web technologies allow inferences and the establishment of relationships in a more natural way. This would be unlikely when dealing with relational and traditional databases.

V. DISCUSSION

In the performed representation model, the first step deals with converting records into machine-readable formats, such as CSV. In this context, several converters can perform this process according to how the database supplies data. Converting to CSV is fundamental for the data structuring step in RDF because the systems that perform this process use machine-readable formats to structure the data. This structuring must be performed using ontologies as a basis to determine how properties and links should be placed and performed. In the context of the images, there is an extensive set of ontologies which can aggregate information about the semantics of an image, contributing to later computational agents being able to extract meaning from the RDF models.

Thus, when inserting ontologies as elements to enrich the datasets, it is necessary to think of ontologies with descriptive and semantic elements defined in RDF format. These ontologies determine the meaning of any element of a query (image or word) according to its context in the ontology. Therefore, the intelligent agents and autonomous programs to trace the web could locate the contents related to a search term associated with the meaning and the other interpreted concepts.

Likewise, the use of ontologies determines new spaces for searches and more effective retrieval of visual documents without limiting itself to the semantic content and covering the syntactic components. Based on these findings, the ontologies are artefacts useful for the semantic enrichment process after converting the metadata into RDF. The use of ontologies is favored since the compatibility between the RDF, the converted metadata, the OWL, and the RDF Schema of the ontologies approximates and facilitates the communication between the data. Moreover, each ontology can enrich the data within a distinct perspective, encompassing visual and non-visual information.

Another enrichment process is the possibility of interconnecting the contents with other databases. In this context, many datasets are used for this purpose, depending

on the context of the connection. Table 1 presents some examples of datasets for making these connections.

Table 1. Examples of datasets for connecting to data.

Databases	Domain	Connection example
VIAF	Authority file	Creators of the images
DBpedia	General domain	Topics of the images
Library of Congress Subject Headings	Controlled vocabulary	Bind folksonomy terms and used terms to describe topics
Geonames	Geolocalization	Use to signal the places referred to in an image
YAGO	General domain	Topics of the images

The bases presented are inserted in the context of Linked Data and Data Web, using the principles of the Semantic Web. In this way, the creation of relations can occur naturally, with the use of RDF, relating the resources of these bases with the resources of the images. It is worth mentioning that one can use more bases apart from the ones inserted here. Therefore, depending on the image's domain, one can add other bases. For example, there is a wide range of available and ready-to-use datasets for the health domain.

The presented case tried to demonstrate the feasibility of the model proposed in Figure 1, using a digital image of the medical scope as an example. In this case, it was possible to determine that non-visual information is mainly associated with metadata patterns, whereas visual information (syntactic and semantic) relates to domain ontologies. Another highlight is the use of Linked Data datasets, verifying the possibility of making relationships with other data sources, content, and the context of the images. Furthermore, to validate the last step of the model exposed in Figure 1, referring to queries in the data published in RDF, it was demonstrated how a SPARQL query works on the presented data. The SPARQL query follows the data structure introduced in Figure 2, using ontologies and datasets that were the basis for the structure.

Therefore, the publication of data must follow the principles of the Web of Data. There is documentation on recommendations to complete this task, such as the W3C document that presents the Good Practices for the Publication of Data on the web [15]. Furthermore, the publication should occur in tools that use RDF as a base, such

as the Virtuoso database (<https://virtuoso.openlinksw.com/rdbms-engine/>), which makes a SPARQL Endpoint interface available, as previously reported.

Finally, the entire model, structure, and queries presented at the end allow digital images to be represented and retrieved more efficiently, with a higher level of formal semantics. Moreover, this model enables the construction of systems using inferences and treating the information of the images with different relationships and data enrichment.

VI. CONCLUSIONS

Digital images are visual documents with a high content of information and can have their contents analyzed by formal models of Information Science, considering their peculiarities in comparison with other classes of existing documents. The evolution and the proposals of new models in the Computer Science and Information Science have in common the different levels of description of the images represented following Panofsky's proposal. It is essential to recognize that these levels are complementary.

The representation of an image is complex because of the consideration of different levels of content, summarized in syntactic and semantic content. The first is associated mainly with low-level visual features such as color, texture, shape, etc. The second is related to concepts determined by the individual interpretation of each and the knowledge created after this interpretation. Additionally, the representation of non-visual or context information from digital images is essential.

Given the above, a digital image representation model was proposed, where the representation of content (syntactic and semantic) and non-visual information (context) is carried out through metadata based on conceptual structures and metadata patterns. The model consists of a second element, which promotes the semantic enrichment of the data through Semantic Web technologies, such as ontologies and semantic links. In the end, the model constitutes a database of digital images information following the Web of Data principles.

It is worth highlighting that, among the SWT, the ontologies analyzed and used in this work can help in the process of enriching the image representation since ontologies

provide an explicit semantic description of the content of a document (text, image, etc.) that is machine-processable and then easily retrieved by users.

The model is applied in a use case to represent a digital image of the medical scope to demonstrate the feasibility of the proposal. In this case, it was possible to determine that non-visual information is mainly associated with metadata patterns, whereas visual information (syntactic and semantic) relates to domain ontologies. Another highlight is the use of Linked Data datasets, in which it was possible to verify the possibility of making relationships with other data sources, content, and the context of the images. Besides, queries using SPARQL allow the retrieval of digital images to occur following the principles of the Semantic Web by promoting inferences and relationships with other bases.

Therefore, the conception of this model contributes to the development of research on the representation of digital images in diverse contexts. Moreover, such a representation aims to improve image retrieval within digital information systems, considering the combination of the content and the context presented in images, and generally regarded as separate information in search engines due to the disaggregation of image representations.

ACKNOWLEDGMENTS

Research Group on Computational Intelligence (GICO) of the Systems Department of the Universidad del Cauca (Colombia), Research Group Novas Tecnologias em Informação (GPNTI) of the Universidade Estadual Paulista Júlio de Mesquita Filho (Brazil) and Universidade de Marília (Brazil).

AUTHORS' CONTRIBUTION

Sandra-Milena Roa-Martínez: Conception, Data collection, Analysis, Writing-original draft.

Caio Coneglian: Conception, Analysis, Writing-review and editing.

Silvana Vidotti: Conception, Writing-review and editing, Supervision.

REFERENCES

- [1] E. Kalkanis, *An Analysis of Erwin Panofsky's Meaning in the Visual Arts*, Macat Library, 2018.
- [2] S. Shatford, "Analyzing the subject of a picture: a theoretical approach," *Cataloging & classification quarterly*, vol. 6, no. 3, pp. 39–62, 1986. https://doi.org/10.1300/j104v06n03_04
- [3] C. Jörgensen, "Indexing images: Testing an image description template," in *ASIS Annual Conference Proceedings*, 1996, pp. 1–8.
- [4] A. Jaimes, S. Chang, "Conceptual framework for indexing visual information at multiple levels," in *International Society for Optics and Photonics*, 1999, pp. 2–15.
- [5] J. Barreto, "Desafios e avanços na recuperação automática da informação audiovisual," *Ciência da Informação*, vol. 36, no. 3, pp. 17–28, 2007. <https://doi.org/10.1590/s0100-19652007000300003>
- [6] F. R. E. Nunes, B. C. M. Maculan, M. B. Almeida, "Os fundamentos da Web Semântica como ferramentas de auxílio para as demandas da Sociedade da Informação," *Em Questão*, vol. 26, no. 3, pp. 224-249, 2020. <https://doi.org/10.19132/1808-5245263.224-249>
- [7] B. Splendiani, *A proposal for the inclusion of accessibility criteria in the authoring workflow of images for scientific articles*, Doctoral Thesis, Universidad de Barcelona, Barcelona, Spain, 2015.
- [8] M. Pritoni et al. "Metadata Schemas and Ontologies for Building Energy Applications: A Critical Review and Use Case Analysis," *Energies*, vol. 14, no. 7, pp. 2024-2030, 2021. <https://doi.org/10.3390/en14072024>
- [9] H. Bannour, C. Hudelot, "Towards ontologies for image interpretation and annotation," in *IEEE*, 2011, pp. 211–216. <https://doi.org/10.1109/cbmi.2011.5972547>
- [10] U. Manzoor et al. "Semantic Image Retrieval: An Ontology Based Approach," *International Journal of Advanced Research in Artificial Intelligence*, vol. 1, no. 4, pp. 1–8, 2015.
- [11] M. Liaqat, S. Khan, M. Majid, "Image retrieval based on fuzzy ontology," *Multimedia Tools and Applications*, vol. 76, pp. 22623-22645, 2017.
- [12] S. M. Roa-Martínez, S. Vidotti, M. Jorente, "Representación conceptual de imágenes médicas digitales: Integración de Contexto y Contenido Visual," *Revista General de Información y Documentación*, vol. 26, no. 2, pp. 651-72, 2016. <https://doi.org/10.5209/rigid.54719>
- [13] K. Munir, M. S. Anjum, "The use of ontologies for effective knowledge modelling and information retrieval," *Applied Computing and Informatics*, vol. 14, no. 2, pp. 116-126, 2018. <https://doi.org/10.1016/j.aci.2017.07.003>
- [14] A. Tello et al. "Rdf-ization of dicom medical images towards linked health data cloud," in *VI Latin American Congress on Biomedical Engineering*, 2015, pp. 757-760. https://doi.org/10.1007/978-3-319-13117-7_193
- [15] B. F. Lóscio, C. Burle, N. Calegari, *Data on the Web Best Practices*, W3C, 2017.