



Promoción del desarrollo de SW libre en un entorno de calidad y confianza adaptando las metodologías, procesos , modelos de negocio y últimas tecnologías

TSI-020301-2008-22

Entregable D 7.3
Especificación de la ontología de descripción de la edición y gestión de documentos

Fundación CTIC
Andago
Instituto Tecnológico de Informática
iSOCO

Fecha límite del entregable: 28/02/2009

Fecha de entrega: 30/03/2009

Este trabajo se licencia bajo Creative Commons Attribution-Share Alike 3.0.
Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-sa/3.0/> o envíe una carta a Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Este trabajo está parcialmente financiado por el Ministerio de Industria, Turismo y Comercio español.

Historial de Cambios

Versión	Fecha	Estado	Autor (Partner)	Descripción
0.1	09/10/2008	Borrador	CTIC	Creación del documento
1.0	30/03/2009	Entregable	CTIC	Versión entregable

RESUMEN EJECUTIVO

En este informe se presenta la ontología DOCOM, que formaliza la composición de documentos y basada en la metáfora de huecos-elementos de relleno. Como complemento a la ontología, se ofrece un ejemplo representativo de la composición de documentos.

Este documento es una versión impresa de la documentación presente en la página web <http://idi.fundacionctic.org/docom>. Todas las actualizaciones posteriores estarán reflejadas en la versión online.

Información del Documento

Proyecto FIT Número	TSI-020301-2008-22	Acrónimo	Vulcano
Título completo	Promoción del desarrollo de SW libre en un entorno de calidad y confianza adaptando las metodologías, procesos, modelos de negocio y últimas tecnologías		
URL	http://www.ines.org.es/vulcano		
URL del documento	http://idi.fundacionctic.org/docom		

Entregable	Número	D 7.3	Título	Especificación de la ontología de descripción de la edición y gestión de documentos.
Paquete de Trabajo	Número	7	Título	Gestión documental colaborativa
Tarea	Número	T 7.3	Título	Ontología para describir la edición y gestión colaborativa de documentos

Fecha de Entrega	Contractual		Entregado	
Estado	Versión 1.0, 30/03/2009		final	
Tipo	Informe			
Nivel de Diseminación	Público <input checked="" type="checkbox"/> Consorcio <input type="checkbox"/>			
Resumen (para diseminación)	Formalización mediante un lenguaje de ontologías de un modelo capaz de representar la información implicada en la edición y gestión colaborativa de un documento, incluyendo de forma no exclusiva, detalles sobre la autoría, el control de cambios y versiones, las dependencias entre documentos, el reparto de tareas, el origen e historial del documento, el vínculo con plantillas de documentos, etc			
Palabras Clave	Ontología, documentación, semántica, composición			

Autores (Partner)	CTIC (editor), Andago, iSOCO, ITI			
Responsable de Autoría	Carlos Tejo Alonso	Email	carlos.tejo@fundacionctic.org	
	Partner	CTIC	Tfno	+34 984 29 12 12

DOCOM Ontology Specification



An ontology to formalize document
composition

Working Draft — 10 December 2008

This version:

<http://idi.fundacionctic.org/docom/spec-20081210/>

Latest version:

<http://idi.fundacionctic.org/docom/spec>

Last Update:

Date: 10 December 2008

Revision

0.9

Editors:

Carlos Tejo Alonso, ([Fundación CTIC](#))

Emilio Rubiera, ([Fundación CTIC](#))

Diego Berrueta, ([Fundación CTIC](#))

Luis Polo, ([Fundación CTIC](#))

Acknowledgements:

See [acknowledgements](#)

Copyright © 2008 by [Fundación CTIC](#). This work is licensed under [Creative Commons Attribution-Share Alike 3.0](#), so you are granted to use, copy, distribute, transmit and adapt this work, but you must attribute the work to the original authors and distribute the resulting work under similar terms.



This work is supported by [Vulcano project](#) and is partially funded by the Spanish Ministry of Industry, Tourism and Commerce under the National Plan for Scientific Research, Development and Technological Innovation 2007-2010 (grant number TSI-020301-2008-22).

Abstract

The DOCOM Ontology provides the main concepts and properties required to describe compound documents made from smaller fragments that are assembled to produce the final document. This report contains a detailed description of the DOCOM Ontology.

Status of this Document

NOTE: *This section describes the status of this document at the time of its publication. Other documents may supersede this document.*

This is a work in progress! This document is changing on a daily if not hourly basis. Comments are very welcome, please send them to the authors. Thank you.

Table of Contents

1. [Introduction](#)
 - 1.1. [Terminology and notation](#)
2. [DOCOM ontology at a glance](#)
3. [DOCOM ontology overview](#)
 - 3.1. [Example](#)
4. [Cross-reference for DOCOM classes and properties](#)

Appendices

- A. [References](#)
- B. [Namespace references and prefixes](#)
- C. [Changes in this version](#) (Non-Normative)
- D. [Acknowledgements](#) (Non-Normative)

1 Introduction

1.1. Terminology and notation

The keywords "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#).

Namespace URIs of the general form "http://www.example.com/" represent some application-dependent or context-dependent URI as defined in [RFC 2396](#).

The XML Namespace URIs that MUST be used by implementations of this specification is:

- <http://idi.fundacionctic.org/docom#> - DOCOM Ontology Namespace

2. DOCOM ontology at a glance

An alphabetical index of DOCOM terms, by class (concepts) and by property (relationships, attributes), are given below. All the terms are hyperlinked to their detailed description for quick reference.

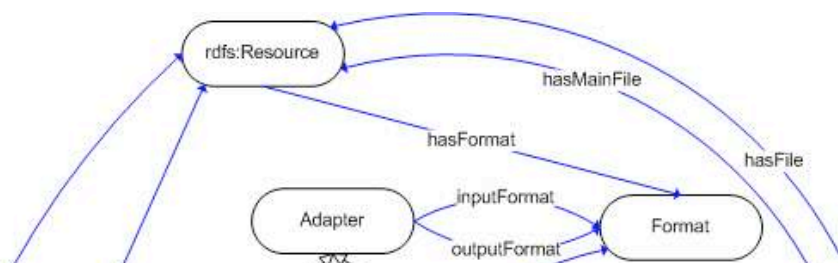
Classes: [Adapter](#), [Annotation](#), [Document](#), [DocumentaryProject](#), [DocumentaryWorkproduct](#), [File](#), [Format](#), [Gap](#), [InputAdapter](#), [IntermediateLanguageDocument](#), [OutputAdapter](#), [PersonProjectRoleAssociation](#), [ProjectRole](#), [SourceDocument](#), [Tag](#), [Ticket](#), [View](#), [Workflow](#),

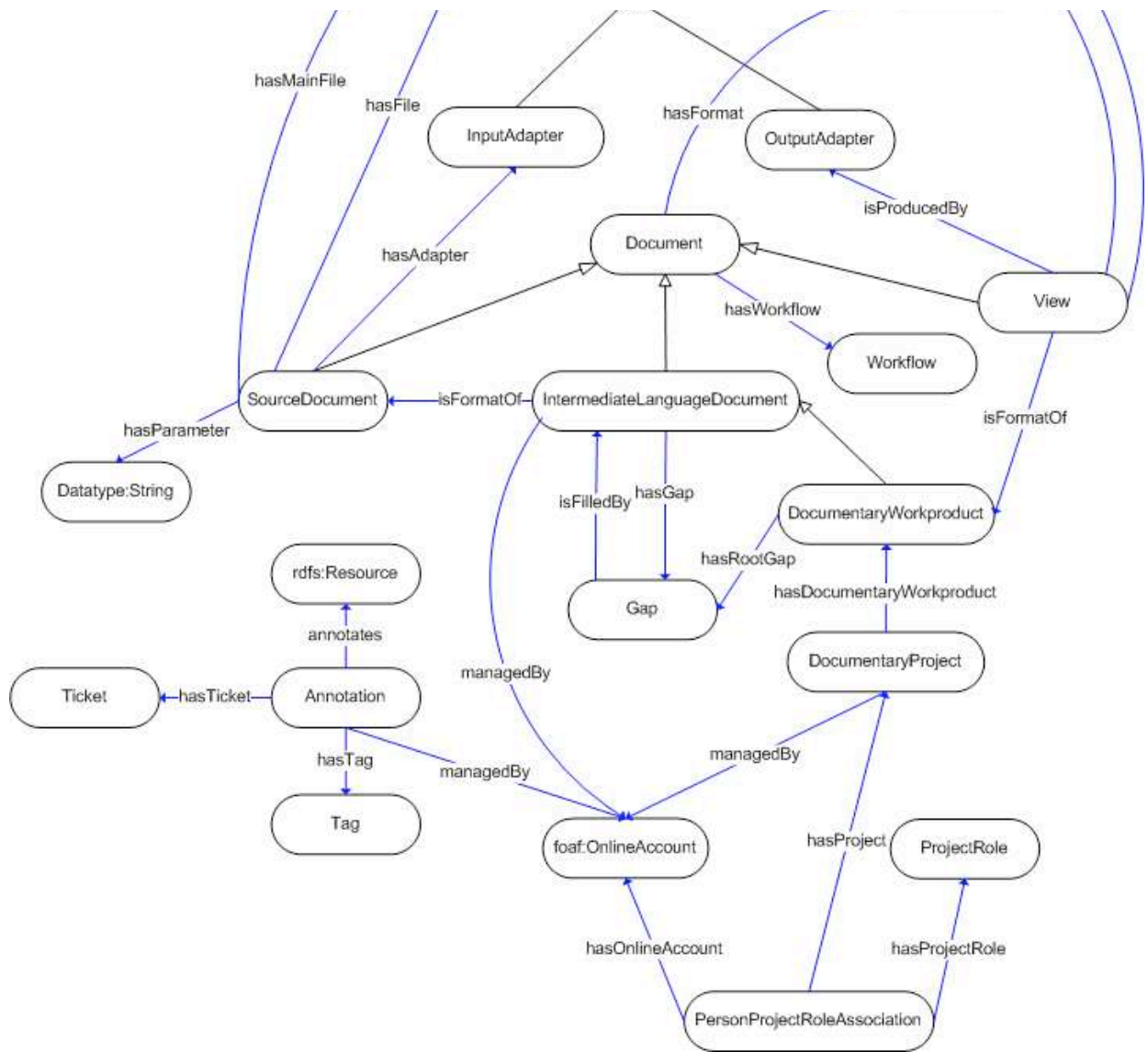
Properties: [annotates](#), [hasAdapter](#), [hasDocumentaryWorkproduct](#), [hasFile](#), [hasFormat](#), [hasGap](#), [hasMainFile](#), [hasOnlineAccount](#), [hasProject](#), [hasProjectRole](#), [hasRootGap](#), [hasTag](#), [hasTicket](#), [hasWorkflow](#), [inputFormat](#), [isFilledBy](#), [isFormatOf](#), [isProducedBy](#), [managedBy](#), [outputFormat](#),

3. DOCOM ontology overview

The DOCOM Ontology definitions presented in this report are written using a computer language ([[OWL](#)]) that makes it easy for software to process some basic facts about the terms in the ontology, and consequently about the things described using the DOCOM vocabulary.

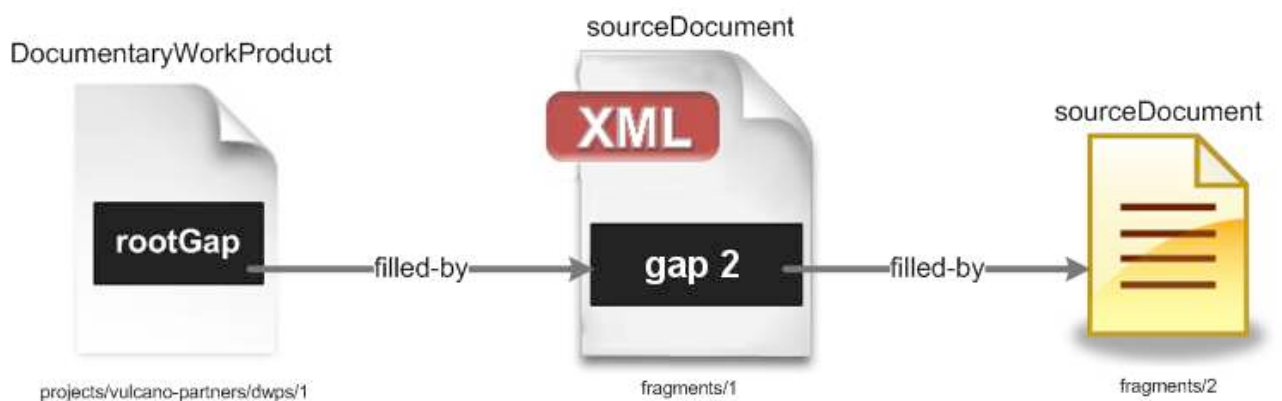
The following figure summarizes the concepts that are part of DOCOM and their relationships. Some auxiliary concepts are referenced by DOCOM, such as `rdfs:Resource` (from [[RDF Schema](#)]) and `foaf:Person` (from [[FOAF](#)]), although they are not part of the DOCOM ontology.





3.1. Example

What follows is a very simple description of a compound document. Note that some resources are not completely described here for the sake of simplicity.



```

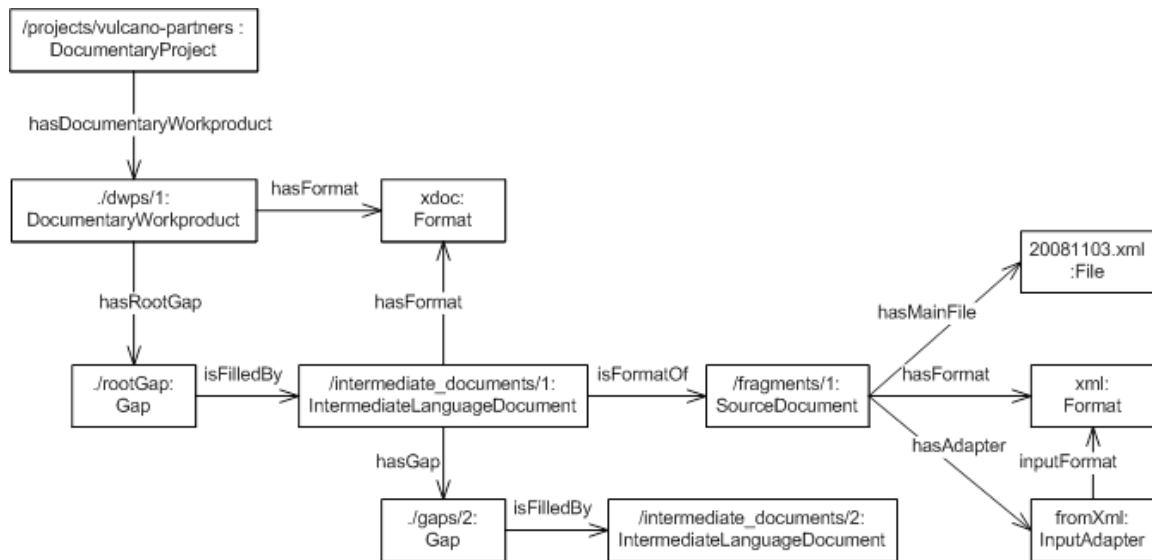
<rdf:RDF ... >
<DocumentaryProject rdf:about="http://example.org/projects/vulcano-partners">
  <hasDocumentaryWorkproduct>
    <DocumentaryWorkproduct rdf:about="http://example.org.org/projects/vulcano-partners/dwps/1">
      <hasFormat rdf:resource="http://example.org/formats/xdoc"/>
      <hasRootGap>
        <Gap rdf:about="http://example.org.org/projects/vulcano-partners/dwps/1/rootGap">

```

```

<isFilledBy>
  <IntermediateLanguageDocument rdf:about="http://example.org/intermediate_docur
    <hasFormat rdf:resource="http://example.org/formats/xdoc"/>
    <isFormatOf>
      <SourceDocument rdf:about="http://example.org/fragments/1">
        <hasMainFile rdf:resource="http://www.example.org/20081103.xml"/>
        <hasFormat rdf:resource="http://example.org/formats/xml"/>
        <hasAdapter about="http://example.org/inputadapters/fromXml"/>
      </SourceDocument>
    </isFormatOf>
    <hasGap>
      <Gap rdf:about="http://example.org.org/fragments/1/gaps/2">
        <isFilledBy rdf:resource="http://example.org/intermediate_documents/2"/>
      </Gap>
    </hasGap>
  </IntermediateLanguageDocument>
</isFilledBy>
</Gap>
</hasRootGap>
</DocumentaryWorkproduct>
</hasDocumentaryWorkproduct>
</DocumentaryProject>
</rdf:RDF>

```



4. Cross-reference for DOCOM classes and properties

Class: docom:Adapter

URI: <http://idi.fundacionctic.org/docom#Adapter>

Adapter - a software element that converts a document from one format to another.

in-domain-of:

[docom:outputFormat](#)

[docom:inputFormat](#)

[back to top](#)

Class: docom:Annotation

URI: <http://idi.fundacionctic.org/docom#Annotation>

Annotation - the action of annotating a resource with a tag by a user.

in-domain-of:

[docom:hasTicket](#)
[docom:annotates](#)
[docom:hasTag](#)
[docom:managedBy](#)

[\[back to top\]](#)

Class: docom:Document

URI: <http://idi.fundacionctic.org/docom#Document>

Document - a writing item that contains information designed and presented as an individual entity.

in-domain-of:

[docom:hasWorkflow](#)
[docom:isFormatOf](#)
[docom:hasFormat](#)

in-range-of:

[docom:isFormatOf](#)

[\[back to top\]](#)

Class: docom:DocumentaryProject

URI: <http://idi.fundacionctic.org/docom#DocumentaryProject>

Documentary project - an initiative that has as objective to achieve documentary workproducts.

in-domain-of:

[docom:hasDocumentaryWorkproduct](#)
[docom:managedBy](#)

in-range-of:

[docom:hasProject](#)

[\[back to top\]](#)

Class: docom:DocumentaryWorkproduct

URI: <http://idi.fundacionctic.org/docom#DocumentaryWorkproduct>

Documentary workproduct - input/output documents obtained as a result of a workflow activity and potentially generated as views. (comes from SPEM 2.0 by OMG).

sub-class-of:

[docom:IntermediateLanguageDocument](#)

in-domain-of:

[docom:hasRootGap](#)

in-range-of:

[docom:hasDocumentaryWorkproduct](#)

[\[back to top\]](#)

Class: docom:File

URI: <http://idi.fundacionctic.org/docom#File>

File - a block of arbitrary information, or resource for storing information.

in-domain-of:

[docom:hasFormat](#)

in-range-of:

[docom:hasFile](#)

[\[back to top\]](#)

Class: docom:Format

URI: <http://idi.fundacionctic.org/docom#Format>

Format - the way in which data is stored or held to be worked on or just to be shown.

in-range-of:

[docom:hasFormat](#)

[docom:outputFormat](#)

[docom:inputFormat](#)

[\[back to top\]](#)

Class: docom:Gap

URI: <http://idi.fundacionctic.org/docom#Gap>

Gap - a space that can be fulfilled by a document.

in-domain-of:

[docom:isFilledBy](#)

in-range-of:

[docom:hasGap](#)

[\[back to top\]](#)

Class: docom:InputAdapter

URI: <http://idi.fundacionctic.org/docom#InputAdapter>

Input adapter - the adapter in charge of the input documents.

sub-class-of:

[docom:Adapter](#)

in-range-of:

[docom:hasAdapter](#)

[\[back to top\]](#)

Class: docom:IntermediateLanguageDocument

URI: <http://idi.fundacionctic.org/docom#IntermediateLanguageDocument>

Intermediate language document - documents formatted in an intermediate language.

sub-class-of:

[docom:Document](#)

in-domain-of:

[docom:hasGap](#)

[docom:managedBy](#)

in-range-of:

[docom:isFilledBy](#)

[\[back to top\]](#)

Class: docom:OutputAdapter

URI: <http://idi.fundacionctic.org/docom#OutputAdapter>

Output adapter - the adapter in charge of the output documents.

sub-class-of:

[docom:Adapter](#)

in-range-of:

[docom:isProducedBy](#)

[\[back to top\]](#)

Class: docom:PersonProjectRoleAssociation

URI: <http://idi.fundacionctic.org/docom#PersonProjectRoleAssociation>

Person/project/role association - the association of three entities: person, role and project.

in-domain-of:

[docom:hasProject](#)

[docom:hasOnlineAccount](#)

[docom:hasProjectRole](#)

[\[back to top\]](#)

Class: docom:ProjectRole

URI: <http://idi.fundacionctic.org/docom#ProjectRole>

Project role - the role of a person involved in a project.

in-range-of:

[docom:hasProjectRole](#)

[\[back to top\]](#)

Class: docom:SourceDocument

URI: <http://idi.fundacionctic.org/docom#SourceDocument>

Source document - an origin document from which the new one derives or is obtained.

sub-class-of:

[docom:Document](#)

in-domain-of:

[docom:hasAdapter](#)

[docom:hasFile](#)

[\[back to top\]](#)

Class: docom:Tag

URI: <http://idi.fundacionctic.org/docom#Tag>

Tag - a keyword or term associated with or assigned to a piece of information.

in-range-of:

[docom:hasTag](#)

[\[back to top\]](#)

Class: docom:Ticket

URI: <http://idi.fundacionctic.org/docom#Ticket>

Ticket - a service request, used to track complaints or requests that an issue be handled.

in-range-of:

[docom:hasTicket](#)

[\[back to top\]](#)

Class: docom:View

URI: <http://idi.fundacionctic.org/docom#View>

View - a possible displayed view of a generated document.

sub-class-of:

[docom:Document](#)

in-domain-of:

[docom:isProducedBy](#)

[docom:hasFile](#)

[\[back to top\]](#)

Class: docom:Workflow

URI: <http://idi.fundacionctic.org/docom#Workflow>

Workflow - a depiction of a sequence of operations.

[\[back to top\]](#)

Property: docom:annotates

URI: <http://idi.fundacionctic.org/docom#annotates>

annotates - the act of annotating some resource.

OWL Type:

ObjectProperty

Domain:

[docom:Annotation](#)

[\[back to top\]](#)

Property: docom:hasAdapter

URI: <http://idi.fundacionctic.org/docom#hasAdapter>

has adapter - a property that brings together a document with an adapter able to transform it into a new format.

OWL Type:

ObjectProperty

Domain:

[docom:SourceDocument](#)

Range:

Property: docom:hasDocumentaryWorkproduct

URI: <http://idi.fundacionctic.org/docom#hasDocumentaryWorkproduct>

has documentary workproduct - a documentary project produces, as a result of its activity, a documentary workproduct.

OWL Type:

ObjectProperty

OWL Type:

InverseFunctionalProperty (uniquely identifying property)

Domain:

[docom:DocumentaryProject](#)

Range:

[docom:DocumentaryWorkproduct](#)

Property: docom:hasFile

URI: <http://idi.fundacionctic.org/docom#hasFile>

has file - a document has a physical representation as one or more related files.

OWL Type:

ObjectProperty

Domain:

[docom:View](#)

[docom:SourceDocument](#)

Range:

[docom:File](#)

Property: docom:hasFormat

URI: <http://idi.fundacionctic.org/docom#hasFormat>

has format - a document or a file is encoded in one specific format.

OWL Type:

ObjectProperty

Domain:

[docom:Document](#)

[docom:File](#)

Range:

[docom:Format](#)

Property: docom:hasGap

URI: <http://idi.fundacionctic.org/docom#hasGap>

has gap - a document can have structural gaps that could be filled by other documents.

OWL Type:

ObjectProperty

OWL Type:

InverseFunctionalProperty (uniquely identifying property)

Domain:

[docom:IntermediateLanguageDocument](#)

Range:

[docom:Gap](#)

[\[back to top\]](#)

Property: [docom:hasMainFile](#)

URI: <http://idi.fundacionctic.org/docom#hasMainFile>

has main file - a source document and a view (being both documents) have, among all the integrating resources, one considered as main file.

OWL Type:

ObjectProperty

sub-property-of:

[docom:hasFile](#)

[\[back to top\]](#)

Property: [docom:hasOnlineAccount](#)

URI: <http://idi.fundacionctic.org/docom#hasOnlineAccount>

has on-line account - indicates the name (identifier) associated with this online account.

OWL Type:

ObjectProperty

Domain:

[docom:PersonProjectRoleAssociation](#)

Range:

[foaf:OnlineAccount](#)

[\[back to top\]](#)

Property: [docom:hasProject](#)

URI: <http://idi.fundacionctic.org/docom#hasProject>

has project - a person having a role is involved in some project.

OWL Type:

ObjectProperty

Domain:

[docom:PersonProjectRoleAssociation](#)

Range:

[docom:DocumentaryProject](#)

[\[back to top\]](#)

Property: [docom:hasProjectRole](#)

URI: <http://idi.fundacionctic.org/docom#hasProjectRole>

has project role - a person involved in a project has a role.

OWL Type:

ObjectProperty

Domain:

[docom:PersonProjectRoleAssociation](#)

Range:

[docom:ProjectRole](#)

[\[back to top\]](#)

Property: [docom:hasRootGap](#)

URI: <http://idi.fundacionctic.org/docom#hasRootGap>

has root gap - a documentary workproduct has a single root gap.

OWL Type:

ObjectProperty

OWL Type:

InverseFunctionalProperty (uniquely identifying property)

sub-property-of:

[docom:hasGap](#)

Domain:

[docom:DocumentaryWorkproduct](#)

[\[back to top\]](#)

Property: [docom:hasTag](#)

URI: <http://idi.fundacionctic.org/docom#hasTag>

has tag - the act of annotating involves some tag.

OWL Type:

ObjectProperty

Domain:

[docom:Annotation](#)

Range:

[docom:Tag](#)

[\[back to top\]](#)

Property: [docom:hasTicket](#)

URI: <http://idi.fundacionctic.org/docom#hasTicket>

has ticket - an annotation could refer to a ticket.

OWL Type:

ObjectProperty

Domain:

[docom:Annotation](#)

Range:

[docom:Ticket](#)

[\[back to top\]](#)

Property: [docom:hasWorkflow](#)

URI: <http://idi.fundacionctic.org/docom#hasWorkflow>

has workflow - a document has a workflow.

OWL Type:
ObjectProperty

Domain:
[docom:Document](#)

[\[back to top\]](#)

Property: docom:inputFormat

URI: <http://idi.fundacionctic.org/docom#inputFormat>

input format - the adapter has an input format (of the document to be processed).

OWL Type:
ObjectProperty

Domain:
[docom:Adapter](#)

Range:
[docom:Format](#)

[\[back to top\]](#)

Property: docom:isFilledBy

URI: <http://idi.fundacionctic.org/docom#isFilledBy>

is filled by - a gap is filled by a document.

OWL Type:
ObjectProperty

Domain:
[docom:Gap](#)

Range:
[docom:IntermediateLanguageDocument](#)

[\[back to top\]](#)

Property: docom:isFormatOf

URI: <http://idi.fundacionctic.org/docom#isFormatOf>

is format of - Created as subproperty of dcterm:isFormatOf in order to be more specific regarding domain and range.

OWL Type:
ObjectProperty

sub-property-of:
[dct:isFormatOf](#)

Domain:
[docom:Document](#)

Range:
[docom:Document](#)

[\[back to top\]](#)

Property: docom:isProducedBy

URI: <http://idi.fundacionctic.org/docom#isProducedBy>

is produced by - a document that has been transformed into a new format is produced by an

adapter.

OWL Type:

ObjectProperty

Domain:

[docom:View](#)

Range:

[docom:OutputAdapter](#)

[\[back to top\]](#)

Property: [docom:managedBy](#)

URI: <http://idi.fundacionctic.org/docom#managedBy>

managed by - a documentary project, an intermediate language document and its annotation are managed by an on-line account.

OWL Type:

ObjectProperty

Domain:

[docom:Annotation](#)

[docom:IntermediateLanguageDocument](#)

[docom:DocumentaryProject](#)

Range:

[foaf:OnlineAccount](#)

[\[back to top\]](#)

Property: [docom:outputFormat](#)

URI: <http://idi.fundacionctic.org/docom#outputFormat>

output format - the adapter has an output format (of the document to be processed).

OWL Type:

ObjectProperty

Domain:

[docom:Adapter](#)

Range:

[docom:Format](#)

[\[back to top\]](#)

A References

FOAF

[FOAF Vocabulary Specification](#), Dan Brickley and Libby Miller.

Available at <http://xmlns.com/foaf/0.1/>

OWL

Peter F. Patel-Schneider, Patrick Hayes, and Ian Horrocks, [OWL Web Ontology Language Semantics and Abstract Syntax](#), W3C Recommendation, February 2004,

RDFS

Dan Brickley and R. V. Guha, [RDF Vocabulary Description Language 1.0: RDF Schema](#), W3C Recommendation, February 2004.

RFC2119

S. Bradner, [RFC 2119: Key words for use in RFCs to Indicate Requirement Levels](#). Internet Engineering Task Force, 1997.

RFC2396

T. Berners-Lee et al., [RFC 2396: Uniform Resource Identifiers \(URI\): Generic Syntax](#). Internet Engineering Task Force, 1998.

RDF

Graham Klyne, Jeremy Carroll, [Resource Description Framework \(RDF\): Concepts and Abstract Syntax](#), W3C Recommendation 10 Feb 2004.

B Namespace references and prefixes

The following namespaces are referenced by the DOCOM ontology:

Prefix	XML Namespace	Specification
dcterms	http://purl.org/dc/terms/	The Dublin Core Metadata Initiative Metadata Terms
foaf	http://xmlns.com/foaf/0.1/	Friend of a Friend (FOAF) Vocabulary
owl	http://www.w3.org/2002/07/owl#	OWL Web Ontology Language
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#	Resource Description Framework (RDF)
rdfs	http://www.w3.org/2000/01/rdf-schema#	RDF Vocabulary Description Language 1.0: RDF Schema
xsd	http://www.w3.org/2001/XMLSchema#	XML Schema (Datatypes)

Regarding the DOCOM ontology itself, the prefix **docom** is suggested as an alias for the namespace <http://idi.fundacionctic.org/docom#>, although applications may use any other prefix of their choice.

C Changes in this version (Non-Normative)

- 2008-12-10: Initial version of the DOCOM Ontology specification.

D Acknowledgements (Non-Normative)

This document was produced with the participation of the [Vulcano project](#) consortium (by alphabetical order):

- [Atos Origin](#)
- [Andago](#)
- [CTIC](#)
- [European Software Institute](#)
- [Instituto Tecnológico de Informática](#)
- [Intelligent Software Components](#)
- [Informática Gesfor](#)
- [Instituto Tecnológico de Aragón](#)
- [Telefónica Investigación y Desarrollo](#)
- [Universitat Oberta de Catalunya](#)
- [Universidad Politécnica de Madrid](#)
- [Universidad Rey Juan Carlos de Madrid](#)
- [Yaco Consultoría](#)



[Vulcano project](#) is partially funded by the Spanish Ministry of Industry, Tourism and Commerce under the National Plan for Scientific Research, Development and Technological Innovation 2007-2010 (grant number TSI-020301-2008-22).