

# Fundació Antoni Tàpies

## Arts combinatòries

### Cap a una xarxa pública de coneixements i idees

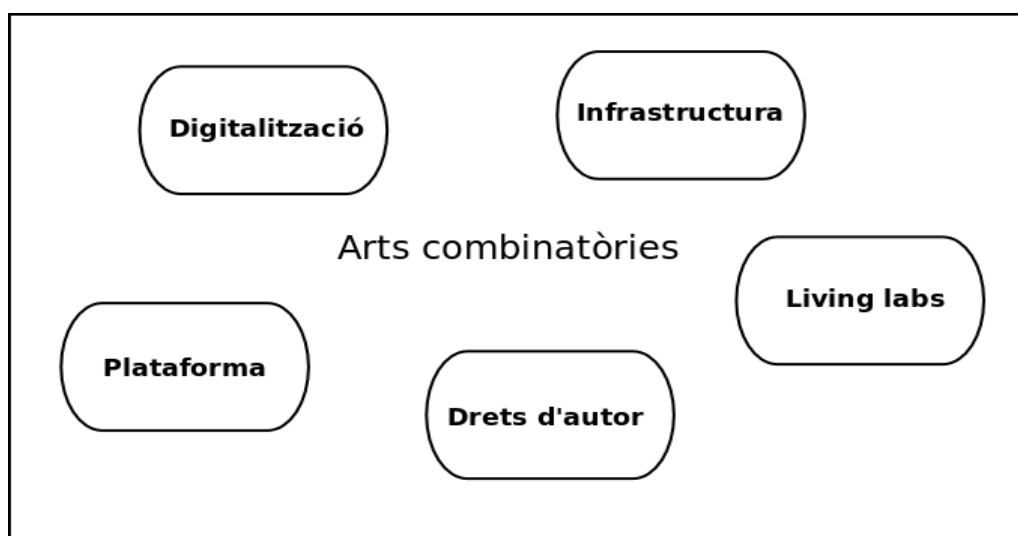
per Jaume Nualart

Introducció.

1. Característiques del projecte.
2. Arxius digitals com a xarxa de dades
3. Funcionalitats i sistema base
4. Visualització i intercanvi de dades
5. Un servei més: decodiFAT
6. Nota final

## Introducció

He col·laborat en el projecte Arts combinatòries de la Fundació Antoni Tàpies 2010-11. Arts combinatòries és un projecte complex que es podria dividir en 5 branques: digitalització, infraestructura, living labs, drets d'autor i plataforma.



*Figura 1: les 5 branques d'Arts combinatòries*

- **Digitalització** de l'arxiu de la Fundació, principalment documents, imatges i àudios vídeos
- **Infraestructura**: obres a la planta de dalt de l'edifici de la Fundació.
- **Living Labs**: reunions i tallers amb persones de diferents perfils interessades en l'art contemporani i en l'activitat de la Fundació. Professors, investigadors i estudiants han ajudat en el disseny de funcionalitats i serveis de la plataforma.
- **Drets d'autor**: el principal problema de les dades en el món de l'art són els drets d'autor que, normalment, no permeten mostrar-ne els continguts. En aquest sentit la Fundació, els experts de *id\_law Partners* i el departament de Programari Lliure de la UPC han dissenyat una estratègia per a poder mostrar el màxim de continguts de l'arxiu malgrat els drets d'autor.
- **Plataforma**: és el cor del projecte. Contempla la construcció d'un sistema en programari lliure capaç de fer útil per al major número de persones les dades que l'arxiu de la Fundació conté; *i és, a més, a on ha col·laborat un servidor.*

La meua participació ha estat el disseny tècnic de la Plataforma. Familiarment, dins la FAT, em diuen el *traductor* entre la FAT i la tecnologia. He participat en els living labs per a conèixer les necessitats i idees d'usuaris. I, en el moment d'escriure aquest article, estem treballant amb la migració de dades, enllestit el sistema ontològic i la integració del sistema de drets d'autor en la plataforma. Així com dissenyant la interacció, l'ergonomia i l'accessibilitat de la plataforma. Es tracta d'una feina que requereix coneixements molt variats, estudi continuu, i consultes a especialistes. Des de ja ara, dono les gràcies a tants amics i amigues als que he fet preguntes específiques en hores intempestives ;-)



Figura 2: Wordle de les àrees de coneixement de la gent que ha paticipat a Arts combinatòries.

En aquest article explico quines són les decisions més importants que hem pres i com hem anat dissenyant aquest nou artilugi pensat per a fruir i treure les essències del coneixement contingut en l'arxiu de la Fundació Antoni Tàpies.

## 1. Característiques del projecte.

Aquests darrers anys moltes institucions culturals han iniciat el complicat procés de digitalització de llurs arxius. El somni és clar: posar el màxim d'informació a disposició de tothom i de la millor manera possible per tal que, creuant tots aquests nous continguts digitals, és pugui enriquir notòriament el coneixement global preservat durant anys. Accessibilitat i interacció amb les dades és el somni.

Arribats a aquest punt, la qüestió és: com fer-ho?

Les tecnologies pensades per a fer realitat aquest somni es van definir fa poc més de 10 anys (XML/RDF/OWL, SPAQL, ...)<sup>1</sup> i tot just fa quatre dies que es comencen a aplicar eficientment. I donat que som en la sub-era de les dades, és veu a prop l'ús massiu d'aquestes tecnologies.

És un moment de canvi a on encara no sabem quines seran, en el futur proper, les millors o les més usades tecnologies d'intercanvi i d'enriquiment d'informació; i, al mateix temps, partim de sistemes analògics o digitals pre-semàntics. La tàctica que hem seguit ha estat escollir la tecnologia que, *a ull d'avui*, permeti un desenvolupament futur de l'aplicació. Aquest és un punt important en la filosofia del projecte, donat que -conscients del moment en el que estem, en la recerca d'una nova era de les dades- quan desenvolupem noves eines que proposen coses noves, cal introduir el vector continuïtat del projecte com a part intrínseca de l'arquitectura dissenyada.

Un resum dels requeriments inicials de la plataforma és:

- Programari amb llicències lliures.
- Produir programaris lliures descarragables i instal·lables.
- Una interfície visualment avançada amb integració d'eines de visualització i representació de resultats de recerca i llistes de continguts.
- Connexions entre dades amb tecnologia de web semàntica.
- Un sistema per a alliberar continguts amb drets d'autor no explícitament reservats.

Hem anat pas a pas definint com acomplir cadascun dels requeriments. Aquí teniu un diagrama de vista general de les funcionalitats i tecnologies escollides.

---

1 Més sobre semantic web: [http://en.wikipedia.org/wiki/Semantic\\_Web](http://en.wikipedia.org/wiki/Semantic_Web)

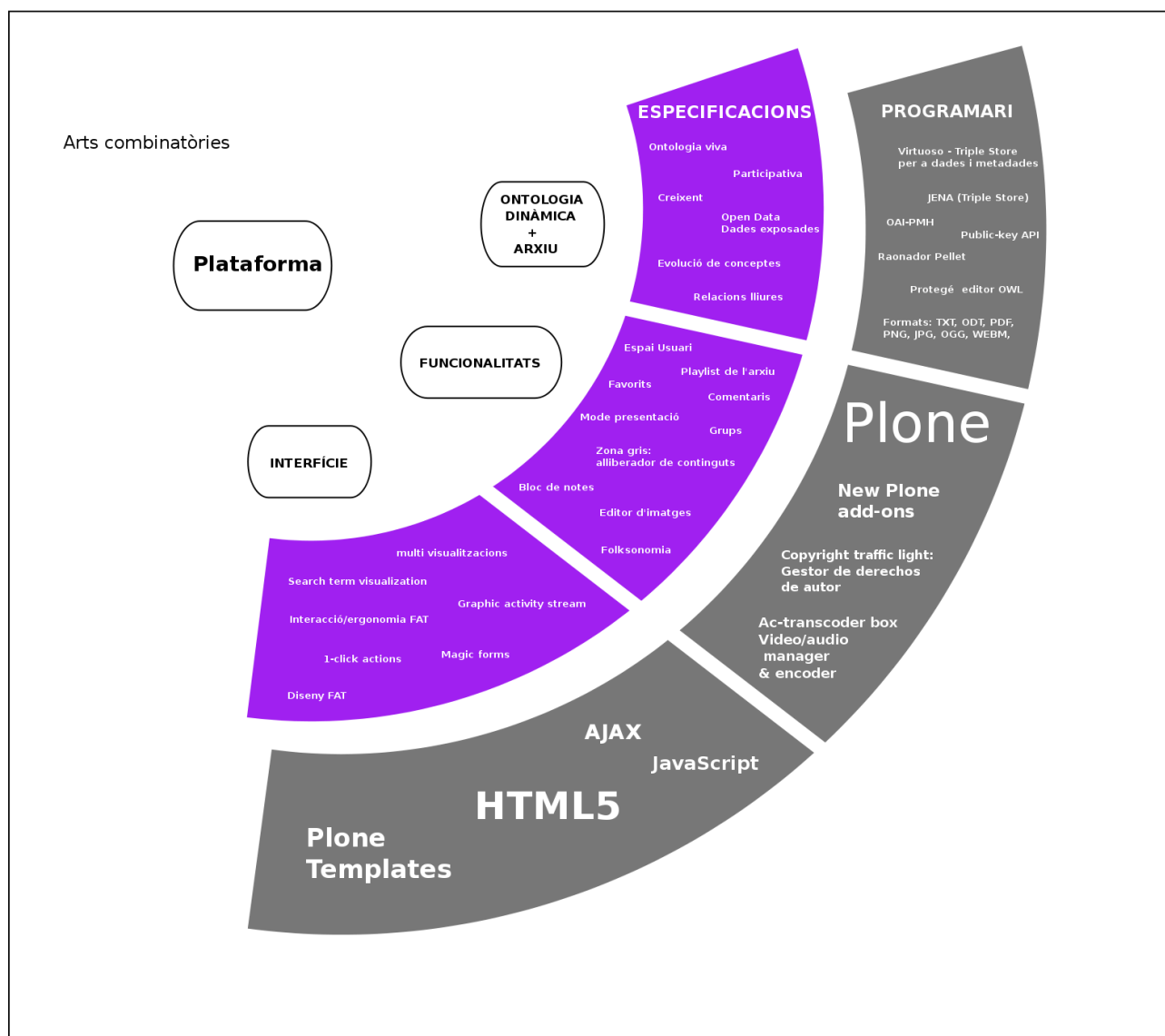


Figura 3: Vista general d'ingredients de la plataforma Arts combinatòries.

## 2. Arxius digitals com a xarxa de dades.

Per definició, cap sistema convencional de classificació, taxonomia, tesaurus o vocabulari soluciona la classificació de la realitat de manera universal. Els arxius que preserven la cultura en són un bon exemple, doncs, a part de preservar el coneixement generat en anys, ens mostren com es percep la realitat en cada moment; ens mostren de quina manera s'ha desat la realitat i com s'ha entès el que ens envolta en cada moment<sup>2</sup>.

Des de fa pocs anys, en aquest procés de recerca embogida de sistemes semàntics o quasi-intel·ligents de processar dades que ens retornin coneixement, es pot dir que patim de **mal d'ontologies**.

<sup>2</sup> [Des d'aquí: #gràciesarxivistes! , però també: #mésarxivistesdigitals! ]

Una ontologia és un sistema per a explicar **que hi ha i com està relacionat** dins d'un domini acotat de coneixement. Per tant, definir una ontologia ens serveix per a sistematitzar una part de la realitat. Així, si diverses persones o màquines adopten una mateixa ontologia, podran comunicar de manera unívoca tot el que faci referència a la realitat representada per l'ontologia. Segurament, el primer problema és que la comunicació amb missatges unívocs, pot ser no prou rica.

Des d'aquest punt de vista, una ontologia és una eina molt pràctica, que pot permetre a persones i màquines comunicar-se de forma limitada segons un semi-llenguatge definit en l'ontologia. De manera molt simplificada la forma física d'una ontologia és un o més fitxers editables tipus text a on s'especifica un arbre d'entitats i sub-entitats, una llista de les relacions possibles que poden tenir aquestes entitats, i, finalment, un llistat de tipus de dades possibles que podem definir per a cada entitat.

Una ontologia és fa més poderosa quan hi ha més gent i dades que la usen. Malgrat això, al món trobem molt poques ontologies d'ús pràcticament universal: ontologies de medicaments, ontologies de malalties,...<sup>3</sup>

Si una ontologia defineix una manera de veure el món, quan entrem en l'ampli terreny de les humanitats, és fa impossible, també per definició, acordar una sola manera de veure el món. Lògicament tenim tantes maneres de veure el món, com persones observant-lo en cada moment.

## La proposta:

Davant d'aquesta situació inevitable i veient com cada institució que s'atreveix a entrar al món de les ontologies ho fa creant-ne una de pròpia, per a Arts combinatòries proposem una **ontologia dinàmica**, una **ontologia viva**, viva com ho és el propi llenguatge i la canviant percepció del món que ens envolta.

De forma semblant a com l'entropia de l'univers només pot créixer, així hem definit l'ontologia d'Arts combinatòries; un llenguatge que només pot créixer. Una complexitat mapejada ontològicament que no pel fet de ser mapejada és menys complexa. Això vol dir que es podrà crear nous noms per a les relacions entre dades de l'arxiu, noves maneres de relacionar conceptes..

Posem un exemple: tenim dos escrits que en l'arxiu estan relacionats com l'*un ve de l'altre*. Resulta que coneixem bé aquests dos documents i creiem més adient que la relació digui que l'*un suggereix l'altre*. Això es podrà fer en el sistema d'ontologia creixent que hem dissenyat.

Algunes conseqüències d'aquest sistema que venen al cap ràpidament són que es pot veure, al cap dels anys, quins conceptes s'han usat més en cada moment. Es pot saber les tendències anomenant les coses en el temps.

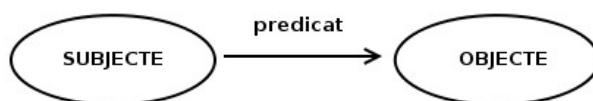
---

<sup>3</sup>Wikipèdia Ontologia: <http://ca.wikipedia.org/wiki/Ontologia>

Ontologies biomediques (via @mem\_somerville) <http://www.bioontology.org/>

Òbviament l'ontologia està feta amb els estàndards recomanats per la W3C<sup>4</sup>, com OWL<sup>5</sup>, de manera que es pot compartir informació amb altres fonts de dades obertes. Existirà una API<sup>6</sup> per a poder fer aplicacions remotes amb les dades de l'arxiu.

Estem en temps de xarxes. Ja es valora més les connexions entre dades que no pas les mateixes dades [7,8]. Tècnicament, com a innovació addicional proposem que totes les dades de l'arxiu estiguin desades en forma de **triplets RDF**<sup>9</sup>. Per entendre'ns, això vol dir que les dades es desen en taules especials de tres columnes a on tenim sempre un subjecte, un objecte i una propietat que els relaciona (A-relació-B). Fins i tot les anomenades *metadades* estaran en RDF.



*Figura 4: Esquema d'un triplet RDF, que relaciona dues classes o entitats definides, per una propietat; per exemple: La imatge A (subjecte) pertany (predicat) a l'expedient 2011 (objecte)*

Amb tot això hem de veure l'arxiu resultant com una **xarxa de dades** amb relacions concretes, carregades de significat semàntic; significat que pot ser interpretat per màquines i persones segons cada versió carregada de l'ontologia canviant d'Arts combinatòries.

---

4 Més info: <http://w3c.org>

5 Més info: [http://ca.wikipedia.org/wiki/Llenguatge\\_d%27ontologies\\_Web](http://ca.wikipedia.org/wiki/Llenguatge_d%27ontologies_Web)

6 Més info: <http://ca.wikipedia.org/wiki/API>

7 Giant Global Graph, T. Berners.Lee (2007) <http://dig.csail.mit.edu/breadcrumbs/node/215>

8 Linked Data - Connect Distributed Data across the Web. <http://linkeddata.org/>

9 Més info: [http://en.wikipedia.org/wiki/Resource\\_Description\\_Framework](http://en.wikipedia.org/wiki/Resource_Description_Framework)

### 3. Funcionalitats

La plataforma *Arts combinatòries* també és un **espai de treball**. Per això hem pensat en dotar d'eines estàndards, però també hem introduït petites funcionalitats menys habituals en plataformes semblants.

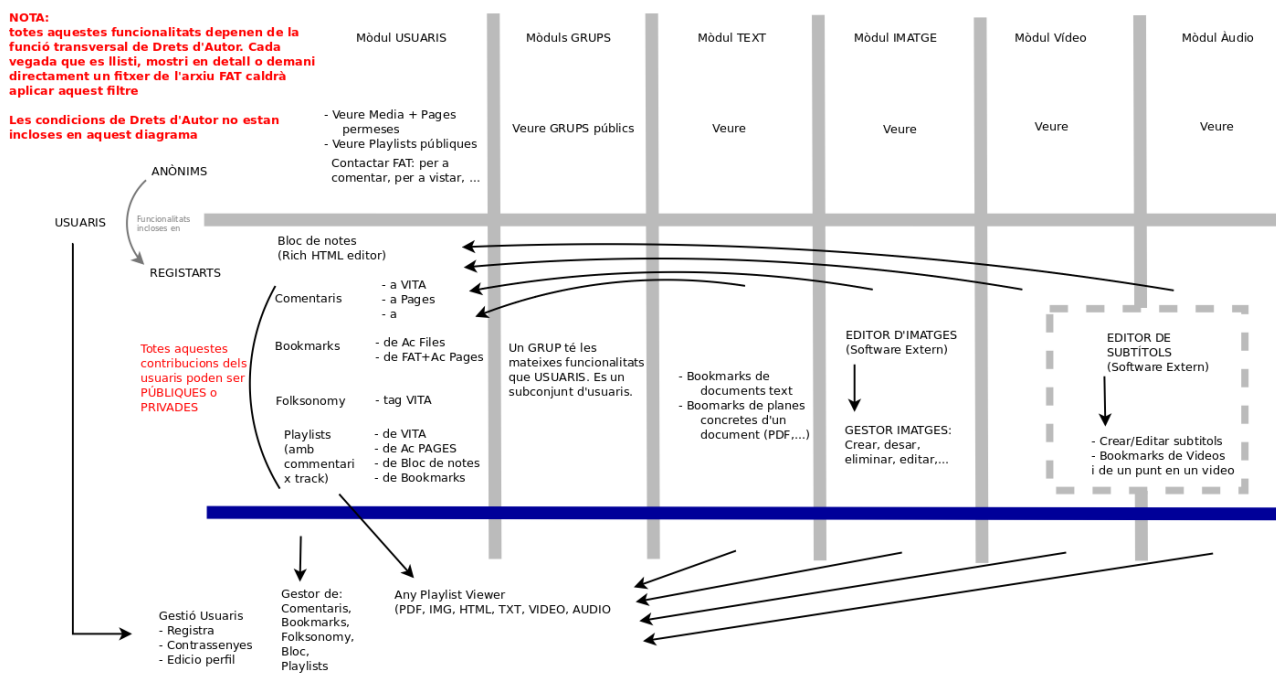


Figura 5: Esquema de treball de funcionalitats de la plataforma *Arts combinatòries* (setembre 2010)

Cada usuari registrat tindrà un espai personal privat a on gestionar:

- **Favorites:** qualsevol plana o document de l'arxiu que pugui ser interessant recordar o tenir-hi accés directe.
- **Bloc de notes:** un editor de text (HTML) per a crear documents.
- Una llista de **paraules clau** per a classificar els continguts de l'espai personal i del propi arxiu *Arts combinatòries*.
- Creació i gestió de **l·listes de reproducció multimèdia:** es podran crear l·listes amb qualsevol document de l'arxiu. Pensants per a presentacions, conferències o recorreguts per l'arxiu.
- **Comentaris adherits** a documents de l'arxiu o a cada ítem d'una llista de reproducció.

Tot el material creat per usuaris podrà ser públic o privat.

Els usuaris registrats podran crear **grups de treball** amb similars funcionalitats que el propi espai personal, però compartit amb els membres del grup.

Seguint amb la filosofia de **reusabilitat** de la feina feta -un principi intrínsec al programari lliure i, per extensió, al coneixement lliure-, vam decidir basar tot el front-end en un sistema de gestió de continguts (CMS) ja existent i, així desenvolupar cada funcionalitat com a mòduls independents del sistema mare. D'aquesta manera serà fàcil qualsevol desenvolupament o adaptació futura de la plataforma. També, aquesta decisió, té l'avantatge de no dependre d'un sol desenvolupador. La institució podrà constinuar innovant i fent créixer el sistema treballant amb qualsevol grup de programadors/es que desitgi.

El món dels CMS, però, és un món molt ampli i divers. S'hi pot trobar des d'eines tant populars com Wordpress (1# en el món dels blocs), fins a eines corporatives (com Liferay) passant per eines per a comunitats o petites corporacions (com Drupal o Joomla). En el nostre cas cercàvem una eina que tingués en compte el compromís entre un sistema molt popular però potser no prou fiable o extensible, i un sistema molt ferm i fiable però potser amb una comunitat petita o d'accés difícil per a usuaris no corporatius (cas del món Java ). En aquesta recerca, finalment, hem apostat per un projecte 100% lliure (GNU General Public License -GPL <sup>10</sup>) i amb la fiabilitat dels sistemes més professionals; estic parlant de **Plone** <sup>11</sup>, una programari que va néixer el 1999, que tècnicament està basat en Zope <sup>12</sup> i programat en Python <sup>13</sup>). Plone és àmpliament usat per governs i ministeris públics acomplint tots els requisits necessaris. Al mateix temps, la gran popularitat de Python com a llenguatge de programació fa que hi hagi molta gent capaç de desenvolupar per a Plone.

És molt important escollir amb quines tecnologies *ens casem* a l'hora de començar un projecte que aspira a tenir futur.

## 4 Visualització i intercanvi de dades

La innovació en interfícies entre persones i dades a internet és lenta inevitablement, doncs el llenguatge visual que podem donar per a entès per una majoria és molt limitat.

Potser és per això que ens trobem molts projectes molt avançats en la capa de dades, amb l'ús d'ontologies i formats estàndards i, en canvi, no impressionen quan en visitem la web perquè usen elements clàssics de presentació de dades, com ara:

- Escasos o cap indicador de les mides del que es té al davant.
- Llistats 1D, uni-dimensionals, tipus text dels quals l'usuari, normalment, no percep el perquè de l'ordre.
- Separació entre cerca i cerca avançada.
- Ús de llargs menús desplegable per a seleccionar.
- Cap indicador del pes de la nostra recerca en tot l'arxiu.

---

<sup>10</sup> Més info: <http://www.gnu.org/>

<sup>11</sup> Més informació sobre PLONE: <http://plone.org>

<sup>12</sup> Més info: <http://www.zope.org/>

<sup>13</sup> Més informació: <http://www.python.org/>



La decisió ha estat donar també l'accés clàssic amb el repte d'atrevir-se a usar visualitzacions de dades inspiradores i ser entesos/es.

## API per a visualització de dades

La plataforma ha estat dissenyada per a poder construir-hi aplicacions -mòduls- que demanen dades i les representen com volen i a on volen. Per això la plataforma permetrà veure resultats de recerca i navegar l'arxiu de maneres diverses. I qualsevol podrà fer noves aplicacions amb noves propostes.

Això serà possible gràcies a una **API** per a accedir a les **dades**. Aquesta API permetrà que un programa demani les dades de l'arxiu que es desitgen i, així, poder representar-les com es vulgui. UPCnet **definirà i documentarà** la creació de mòduls de visualització usant una API amb clau d'usuari tant per a ús interns, com extern.

Aquesta API també permetrà compartir les entrades de l'arxiu de la FAT amb altres arxius, com **Europeana**<sup>14</sup>, usant protocols estàndards com **OAI-PMH**<sup>15</sup>.

Entre les visualitzacions que estem preparant hi ha visualització del model informacional de l'arxiu, es a dir de l'ontologia amb mides de cada classe proporcionals al número d'entrades de l'arxiu.

## 5. Un servei més: decodiFAT

La cirereta del projecte serà un projecte independent de la plataforma fet amb la Fundació I2CAT<sup>16</sup>, decodiFAT: un motor per a donar serveis en-línia on-demand de codificació de video. DecodiFAT serà una eina que les insitucions amb certs recursos en servidors i connexió poden oferir a les comunitats educatives, artistiques i, per extenssió, a tota persona que ho necessiti.

Canviar els videos de format i de qualitat i publicar-los a internet amb la seguretat que la gran majoria de màquines el mostrin correctament, no és una feina trivial. Amb el decodiFAT és podrà sol·licitar versions amb formats OGG i WebM tant amb baixa qualitat com en mitja/alta.

En el cas de la plataforma Arts combinatòria, I2CAT desenvolupa una versió integrada en la plataforma Plone i una versió independent (*standalone*) que podrà ser descarregada i usada lliurement.

---

14 Més info: <http://www.europeana.eu>

15 Més info: <http://www.openarchives.org/pmh/>

16 Fundació I2CAT: <http://www.i2cat.net/ca>

## 6. Nota final

Si ens mirem aquest projecte d'aquí a pocs anys, es probable que el veiem antic, inútil o, fins i tot, ridícul. El temps, en tecnologia, corre molt de pressa; es pot considerar una obsolescència creixent positiva.

El món dels arxius digitals, semàntics i interconnectats tot just comença; en sabem molt poc de tot plegat. En Arts combinatòria s'ha volgut arriscar lo just; ser una font d'inspiració, un projecte amb futur, una eina bastant usada, un punt en un timeline, són somnis que em farien feliç. Tothom hem après moltes coses diferents durant aquest projecte, i a mi, això, ja m'ha especialment anat bé, doncs sóc dels que miren de saber poc de molt, més que molt de poc.



*Figura 6: Aby Warburg, Mnemosyne-Atlas, Nr.32  
1926 - Photography.*

A partir d'aquí, el repte comença i el joc d'imaginar nous mons continua preservant-se en el temps. Veiem i veurem moltes provatures respecte a com gestionem i usem el material contingut als nostres arxius. Cal que sapiguem entendre què en volem treure, què en volem aprendre. O que en voldran extreure en el futur. M'agradaria pensar que hem dissenyat un sistema que permetrà saltar cap a l'esglaó següent fàcilment. Un sistema preparat per a ser canviat completament. Arts combinatòries com un camí proposat, no pas com un programari acabat.

Jaume Nualart  
[jaume@nualart.cat](mailto:jaume@nualart.cat)  
@jaumetet