



ICA ROMA

Archives bridging the digital gap. New and disruptive technologies, methods and environments are constantly making inroads into our societies. Blockchain, Big Data analytics, Internet of Things, virtual reality, augmented reality, Artificial Intelligence, algorithmic techniques and gamification are changing not only the archival scene, but also our way of living. What is the impact of these new technologies on records and archives? What are the challenges that need to be faced in relation to technological development, obsolescence, software design or long-term digital preservation? What role can archives perform in helping users to deal with new technologies? How are archivists and records managers changing their professional profile? How is the distance between the different professionals working on records and archives (including, but not limited to, records managers, archivists, ICT professionals, lawyers and policy makers) changing?

Els deadlines :

29 November 2021 Opening Call for papers

21 March 2022 Deadline Submission of proposals

21 May 2022 Notification of acceptance

August 2022 Submission of final slides and/or text for display at the conference

September 2022 Proceedings available for online distribution at conference

November 2022 Final proceedings available for online distribution

Poster 46 - Building Interoperability in Archival Management Systems, Between Data and Records Long-term Digital Preservation for Joint Research. Challenges in a Fast Changing World. Prof. Josep Ballber, Dr. Pepita Raventós, Dr. Roberto García

<https://www.ica.org/es/ica-roma-2022-programa-preliminar>

https://www.ica.org/sites/default/files/preliminary_programme_for_publication_29_july_2022.pdf

Abstract ID : **438**

“BUILDING INTEROPERABILITY IN ARCHIVAL MANAGEMENT SYSTEMS, BETWEEN DATA AND RECORDS LONG-TERM DIGITAL PRESERVATION FOR JOINT RESEARCH.”

Josep Ballber¹, doctorand Industrial de la Universitat de Lleida.

Pepita Raventós², cap de servei Arxiu i Gestió de Documents. Universitat de Lleida. pepita.raventos@udl.cat

Roberto García³

Roberto González⁴

¹ Doctorand Industrial de la Universitat de Lleida

² Cap de servei Arxiu i Gestió de Documents. Universitat de Lleida. pepita.raventos@udl.cat.

³ García

⁴ González

1. INTRODUCCIÓ

Amb aquest pòster es vol donar projecció a la recerca que estem fent en preservació digital. L'objectiu de la recerca és trobar la millor forma per preservar digitalment els documents i les dades dins d'un context universitari partint d'uns estudis de cas reals. Ens hem trobat que per fer una bona preservació digital a llarg termini (les evidències documentals es depositen en un arxiu digital), abans hem de fer una bona gestió documental i durant el transcurs de la investigació hem anat orientant la recerca.

Això comporta fer una descripció de l'objecte que volem preservar a partir de les seves metadades, però ens hem trobat que els standards actuals com ISAD(g) o ISO23081 tenen una estructura massa jeràrquica/rígida i hem decidit modelar cap a RiC amb una estructura de grafs, amb una visió més semàntica basada en ontologies. També ens hem trobat que en aquest procés de preservació digital intervenen diversos sistemes o dispositius (agents), i a través de la gestió de les metadades és vol facilitar la interoperabilitat entre els diferents sistemes/dispositius (agents).

La metodologia que farem servir per intentar innovar és primer utilitzar la capacitat d'observar enlloc d'anar desde la tecnologia directament a una solució. L'objectiu és tenir ben perimetrada abans la solució perquè hem observat bé el problema o necessitat (el que hem fet durant aquest primer any) i en aquesta segona fase anirem a buscar la tecnologia que ens ajudi a crear-la (durant el segon any).

Ara mateix estem explorant la tecnologia semàntica per poder analitzar com encaixen els models amb RiC (Records in Context) i les especificacions d'eArchiving. També l'ús de l'arquitectura de microserveis per processar la informació (recol·lectant documents i (meta)dades ...) i creant un model de gestió de metadades comú que ens ajudarà a utilitzar patrons precarregats d'esquemes de metadades. També s'exploraran solucions aplicant intel·ligència artificial i Machine learning.

Es proposen 3 estudis de cas :

- Actes Acadèmiques (Document administratiu/academic UdL)
- Guies Docents (Actius digitals del campus virtual)
- Quaderns Docents (Documents històrics Escola Normal)

Però en aquest poster ens centrarem en el primer estudi d'un cas real de les actes acadèmiques de la UdL

En aquest estudi de cas tenim un ecosistema d'origen amb diferents sistemes d'informació, que podem dividir entre els de gestió de negoci i els d'arxiu.

a) Gestió de negoci :

1. Gestor acadèmic (UXXI OCU) > Mòdul de actes acadèmiques
2. Portasignatures. (@Firma)

b) Gestió d'Arxiu :

1. Gestor documental. GDC. CSUC
2. Arxiu digital. eArxiu CSUC

Les metadades de les aplicacions de negoci s'han recopilat en un fitxer (metadades_origen.xlsx) i han estat obtingudes des del pdf de l'acta acadèmica i de la informació/metadades que hi consta i que prové de les dues aplicacions abans esmentades.

Les metadades de gestió documental i arxiu digital es concreten:

- Gestor Documental (GdC CSUC) vocabulari de metadades de gestió documental (fons, serie,uds..) que compleix amb l'esquema e-EMGDE (Metadatos para la Gestión del Documento Electrónico de ENI). metadades_desti_GdC.xlsx
- Arxiu Digital (eArxiu CSUC) metadades de preservació digital que segueix l'esquema de metadades PREMIS(Preservation Metadata Maintenance Activity) metadades_desti_eArxiu.xlsx
- L'acta d'exemple és una del GRAU EN TREBALL SOCIAL : 2020-21(UNI) Assign. 100948 - PSICOLOGIA SOCIAL Grup 1 Ordre 0 .pdf_firmado-1.pdf
 - L'enllaç directe al document dins del GdC és : <https://gdc.udl.cat/alejandria/services/gestordocumental/fitxer/866a957f-04ad-47cf-bd20-2d0f9a6ed4a5>
- L'exemple d'acta acadèmica segueix el model documental (fons, serie i unitat documental simple) i usa un quadre de classificació jeràrquic QdC(quadre de classificació) i per tant, existeixen relacions d'herència : (Annex 1)
 - F(fons) -> Fons de la Universitat de Lleida
 - CL(clase) -> M1000 Ordenació de la Docència
 - SCL (Subclase) -> M1024 Avaluació acadèmica i docent
 - D4 (nivell 4 del quadre de classificació jeràrquic) -> M1025 Avaluació dels estudiants
 - D5 (nivell 4 del quadre de classificació jeràrquic) -> M1030 Actes de qualificacions
 - S (serie) -> M1030N4195 GRAU EN TREBALL SOCIAL
 - UDS (Unitat documental Simple) -> 2020-21(UNI) Assign. 100948 - PSICOLOGIA SOCIAL Grup 1

La sèrie documental està composta pel codi de classificació (M1030 Actes de qualificacions) i la subdivisió nominativa (N4195 GRAU EN TREBALL SOCIAL). Aquesta darrera correspon al nom del pla docent (estudi on aquest pla està associat).

Algunes de les metadades de negoci (UXXI i @firma) es traslladen al GDC i a eArxiu.

- L'arxiver de forma manual a través de la interfície gràfica de l'aplicació.
- Amb l'automatització del procés a través de la API.
- Generades automàticament pel GdC al fer l'ingrés a l'arxiu (com serien timestamps, auditoria, l'accés, disposició, etc ..)

L'exemple de l'acta acadèmica segons el model documental jeràrquic i aquesta és l'estructura que reproduïx l'aplicació de gestió documental i eArxiu.

L'acta es troba el gestor documental(GdC), l'estudi preveu fer la càrrega automàticament un cop signada pel professor. L'exemple d'acta, encara no s'ha traspasat a l'arxiu digital (eArxiu) però està previst de fer-se amb les metadades informades per crear el paquet d'informació d'arxiu (PIA).

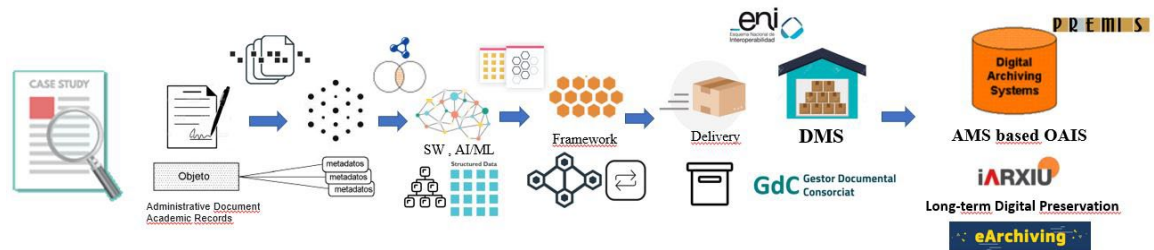


Figura 1. Procés definit per la preservació digital de les actes acadèmiques

[MODEL DOCUMENTAL aplicat] segons ISO 23081, atenent la multientitat, document – agent---activitat---regulació—relació segons el model multinivell segons la norma ISAD (G).

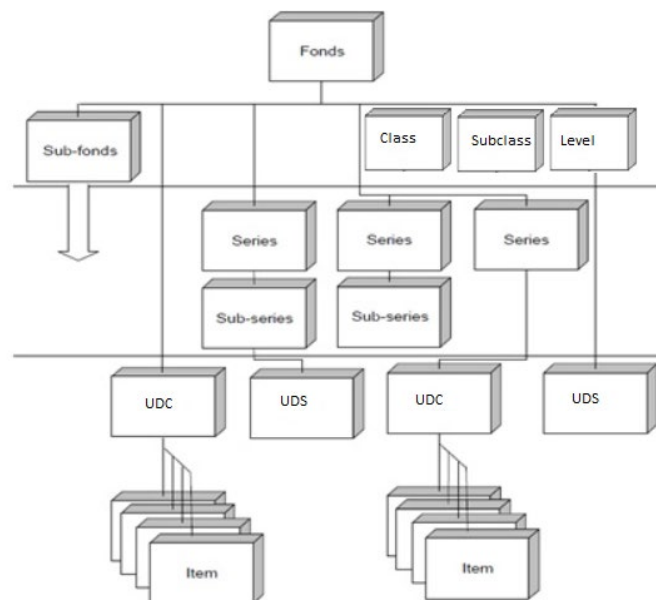


Figura 2. Dibuix reproduït sobre el model multinivell pel Josep en el correu de 28/7/2022]

Aquest model està demostrant que s'està esgotant perquè no permet tenir en compte els diferents contextos de pels quals passa el document i per tant queden sense documentar. [En aquest punt ens pot ajudar el mateix document de R&C quan dona els arguments de perquè aquesta evolució.]

L'exemple de l'acta acadèmica i seguint Records in Context, segons el disseny a través de la web semàntica,

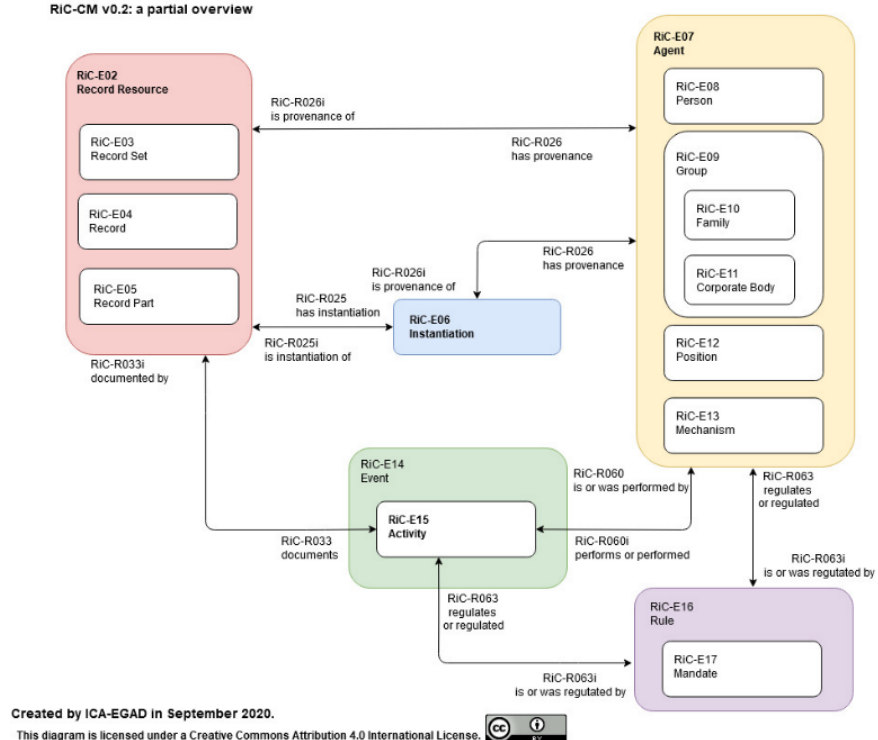


Figura 3 : Representació del model conceptual del RiC (Records in Context)

S'estudia si el model es converteix en multidimensional i permet complir amb el model multientitat definit per ISO 23081.



Figure 2. Left: Representation of data in a hierarchical structure like XML or other markup-language. Right: Representation of data in triples that results in a graph-structure.

El dibuix del diagrama respon a aquesta evolució, però s'ha de veure el seguiment del model documental i la seva evolució d'acord amb aquest model multientitat i després veure la interoperabilitat entre sistemes.

Per la identificació de l'acta acadèmica com a estudi de cas, s'ha tingut en compte per a modelar el diagrama la titulació, l'assignatura, curs acadèmic, grup, etc. També cal tenir en compte els agents. En aquest cas no són persones, son grups d'agents. I s'ha modelat tots els grups que apareixen a les dades originals: facultat, titulació, assignatura, assignatura segons el curs, per ex. el curs 20-21, grup dins de l'assignatura, etc. També s'ha afegit el curs, com a grup de totes les assignatures impartides en aquell any, aprofitant que amb aquesta estructura es pot tenir un graf i es pot anar més enllà de l'estructura d'arbre. De fet, mantenint-lo independent de la jerarquia de grups de la UdL, es podria fer servir per diferents institucions i per tant, fer-lo servir des de la vessant per recuperar assignatures del mateix any a diferents institucions...

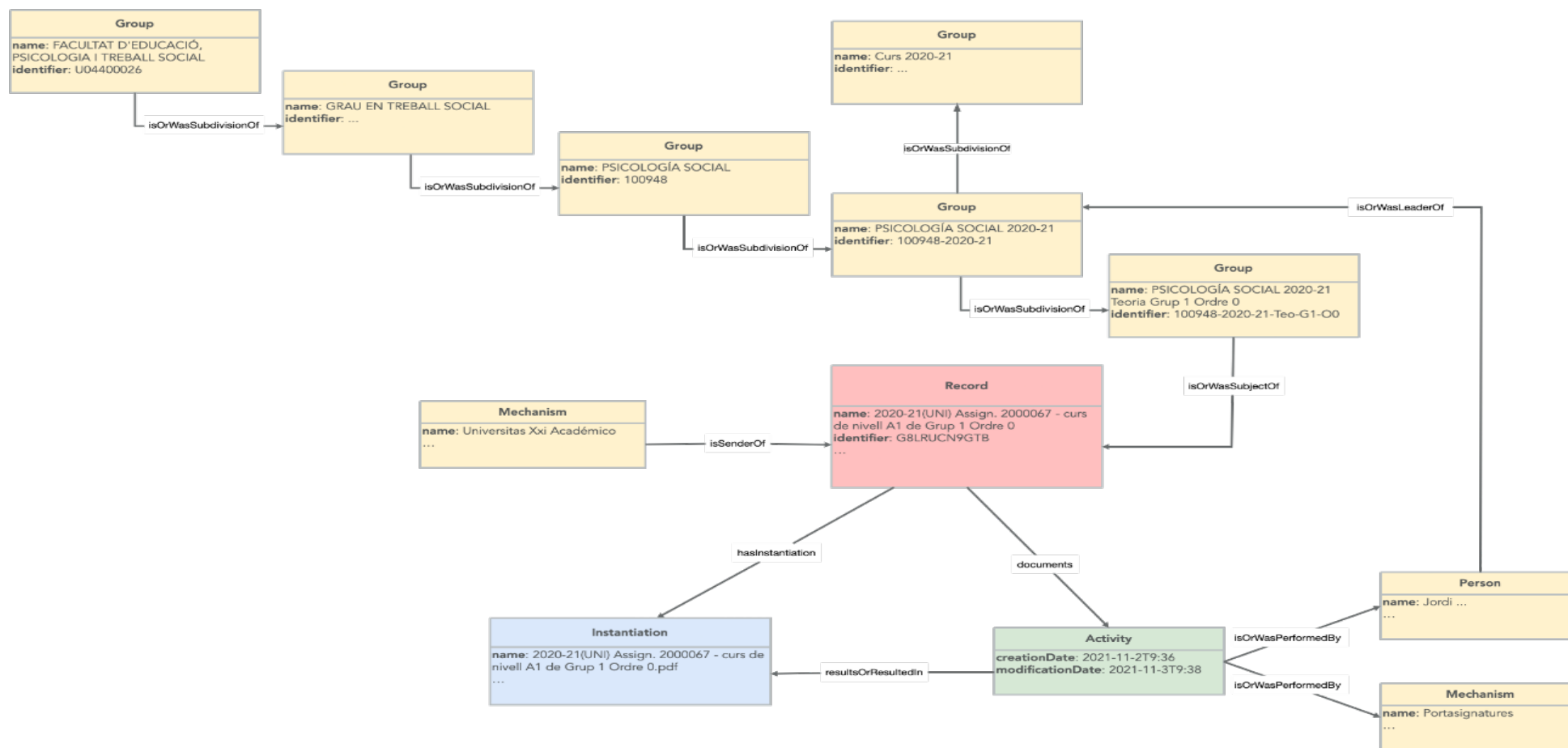


Diagrama 1. Versió inicial de l'adaptació del RiC-CM a les actes acadèmiques

UNE ISO 23081:2021, en la p. 15 fa referència al model multientitat indicat més amunt i deixa endarrere el model multinivell.

La proposta és estudiar si la web semàntica permet ampliar aquests contextos partint del model RiC Records in Context).

De moment l'estudi se centra en les dades per una única acta. Si es tenen en compte les dades de gestió documental i arxiu digital s'afegeixen els record sets pel fons i la sèrie. També els agents per la UdL i gestió acadèmica identificats [indicar el Negociat, el professor o professora com autor o autora]

Finalment, sobre les regles. En el marc del model multientitat les regles es refereixen a l'entitat normatives o regulacions. A més de la normativa d'accés, s'ha d'atendre també altres regulacions que afectin l'acta acadèmica, com per exemple ho seria una normativa nova que en regulés el seu contingut. També i molt significativa és la regulació de la preservació, que per les connotacions del seu contingut és de conservació permanent i accés restringit (D1. Accés restringit, per contenir majoritàriament dades personals ordinàries. Dades personals no especialment protegides ni merament identificatives contingudes a informació pública directament relacionada amb l'organització, funcionament o activitat pública de l'Administració. Aquesta exclusió queda sense efecte als trenta anys des de la producció del document)

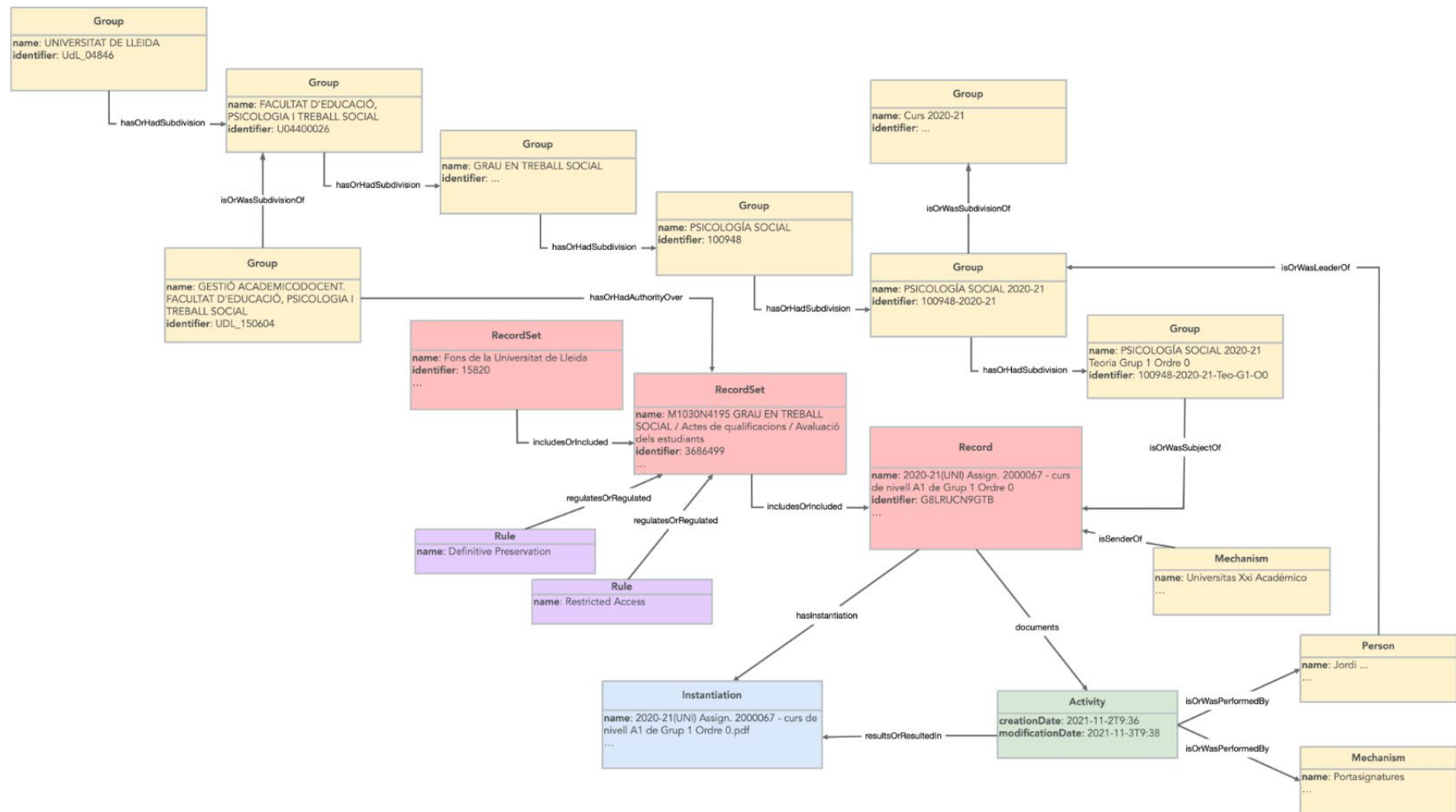
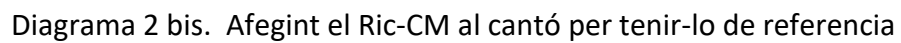


Diagrama 2. S'afegeixen "Rules" de preservació i accés



Encara que la informació en alguns nivells (com els grups) sigui semblant, no és la mateixa. Potser afegint una petita descripció de l'entitat a cada element (com als "Group", per exemple si es estudi, assignatura, grup treball) donaria una millor comprensió de cada element. No se si potser hi faltaria un Group que sigui el "pla d'estudi" com a subdivisió d'estudi ja que a partir d'aquest creem la serie documental i és com agrupem les que fan el RecordSet, adjunto la imatge :

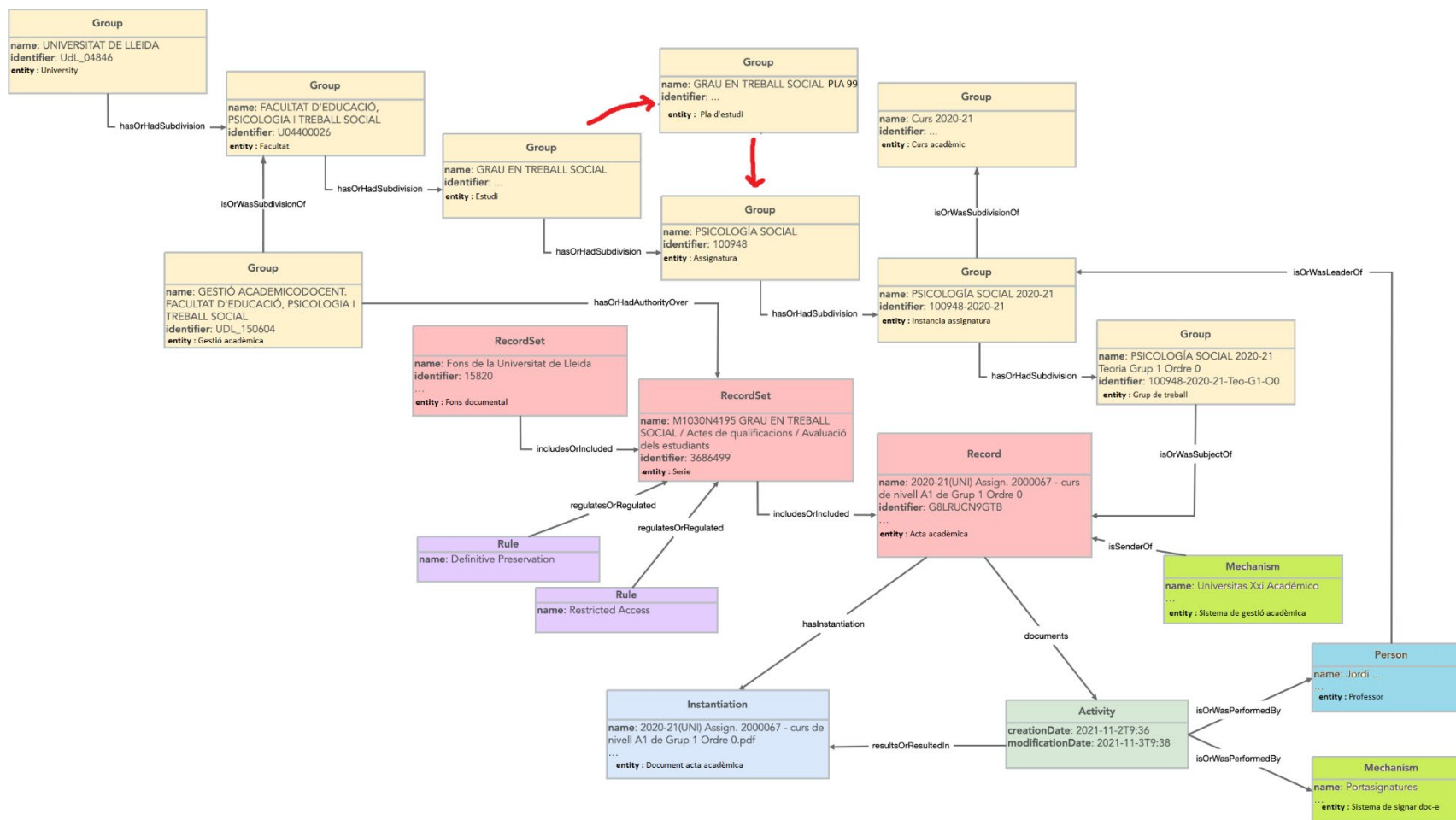


Diagrama 3. Proposta de creació grup "Pla d'estudi"

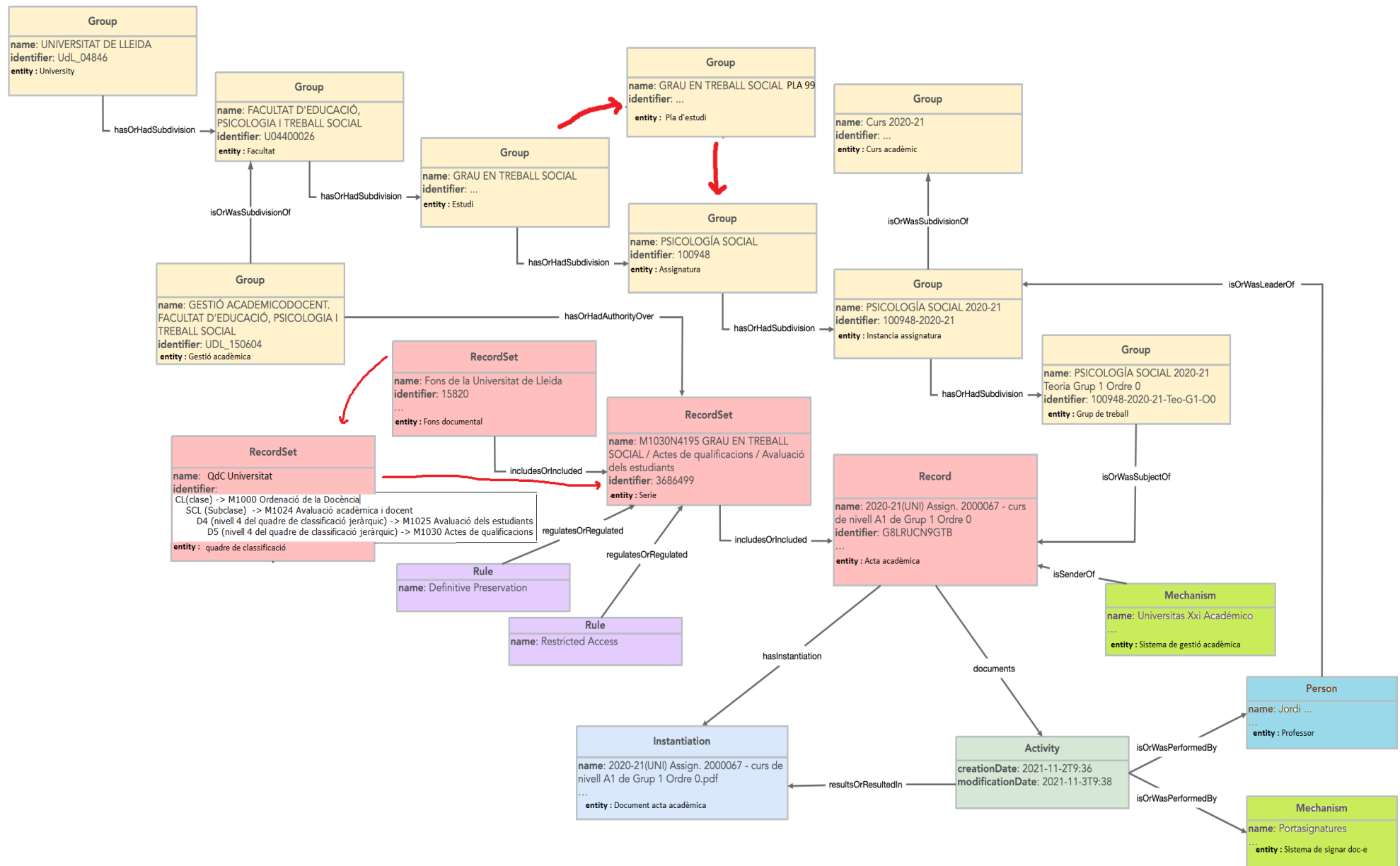


Diagrama 4 . Proposta de crear un RecordSet amb el QdC de la universitat (Pepita: Josep, això no em quadra. El quadre de classificació respon a RiC-E14. Event de segon nivell i a RiC-E15Activity)

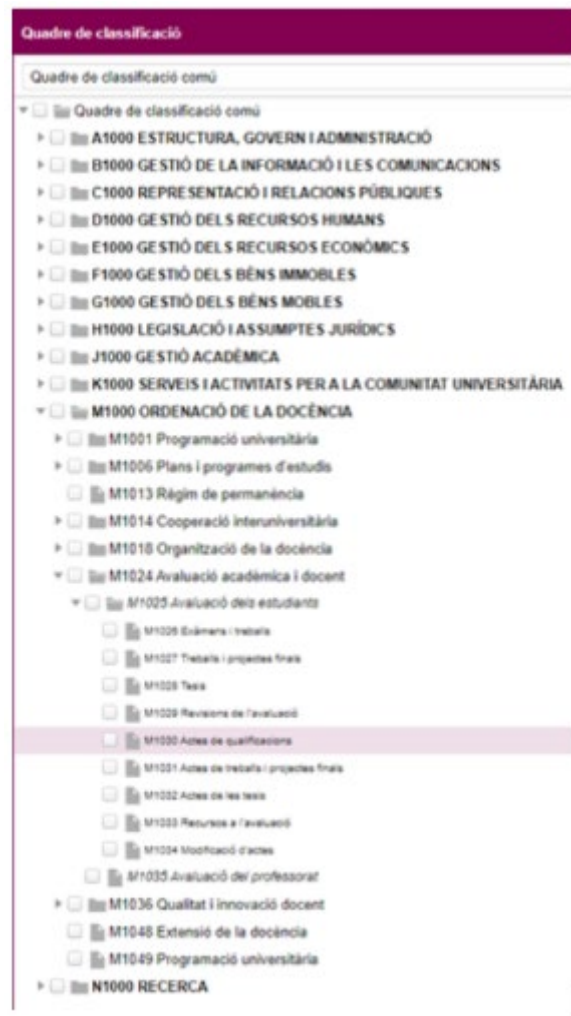
ANNEX 1

L'acta d'exemple és una del GRAU EN TREBALL SOCIAL : 2020-21(UNI) Assign. 100948 - PSICOLOGIA SOCIAL Grup 1 Ordre 0 .pdf_firmado-1.pdf

L'enllaç directe al document dins del GdC és :

<https://gdc.udl.cat/alejandria/services/gestordocumental/fitxer/866a957f-04ad-47cf-bd20-2d0f9a6ed4a5>

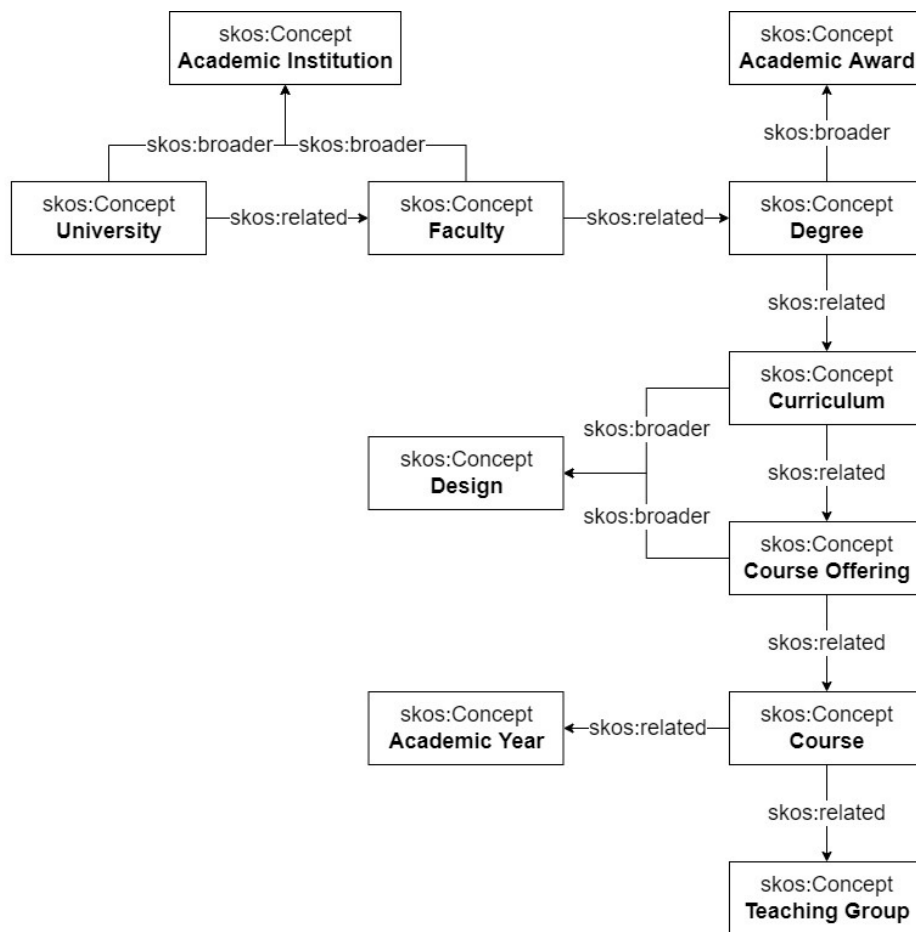
2. L'exemple d'acta acadèmica segueix el model documental (fons, serie i unitat documental simple) i usa un quadre de classificació jeràrquic [QdC\(quadre de classificació\)](#) i per tant, existeixen relacions d'herència :
3. **F(fons)** -> Fons de la Universitat de Lleida
CL(clase) -> M1000 Ordenació de la Docència
SCL (Subclase) -> M1024 Avaluació acadèmica i docent
D4 (nivell 4 del quadre de classificació jeràrquic) -> M1025 Avaluació dels estudiants
D5 (nivell 4 del quadre de classificació jeràrquic) -> M1030 Actes de qualificacions
S (serie) -> M1030N4195 GRAU EN TREBALL SOCIAL
UDS (Unitat documental Simple) -> 2020-21(UNI) Assign. 100948 - PSICOLOGIA SOCIAL Grup 1



La **sèrie documental** està composta pel codi de classificació (M1030 Actes de qualificacions) i la subdivisió nominativa (N4195 GRAU EN TREBALL SOCIAL). Aquesta darrera correspon al nom del **pla docent** (estudi on aquest pla està associat).

ANNEX 2

Definim un model de referencia SKOS (Simple Knowledge Organization System)



REFERENCIES

- E-EMGDE (Esquema de Metadatos para la Gestión del Documento Electrónico)
<https://administracionelectronica.gob.es/ctt/eemgde>
- PREMIS (Preservation Metadata Maintenance Activity) Library of Congress
<https://www.loc.gov/standards/premis/>
- MEL (Metadata Extractor & Loader) Australian National University
<http://ceur-ws.org/Vol-2980/paper399.pdf>
- J2RM: an Ontology-based JSON-to-RDF Mapping Tool
<http://ceur-ws.org/Vol-2721/paper593.pdf>
- AGRIF (The Australian Government Records Interoperability Framework) ontology
<http://linked.data.gov.au/def/agrif>
- ISO 23081:2021. Information and documentation — Metadata for managing records — Part 2: Conceptual and implementation issues. Geneve: ISO, p. 15
- RiC (Records in Contexts). A Conceptual Model for Archival Description (2021). París: International Council on Archives.
<https://www.ica.org/sites/default/files/RiC-CM-0.1.pdf> [access 28/7/2022]
- SKOS (Simple Knowledge Organization System)
<https://www.w3.org/2004/02/skos/>
- 2018 Semantizing and visualising archival metadata: the PIAAF French prototype online PIAAF (Pilote d'interopérabilité pour les Autorités Archivistiques françaises)
<https://www.ica.org/es/node/17606>
<https://piaaf.demo.logilab.fr/>
- 2017 Records in contexts: the road of archives to semantic interoperability
<https://www.emerald.com/insight/content/doi/10.1108/PROG-03-2017-0021/full/html>
<https://core.ac.uk/download/pdf/132545189.pdf>
- 2017 Daniel Pitti - Records in Contexts (RiC) presentation
<http://www.alaarchivos.org/wp-content/uploads/2018/01/1.-Daniel-V.-Pitti.pdf>
- Repositori de github amb alguns exemples de RiC-O de Matterhorn-Switzerland ,
NationalArchivesOfFrance i University-of-Strathclyde-Archives_Scotland :
<https://github.com/ICA-EGAD/RiC-O>
- ICA -Section on University and Research Institution Archives - SUV
<https://www.ica.org/en/section-on-university-and-research-institution-archives-suv>
-

Building interoperability in archival management systems

ICA
ROMA 2022

ARCHIVES: BRIDGING THE GAP
ARCHIVES: COMBLER LES FOSSES
September 19th-23rd, Italy

Josep Ballber
josep.ballber@udl.cat

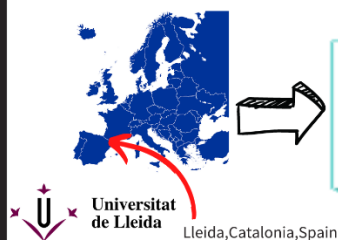
Pepita Raventós
pepita.raventos@udl.cat

Roberto Garcia
roberto.garcia@udl.cat

*Between data and records **long-term digital preservation** for joint search. Challenges in a fast changing world*

What

Observation from use cases



- 1 Academic Records
- 2 Virtual Campus Assets
- 3 Historical Notebooks

Case Study

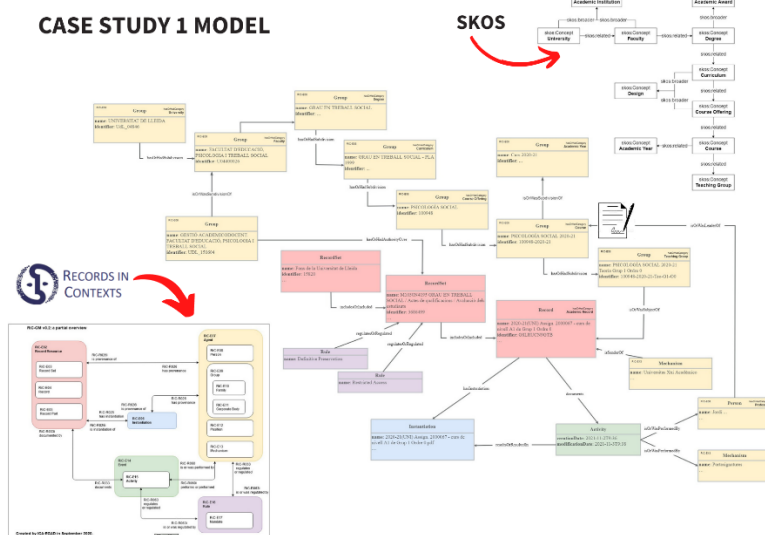
The Archive and Records Management Service of the University of Lleida is working in this direction based on real use cases.

Collaboration and knowledge transfer between university and business are making it possible and helping focus their efforts.

How

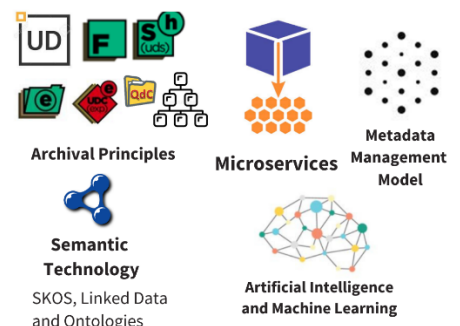
Implement a technological solution

CASE STUDY 1 MODEL



Archival Science & Computer Science Alliance

Exploring synergies between them to facilitate interoperability

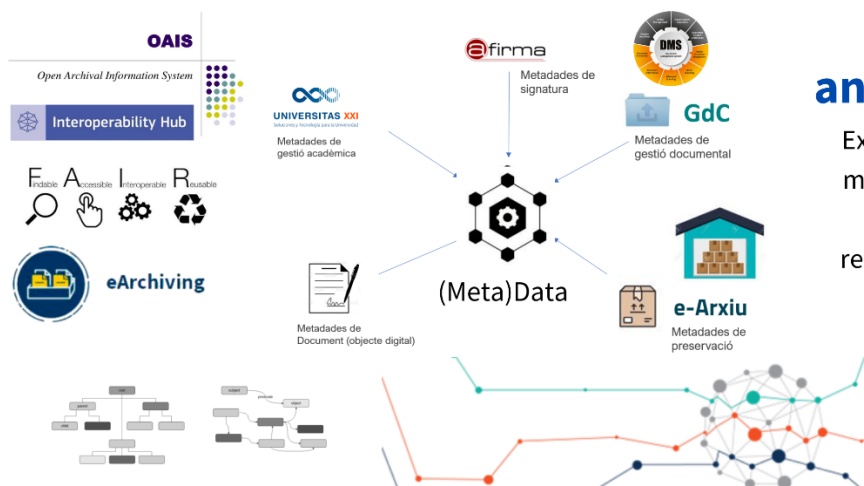


Benefits

Application of research and results

Interoperability and digital preservation

Experience based on a university Records model in accordance with a standardised management system. Preservation of records and data following the eArchiving specifications



nageru solutions



Industrial PhD programme Thesis :
Digital preservation of documents/
records and data for research



SUVOICA
Sección on University and Research Institutions Archives
Recherches des Archives des Universités et des Instituts de Recherche

POSTER CIENTÍFIC

Document gràfic de grans dimensions que serveix per a sintetitzar els resultats d'un treball, projecte o investigació que s'està presentant davant del públic, sigui de manera directa durant un congrès o presentació, o de manera indirecta a una exposició.

ESTRUCTURA

- S'hauran de respectar les indicacions del vostre tutor, facultat o organisme.
- Normalment consta de: títol, autor, logo, introducció, metodologia, resultats, conclusions i bibliografia.

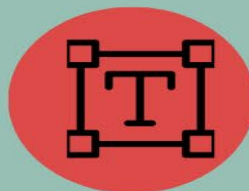


CARACTERÍSTIQUES

Brevetat
Senzillesa
Ordre
Fluïdesa
50% dades visuals
30% text
20% espai buit

CARACTERÍSTIQUES FORMALS

- Format: vertical o horitzontal
- Dimensions: seguiu les instruccions donades pel vostre tutor o facultat
- Tipografia: Arial, Sans Serif, Verdana
- Colors: harmònics, sense estridències



PROGRAMARI

Canva
Power Point
Piktochart
Venngage

PRESENTACIÓ

- Coneix el teu públic
- Disposaràs d'uns minuts per efectuar la teva presentació: assaja!
- Compta el llenguatge verbal i el llenguatge no verbal



**MES INFO:
BIBLIOQUIES.UDL.CAT/POSTER**