

El Taller Digital de la Universitat d'Alacant

bressol de biblioteques digitals

Francesca MARÍ DOMENECH, Laura SÁNCHEZ MENASANCH

RESUM: *El Taller Digital de la Universitat d'Alacant és una empresa dedicada a organitzar i presentar informació digital. Ha desenvolupat dues biblioteques virtuals reconegudes, la Biblioteca Virtual Miguel de Cervantes i la Biblioteca Virtual Lluís Vives. S'explica com s'han concebut, evolucionat i executat els dos projectes, que s'han convertit en una mostra de conciliació entre les biblioteques tradicionals i el món digital. L'evolució tecnològica pel que fa a l'accés a la informació a través d'Internet obliga, tant les empreses com les institucions, a investigar i crear noves formes de presentar-la. La tasca hui dia és la d'oferir, de manera eficient, un sistema de visualització compatible amb les eines que s'estan utilitzant en l'anomenada societat de la informació.*

MOTS CLAU: projectes de digitalització; digitalització; empreses de serveis documentals; biblioteques digitals; Biblioteca Virtual Miguel de Cervantes; Biblioteca Virtual Lluís Vives; Universitat d'Alacant. Taller Digital

EL TALLER I ELS SEUS PROJECTES

El Taller Digital de la Universitat d'Alacant <<http://www.eltallerdigital.com>> és una empresa de base tecnològica creada el 2004 per la Universitat d'Alacant a partir del projecte d'investigació anomenat Biblioteca Virtual Miguel de Cervantes <<http://www.cervantesvirtual.com>>, que es va iniciar en 1999. Actualment, la Biblioteca Virtual Miguel de Cervantes és considerada la biblioteca virtual de referència universal en llengua castellana; disposa d'un catàleg que supera els 18.000 títols i ha sigut consultada en més de 240 milions d'ocasions.

Francesca Marí Domenech. Directora de Producció. servicios@eltallerdigital.com

Laura Sánchez Menasanch. Gerente. gerencia@eltallergigital.com

Universitat d'Alacant. Taller Digital. <http://www.eltallerdigital.com>

Article rebut el 19 de gener de 2006.



Lloc web del Taller Digital <<http://www.eltallerdigital.com/>>

L'objectiu principal amb el qual naix el Taller Digital de la Universitat d'Alacant és el de col·laborar d'una manera eficient amb les institucions o les empreses que desitgen fer una millora contínua i innovadora en la prestació dels seus serveis a la societat o al mercat en general, i apropar-los a l'usuari amb criteris d'innovació, qualitat, adaptabilitat i garantia.

La voluntat d'aquest article és la de fer un repàs històric i metodològic de l'inici, execució i posterior enlairament de la Biblioteca Virtual Miguel de Cervantes i de la Biblioteca Virtual Lluís Vives.

La Biblioteca Virtual Miguel de Cervantes

La Biblioteca Virtual Miguel de Cervantes, inaugurada al juliol de 1999, és un projecte d'edició digital del patrimoni bibliogràfic, documental i crític de la cultura tant espanyola com hispanoamericana.

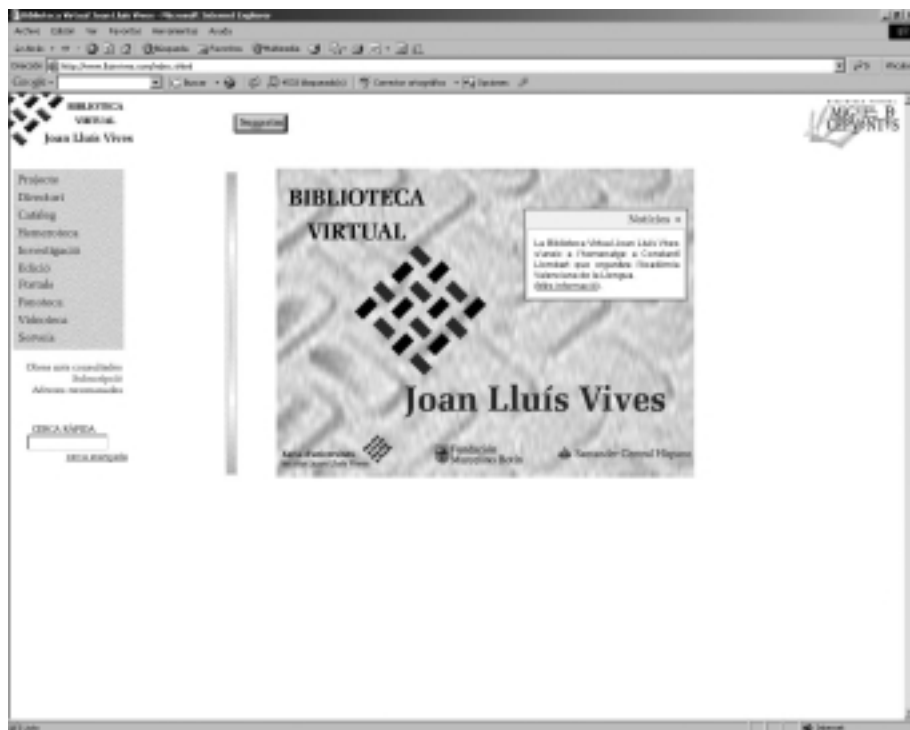
La iniciativa de la Universitat d'Alacant i el Banco Santander Central Hispano, amb el suport de la Fundación Marcelino Botín, intenta pro-

moure el sumatori d'esforços i projectes d'altres institucions, tan públiques com privades, interessades en l'estudi i difusió de la cultura per tot arreu.

La idea final és la de posar a l'abast de tot el món les obres i els documents més importants, mitjançant la utilització de les noves tecnologies, i està destinada a qui vulga investigar la cultura espanyola i hispanoamericana, específicament en els aspectes literaris i històrics.

Tant el rigor científic com l'operativitat de la seua difusió electrònica és responsabilitat d'un equip d'investigadors universitaris, oberts a la col·laboració i a les innovacions que en aquest camp es fan cada dia. La Biblioteca Virtual Miguel de Cervantes és un projecte internacional atès l'origen dels seus fons i les entitats col·laboradores, així com pel seu caràcter de difusió universal.





Biblioteca Virtual Joan Lluís Vives <http://www.lluivives.com/index.shtml>

La Biblioteca Virtual Joan Lluís Vives

La Biblioteca Virtual Joan Lluís Vives és fruit d'un conveni entre la Fundació Biblioteca Virtual Miguel de Cervantes i la Xarxa d'Universitats que formen l'Institut Joan Lluís Vives. Amb aquest acord es pretén, d'una banda, aprofitar, potenciar i difondre al màxim els fons digitalitzats existents a la xarxa i, d'una altra, proposar i fer efectiva l'edició digital de les obres més representatives de la cultura catalana, valenciana i balear.

DESENVOLUPAMENT D'UN PROJECTE DE DIGITALITZACIÓ

El Taller Digital de la Universitat d'Alacant ha desenvolupat un mètode de treball que rendibilitza al màxim el nostre procés de producció. Donarem ací, a grans trets, el procés que hem seguit per a dos projectes com la Biblioteca Virtual Miguel de Cervantes i la Joan Lluís Vives.

La selecció de materials

La primera fase del procés de producció és la localització i selecció de materials que seran migrats a formats digitals. Parlem de *materials* perquè la documentació no és únicament i exclusiva textos en paper, sinó tot tipus de suports: vídeos, so, imatges, etc. La selecció de materials la fan les institucions que mitjançant un conveni decideixen incloure documentació valuosa a les nostres biblioteques.

Gestió dels drets de publicació

És necessari comptar amb les escaients autoritzacions dels autors del documents, per això se sol·liciten el permisos als titulars del drets d'autor morts després de 1935, *post mortem auctoris* (les obres dels morts setanta anys abans de l'any en curs passen al domini públic). Així mateix també es gestionaran les autoritzacions *inter vivos*.

En la legislació espanyola (la seu de la Biblioteca és a Espanya) els drets d'autor es troben protegits pel text refós de la Llei de Propietat Intel·lectual (RDL 1/1996 del 12 d'abril) i per la Llei 5/1998 de 6 de març. Quant als autors estrangers, es regeix per tractats i convenis internacionals, com el Conveni de Berna o la Convenció Universal de Ginebra.

És important, abans de començar qualsevol treball de digitalització, haver tramès tota la documentació legal per a poder-ho fer. Hem de recordar sempre que no es fa una cessió de drets d'autor, sinó un permís de publicació i transmissió pública.

Transmissió de materials al centre de producció

Després d'aconseguir les autoritzacions escaients, els documents poden ser tramesos al centre de producció, en aquest cas al Taller Digital perquè passe a edicions digitals; però hi ha casos en què, a petició de les institucions, la digitalització es fa *in situ* perquè el material és valuós o per causa del seu estat de conservació. La manera de procedir en aquestos casos és la següent: es desplaça una cèl·lula de digitalització al lloc on es troben els documents i es digitalitzen amb càmera digital o escàner tradicional segons les característiques físiques dels documents.

Catalogació. No podem oblidar que una biblioteca digital no ha de perdre mai el nord, i és que s'han de catalogar i registrar els textos de la mateixa manera que es fa a les biblioteques tradicionals. La catalogació que es fa és una descripció bibliogràfica del document, amb la finalitat d'oferir-ne una imatge sintètica i permetre'n la recuperació des dels catàlegs i cercadors.

L'anàlisi documental de la catalogació se centra fonamentalment en un curiós tractament dels punts d'accés, és a dir, autors, títols i matèries, que són els elements clau per a les recerques. Els punts d'accés es tracten com fitxers d'autoritats on se'n normalitza la forma i la redacció d'acord amb les normes de descripció internacionals International Standard Bibliographic Description (ISBD). Per a la redacció de matèries, l'encarregat de la catalogació fa un treball d'anàlisi i síntesi del contingut amb el propòsit d'establir els temes per als quals el document és d'interès.

Els registres catalogats s'integren en una base de dades que està en SQL-Server i que suporta l'estàndard MARC en la versió MARC 21, que és l'estàndard internacional de catalogació més universal i el que el Taller Digital ha triat per tal de gestionar els continguts tant de la Biblioteca Virtual Miguel de Cervantes com de la Biblioteca Virtual Joan Lluís Vives.

La base de dades que hem creat permet generar altres estàndards gràcies a l'exportació de dades mitjançant programes de creació pròpia. Com a exemple podem dir que transmet algunes dades que s'utilitzen en les edicions textuals obtingudes després del marcatge estructural en XML; en concret, en la generació del TEI-header, un altre estàndard de catalogació de gran importància en les biblioteques digitals.

Edició

Una vegada finalitzat el procés de catalogació, la labor de producció de continguts es dividirà en dues vessants, directament relacionades amb els formats d'edició. D'una banda, els materials que seran publicats en text i imatge (edició facsímil). I per una altra, els que acabaran convertint-se en arxius d'àudio (de la fonoteca) o audiovisuals (de la videoteca).

Digitalització

En lloc de la conservació habitual, la conversió a formats digitals dels suports tradicionals (impresos, fotografies, llibres, etc.) ofereix innumbrables avan-

tatges en la gestió i la recuperació de la informació, redueix la necessitat de destinar espai a l'emmagatzematge físic i facilita la transferència i la difusió dels arxius.

L'àmplia experiència del nostre servei, així com el nostre nivell tecnològic, ens permeten tractar la digitalització de qualsevol suport i oferir la possibilitat de diferents tractaments digitals d'imatge (restauració virtual de documents deteriorats, reconstrucció digital d'originals amb dificultats de llegibilitat, etc.) i obtenir un producte de gran qualitat

Tret del cas especial indicat en l'apartat de transmissió de materials al centre de producció, el procés de digitalització es portarà a terme a les instal·lacions del Taller Digital a la Universitat d'Alacant. El procediment triat per al tractament dependrà bé de l'estat del material, bé de l'expressa decisió de publicarlo en un determinat format, o bé de la valoració d'ambdues circumstàncies.

En relació amb l'estat del material, tindrem en compte el tipus de lletra (tant en manuscrits com en obres impreses), el tipus i el color del paper, l'estat de conservació i l'antiguitat, el valor i la qualitat de l'original.

El procés de digitalització es farà mitjançant diferents tipus d'escàners:

- *Escàner pla o de taula*: és el tipus més conegut. S'utilitza de manera semblant a una petita fotocopiadora: el document es col·loca sobre un vidre sota el qual es desplaça la lent lluminosa que digitalitza el document.



- *Escàner de trajectòria aèria*: permet digitalitzar volums enquadernats amb les fulles cap amunt gràcies a que la font de llum i el sensor CCD (dispositiu que converteix la llum rebuda en informació analògica) es troben acoblats a un braç de trajectòria aèria; és apte per als llibres molt gruixuts o amb una enquadernació molt delicada.
- *Escàner microfilmadora*: es tracta d'un tipus especial que s'utilitza per a la digitalització de pel·lícules en rotllo i microfites; la seva qualitat depèn, sobretot, de l'estat en el qual es trobe el microfilm.
- *Escàner de mà*: escàner generalment de reduïdes dimensions, que el mateix usuari ha de passar sobre la zona a digitalitzar. És poc fiable, ja que depèn de l'habilitat i el pols del digitalitzador.
- *Escàner de diapositives*: és un escàner especial amb doble vidre, que serveix també per a digitalitzar tot tipus de document.
- *Càmera digital*: la càmera digital s'adapta a una àmplia varietat de documents i objectes; permet capturar de manera segura fins i tot els materials més fràgils.

El tractament del material digitalitzat dependrà del seu format d'edició. En aquest sentit diferenciarem la digitalització de textos i la d'imatges.

Per al primer dels casos, utilitzarem tecnologia d'OCR (sistema de reconeixement òptic de caràcters impresos), a través d'un programa comercial que convertirà la imatge digitalitzada en un arxiu de text editable que podrà ser utilitzat per qualsevol processador de textos.

El segon dels casos citats és la digitalització d'imatges, enfocada a la producció d'edicions facsímils. Aquestes publicacions es fan o bé quan no es pot aplicar a un escrit la tecnologia d'OCR (per les característiques físiques del paper, per la seva tipografia irregular, per la seva baixa qualitat d'impressió, per l'ús d'un vocabulari antic), o bé quan l'aspecte original és un element de valor afegit de l'obra que es digitalitzar. El procediment de digitalització d'imatges orientat a la producció d'edicions facsímils és considerablement diferent del descrit per a les edicions en versió textual.

Tant per a l'escanejat com per a l'adequació de les imatges, el Taller Digital utilitza programes comercials. El procés està delimitat en quatre fases diferents: la planificació de la digitalització; la digitalització de les imatges; el tractament dels arxius amb l'objectiu fonamental de la fidelitat als originals i l'enregistrament.

Correcció i edició

Després de la utilització de l'OCR per a les obres que seran publicades en format text, el següent pas en la cadena productiva és la correcció tant de les errades d'interpretació generades pel programa de reconeixement òptic de caràcters com de les errates de l'edició original. Aquesta tasca, una de les més laborioses del procés, és feta pels editors acarant la versió digitalitzada amb l'obra en format «físic». Requereix alts graus d'atenció i concentració. Amb aquestes edicions es pretén assolir un model que conserve tant les característiques de l'edició digital com els tipus més representatius de l'edició tradicional impresa, de manera que planteja i utilitza diferents criteris i normes de caire crítica, fonètica i paleogràfica. Per a les edicions crítiques, amb el seu aparat de variants o la fixació textual d'un editor, compta amb edicions externes, és a dir, amb textos aportats per especialistes. I per a la resta de casos, ha establert uns criteris de correcció i edició operatius per a aplicar-los a un corpus de textos molt extens pel que fa a àmbit cronològic i geogràfic.

Edició facsímil

Davant de la necessitat de reproduir en format imatge diversos originals, la tecnologia ofereix la possibilitat de fer rèpliques digitalitzades fidedignes (d'un manuscrit, una edició valuosa, una pintura, etc.) o d'altres retocades (d'una fotografia antiga mal conservada, un llibre antic amb taques que en dificulten la lectura, etc.). Aquest nou concepte d'edició facsímil adaptat a les novetats tecnològiques no solament no perd rigor i qualitat, sinó que guanya en maniobrabilitat, riquesa de continguts i processos de conservació, cosa que ens ha dut a aprofundir en el procediment d'edició en suport digital. L'edició d'obres en format imatge també segueix les pautes marcades per uns criteris d'edi-

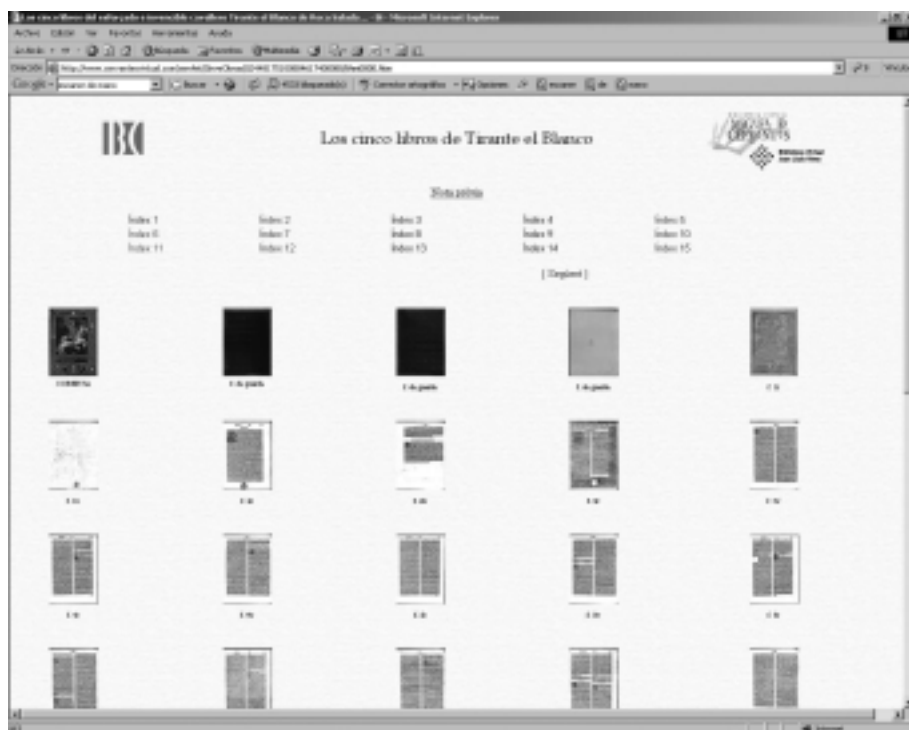


Coberta d'un llibre digitalitzat a la BVMC

ció facsímil, que regulen principalment la reproducció de manuscrits i primeres edicions. La falta de l'objecte llibre comporta la necessitat d'haver de crear un motlle per a publicar les imatges dintre d'una estructura organitzada, respectant i corregint a un temps cada errada de paginació o foliació per a ser fidel a l'original sense entorpir la consulta del text. S'ha de respectar així l'estructura de l'obra i l'ordre dels folis mitjançant un sistema d'índexs de foliació basat en l'estructura interna de l'obra, índexs de miniatures i ampliació de cadascun dels folis. S'inclou també una nota prèvia on es consignen les dades relatives a l'exemplar reproduït i es descriu el procés digital que s'ha seguit amb les imatges, així com qualsevol dada que pugui afegir informació a l'usuari.

Control de qualitat

L'última fase de producció tant de les edicions en text com de les edicions en imatge és el control de qualitat. En les primeres, la generació automà-



Pàgina web de la BVMC amb les pàgines del llibre digitalitzat, en format petit

tica redueix la possibilitat que existeixen errors. Major atenció necessiten les edicions en format text. Aquestes són sotmeses a diverses proves que, posteriorment, permeten el perfeccionament del procés de digitalització i generen idees que poden ser aprofitades de cara al desenvolupament i la integració d'altres eines que facilitaran les tasques de correcció i edició.

Publicació

Una vegada finalitzat el procés de control de qualitat, es procedeix a la publicació dels materials en la web.

En el cas de les edicions textuais, cada obra es correspon amb un fitxer XML, que és sotmés a un doble tractament de producció automàtica. En primer lloc, mitjançant un full d'estil *XSL és convertit en un fitxer HTML únic (amb marques de format especials). I, a continuació, és transformat en llibre digital gràcies a l'execució d'un programa de creació pròpia que genera un llibre electrònic. Una vegada «construïts» els continguts digitals, són generats com a «novetat bibliogràfica», és a dir, són publicats en el web.

La publicació d'edicions facsímils segueix un procediment similar, encara que considerablement més senzill. Aquest tipus d'obres no requereix de cap procés de postproducció. El programa de generació automàtica deixa els continguts llestos per a ser publicats.

Edició d'àudios i audiovisuals

El segon tipus d'edicions fa referència als formats de so i vídeo. El procés productiu d'aquestes obres és radicalment diferent al de les textuais i facsímils. Consisteix fonamentalment en cinc fases consecutives: producció, enregistrament, edició, postproducció i publicació.

Edició de webs

A més de poder ser recuperades des del catàleg i els cercadors habilitats a aquest efecte, les obres podran ser consultades des d'agrupacions temàtiques de continguts, que tindran la forma genèrica de web.

S'elabora i es presenta el pla del web i se li proporcionen els materials que hi hauran d'anar, així com l'URL que tindrà el web en Internet. Tenint en compte tota aquesta informació, es fixa l'estructura del web que anirà enllaçant automàticament, gràcies a la creació de sèries del catàleg, totes les obres i tots els materials que l'integren.

Control d'ús i accessibilitat

La producció de continguts digitals ha de tenir sempre present les normes fonamentals d'ús i accessibilitat, ja que d'ambdós factors depén, en gran mesura, la seva utilitat real en el context de la xarxa.

Una obra i, en sentit més ampli, una web són accessibles quan estan dissenyades i codificades perquè els seus continguts i serveis estiguen disponibles per a qualsevol persona, amb independència de quin siga el seu context de navegació.

Per a assolir aquest objectiu s'utilitzen estàndards que ajuden a identificar i resoldre problemes d'accessibilitat. La Iniciativa d'Accessibilitat a la Web (WAI) proporciona una sèrie de catorze pautes generals que són considerades en la Unió Europea com les normes *de facto*.

En el Taller Digital vam seguir les pautes d'accessibilitat per a desenvolupar les pàgines de les nostres biblioteques i la resta de totes les pàgines. A més, vam proporcionar un disseny accessible a través d'eines de navegació.

S'entén per control d'ús el conjunt de tècniques que ajuden els usuaris a navegar i interactuar de manera senzilla i segura en el web. El control d'ús de les seves pàgines es controlarà mitjançant una sèrie de tests consistents bàsicament en l'observació de l'evolució d'un usuari davant una sèrie de tasques i cerques pel lloc. L'objectiu és arribar a esbrinar els principals problemes d'ús d'aquest lloc virtual, detectant errades de disseny. Les proves efectuades han permés introduir en els llocs web de les biblioteques virtuals Miguel de Cervantes i Lluís Vives millores que n'han afavorit l'ús. Heus ací algunes mostres: les visites guiades (en les quals s'explica com utilitzar la Biblioteca), la possibilitat de convertir aquesta web en la pàgina d'inici del navegador de l'internauta, l'opció de col·locar un enllaç a la Biblioteca en la llista de «millors», personalització del lloc d'acord amb les preferències de l'usuari, la subscripció a butlletins periòdics d'informació general o segmentada i, finalment, el desenvolupament d'un sistema de consultes, centralitzat en una pàgina anomenada El Bibliotecari, que promou l'atenció dels administradors de continguts pels usuaris.

Manteniment

Els servidors de producció i web de la Biblioteca Virtual Miguel de Cervantes necessiten ser actualitzats contínuament. Aquesta labor de manteniment suposa la realització de diverses tasques: l'actualització del sistema operatiu i del sistema de seguretat; el servidor de correu, les bases de dades, etc.

Preservació

La preservació implica la realització de còpies de seguretat i la seva conservació en llocs adequats. Des del Taller Digital s'ha estipulat un protocol de còpia de dades dels servidors de producció i dels servidors web, amb el propòsit de poder recuperar-los davant qualsevol errada humana o tècnica.

Eines de cerca i valoració

Aquesta és, a grans trets, l'organització del treball per a mantindre des del Taller Digital els webs de les dues biblioteques. L'èxit d'aquestes biblioteques no és degut solament a la qualitat dels materials que s'ofereixen, triats per les institucions que proporcionen tota la documentació, i a l'estreta col·laboració d'un extens equip d'investigadors universitaris i tècnics, vinculats a la Universitat d'Alacant –que formen un grup interdisciplinari d'innovació permanent en la direcció d'àrees relacionades amb el desenvolupament tecnològic en informàtica, filologia, humanitats i biblioteconomia i documentació, entre altres disciplines–, sinó també al sistema de cerca de les dues pàgines que permet fer cerques ràpides o més específiques.

En el motor de cerca podem trobar l'anàlisi i descripció dels objectes digitals (documents de text, àudio, vídeo, imatge, etc.), recuperar per descriptors, paraules clau, combinar camps o paraules amb operacions booleans.

Ambdues pàgines tenen una organització molt intuïtiva que respon a les cerques i camins lògics que un investigador i el públic en general seguiria per a trobar un document. Les dues biblioteques consten d'una capçalera on podem trobar totes les opcions generals: visites guiades, tipus de cerques i ajuda; en segon lloc la pàgina principal té suggeriments, últimes notícies i un apartat amb nous projectes perquè les novetats més importants es puguin trobar de seguida; després considerem que una biblioteca ha de tindre en primera instància un catàleg, organitzat en catàleg general per ordre alfabètic i

d'autors i després una sèrie de catàlegs secundaris classificats per format de publicació: així trobem un catàleg de vídeos, d'àudios, d'edicions facsímils, edicions multimèdia, arxius, articles crítics, hemeroteca i tesis doctorals. Arribem després a les seccions d'investigació, eines lingüístiques, Biblioteca de Signes i portals temàtics i institucionals, i finalment trobem un apartat de servicis a l'usuari.

L'activitat del Taller Digital de la Universitat d'Alacant se centra de manera fonamental en la producció de continguts digitals en nombrosos formats, en el desenvolupament d'eines i entorns informàtics que gestionen grans masses d'informació i en l'anàlisi de demandes dels continguts en la xarxa, ja que, per causa de l'origen de la nostra tasca, som especialistes en la gestió de documents d'investigació. És una feina que requereix uns criteris que constitueixen els valors i l'estil de funcionament de l'empresa: treball en equip, adaptabilitat i millora contínua, responsabilitat i garantia davant els compromisos. El manteniment d'aquestes dues biblioteques virtuals n'és un exemple, a banda de representar una reconeguda contribució al món de la cultura.



Instal·lacions del Taller Digital, a la Universitat d'Alacant