

Recursos para a clasificación da produción editorial na Galiza durante a etapa franquista: deseño e alimentación da base de datos¹

CARME FERNÁNDEZ PÉREZ-SANJULIÁN
M^a ANTONIA PÉREZ RODRÍGUEZ
EDUARDO RODRÍGUEZ LÓPEZ
ÁNXELES SAAVEDRA PLACES
Universidade da Coruña
carme@udc.es

1. INTRODUCCIÓN

O proxecto multidisciplinar que presentamos é o resultado da colaboración entre especialistas en Filoloxía, Documentación, Historia e Informática; neste sentido, para alén do interese que poda suscitar polos resultados intrínsecos da propia pesquisa, representa un exemplo paradigmático de como as Humanidades Dixitais nos levan a reflexionar sobre as abordaxes metodolóxicas clásicas no campo das ciencias humanas e, ao tempo, a xerar novos modelos de investigación máis adaptados aos novos contornos dixitalizados de traballo. Estas propostas metodolóxicas permiten a descuberta e a análise eficiente de elementos culturais e patrimoniais que moi dificilmente (ou, ao menos, de xeito moito máis laborioso) poderían ser enfrontados desde a óptica clásica das Humanidades.

Así, para alén de desenvolver a metodoloxía seguida en concreto no proxecto sobre a produción editorial na Galiza durante a etapa franquista e os resultados acadados, revisaremos os erros e acertos que tivemos no curso da investigación, coa vontade de contribuír á reflexión sobre os problemas que xorden neste novo marco de investigación, na procura do desenvolvemento e/ou estandarización de novas metodoloxías de traballo no ámbito das Humanidades Dixitais.

Iniciaremos a nosa exposición cunha presentación dos obxectivos e liñas xerais do noso proxecto de investigación para, a seguir, explicarmos a metodoloxía e os criterios que se utilizaron para desenvolver o plano de traballo. No seguinte

¹ Este traballo está realizado ao abeiro de dous proxectos complementares: "La producción editorial en Galicia durante la etapa franquista (1939-1975)" (FFI2010-16924) e "Producción e mercado editorial en Galiza durante a etapa franquista: a súa incidencia na industria editorial actual" (10SEC104017PR).

apartado centrarémonos na explicación das ferramentas informáticas que se desenvolveron especificamente; no último punto analizaranse os problemas atopados no seu desenvolvemento e as solucións adoptadas para, finalmente, presentarmos as conclusións.

2. PRESENTACIÓN DO PROXECTO

O obxectivo fundamental do noso proxecto de investigación consiste no estudo da produción editorial na Galiza durante a etapa franquista. Como no resto do Estado, esta actividade veuse condicionada e limitada pola existencia dun control político que considerou como esencialmente perigosa toda a produción cultural que non respondese ao seu ideario, de modo especial a que puidese defender ou promover as daquela denominadas culturas “vernáculos” (sobre todo se se editaba en linguas diferentes ao español), xa que, neste caso, esa actividade era vista directamente como sospeitosa de separatismo.

Con todo, a pesar destas limitacións e da difícil situación económica do país, o labor editorial non foi testemuñal. Foron numerosos os proxectos editoriais que xurdiron nesta etapa na Galiza, algúns de longa duración (que chegan, mesmo, até a actualidade), outros fugaces, como pode ser o caso da colección Benito Soto, mais todos certamente importantes na historia das nosas letras. Existiron mecenas, diferentes plantexamentos editoriais e comerciais, variados modos de conseguir un público lector, e opcións estéticas e ideolóxicas diferenciadas. Contemplada no seu conxunto, toda esa actividade editorial cumpriu un papel significativo na resistencia cultural que foi chave na loita antifranquista e na procura dunha sociedade libre e democrática²; esta circunstancia non foi, na nosa opinión, suficientemente valorada até agora e, de feito, estamos a comprobar o profundo descoñecemento que existe na actualidade de moitas destas iniciativas editoriais, mesmo dalgúns de indiscutíbel suceso comercial³.

No campo da literatura galega, a historiografía literaria deste período consolidou a idea da centralidade cultural e político-ideolóxica dalgunha desas editoras (debido en parte ao número de autores publicados e á incontestábel calidade literaria dalgúns dos textos alí editados). Porén, a nosa hipótese é que outros proxectos tiveron un peso moi importante na revitalización da cultura galega, nuns casos pola orixinal liña editorial desenvolvida (pénsese nas iniciativas de Xosé M^a Álvarez Blázquez), noutros polo perfil dalgún mecenas (como Isaac Díaz Pardo e Luís Seoane) ou, tamén, pola actividade levada adiante por autores que funcionaron como editores e dinamizadores culturais (como, entre outros, os poetas Manuel María ou Celso Emilio Ferreiro), sen esquecer o importante papel que tivo a autoedición para moitos dos títulos desta época e o apoio constante de imprentas e talleres de artes gráficas de moitas das vilas do noso país.

² Ao tempo, non ignoramos o extraordinario labor editorial levado a cabo nesta etapa polo exilio galego en diferentes países (sobre todo na Arxentina), nin tampouco iniciativas fundamentais para as nosas letras como as de Akal ou de Brais Pinto en Madrid. Aínda que non formen parte do noso estudo polo seu marco xeográfico, deberán ser analizados tamén nunha fase posterior.

³ Un exemplo sería o da Editorial Cíes de Vigo que, malia o seu suceso comercial, é practicamente inexistente nos andeis das bibliotecas públicas polo seu carácter de literatura “popular”.

O noso obxectivo fundamental foi recoller con vocación de exhaustividade a produción de todo tipo de iniciativas editoriais que tiveran lugar na nosa terra, aínda que parte delas non fosen literarias (velaí o exemplo da editora técnica Valle-Collantes), sen esquecermos o fundamental labor editorial de moitas das entidades públicas do momento (Academias, Universidade, museos...). Cómpre ter en conta que unha parte central da nosa investigación é o estudo das empresas que editaron en galego durante ese período (as súas características, produción, catálogos...); porén, para podermos levala adiante, só unha pormenorizada descrición do mundo da edición na Galiza daqueles anos vai permitir desenvolver con precisión a análise de moitas das circunstancias e condicionamentos en que se retoma a escrita e a edición en galego durante o franquismo na Galiza.

Na nosa opinión, neste período atópanse moitas das chaves que explican a evolución da industria editorial galega durante o posfranquismo e, aínda, en etapas máis recentes ou na actualidade. A análise en profundidade do mundo editorial, desde unha perspectiva histórica e multidisciplinar, pode ser unha fonte de ideas para o desenvolvemento de estratexias comerciais de interese no presente para este sector económico na Galiza.

3. METODOLOXÍA PARA A RECUPERACIÓN E CATALOGACIÓN DA PRODUCCIÓN EDITORIAL GALEGA (1939-1975)

A novidade do proxecto de investigación que estamos a levar adiante estriba en que pretende, por primeira vez, enfrontar a análise de todos os proxectos editoriais da Galiza da posguerra individualmente, tanto nos seus aspectos ideolóxico-culturais como comerciais, para, a seguir, comparalos entre si. Pretendemos analizar a súa relevancia na recuperación e posta en valor da cultura galega e o seu papel na resistencia cultural antifranquista durante esa etapa nos casos en que resulte pertinente.

Para cumprirmos estes obxectivos estamos a levar a cabo diversas tarefas, a máis laboriosa das cales está a ser a busca e recuperación de información da produción editorial galega entre 1939-1975, coa relación de editoriais e impresas, delimitación do lugar de edición, relación de títulos, de autores/as, etc. Este labor de fixación do catálogo resulta imprescindíbel para podermos desenvolver axeitadamente a caracterización e análise dos diferentes proxectos editoriais (equipos editoriais, coleccións, liña editorial, características específicas, tiraxes, axudas, vinculacións con outros proxectos empresariais, papel dentro do sistema literario da época, etc.), así como o estudo dos protagonistas e os seus perfís ideolóxicos (biografías, epistolarios, entrevistas...)⁴.

Neste proceso foi fundamental establecer cales ían ser os criterios a seguir na selección dos documentos (neste casos libros/obras) e dos servizos de información onde están depositados. Así, tivemos que escoller a que centros de información, bases de datos, repertorios e outras publicacións acudir para recoller o editado en Galiza entre 1939 e 1975, e que publicacións elixir.

⁴ Coa finalidade de difundir o traballo realizado, contamos con abrir ao público proximamente unha web con todos os datos e materiais recollidos. Aquí inclúiranse, para alén do catálogo e dun útil e complexo buscador, outros documentos como, por exemplo, as entrevistas e filmacións dos protagonistas aínda vivos e dos seus colaboradores directos. O enderezo será: <<http://ediciongalizafranquista.com>>.

Optouse, pois, por combinar varios procedementos; en primeiro lugar, decidimos aproveitar as vantaxes que nos ofrecen as tecnoloxías e os recursos de que hoxe dispomos na sociedade da Información; foi así como comezamos a valorar a posibilidade de recuperar as publicacións editadas neste período que custodian as bibliotecas públicas galegas (ou con fondos galegos) mediante a importación e integración de catálogos dixitalizados. Finalmente, unha vez que nos puxemos en contacto coas persoas responsábeis das bibliotecas, e tras avaliarmos as dificultades e os recursos necesarios para poder levar adiante este plano de traballo, tomouse a decisión de proceder á importación masiva de datos dos catálogos das seguintes bibliotecas públicas: Biblioteca de Galiza e Bibliotecas nodais de Coruña, Lugo, Ourense, Pontevedra, Ferrol, Vigo e Santiago.

En total, coa importación automática destes fondos documentais das bibliotecas, partimos de 22.495 rexistros bibliográficos en formato *MARC 21*. Este número de rexistros era, evidentemente, demasiado elevado e non se correspondía coa cantidade de materiais⁵ que intuíamos podían ter sido publicados naquela etapa.

Cómpre sinalar que, cando decidimos facer a importación masiva, xa eramos conscientes de que ían aparecer moitos rexistros duplicados (por estaren os exemplares repetidos en varias bibliotecas), mais contabamos con que, ao viren en formato *MARC*, sería doada a súa identificación e eliminación automática; daquela, a revisión manual ficaría só para casos problemáticos e, en principio, excepcionais. A sorpresa veu cando constatamos que os rexistros *MARC 21* dunha mesma obra presentaban numerosas diferenzas que impedían a súa identificación automática como duplicados. Esta circunstancia obrigou a un intenso traballo manual (aínda que soportado polas ferramentas informáticas desenvolvidas) de eliminación da redundancia, un labor que ocupou grande parte da actividade do grupo de investigación.

Ao tempo, fíxose a revisión e cotexo doutros materiais bibliográficos que serviron para ampliar a base de datos inicial pois, obviamente, partiamos da base de que toda a produción editorial non se conservaba nas bibliotecas⁶. Foi este un traballo laborioso que nos obrigou a revisar manualmente repertorios bibliográficos (Palau, 1948); catálogos do Instituto Nacional del Libro Español (1957-1961, 1972, 1974, 1978) e do Servicio Nacional de Información Bibliográfica (1958-168), libros e revistas editados dentro e fóra de Galiza que de forma parcial recompilan as nosas publicacións (apéndices de *Cuadernos de Estudios Gallegos*, *Galaxia Almanaque...*) ou os catálogos de diversas mostras bibliográficas feitas en distintas comarcas ou vilas galegas (Alonso y Ramos, 1963; Vilariño y Pardo, 1981; Bassave, 1989; Álvarez, 1995). Nesta fase, obviamente, resultou esencial o cotexo e revisión dos catálogos, nos casos en que se conservan, e da documentación conservada das organizacións editoriais en activo entre 1939 e 1975.

Este traballo, aínda sen rematar por completo, estase a realizar, fundamentalmente, a partir da consulta dos arquivos das propias editoriais estudadas e de bibliotecas galegas con fondos históricos (Real Academia Galega, Biblioteca

⁵ Ao falarmos de rexistros, referímonos a libros, folletos, revistas ou separatas.

⁶ Esta recuperación manual foi especialmente útil, ademais, para localizarmos moitos dos libros publicados antes do establecemento do Depósito Legal (23-XII-1957) e nos primeiros anos en que este se implantou, xa que é notorio que en todo o Estado español se publicaron bastantes libros sen el durante os primeiros anos da entrada en vigor da lei.

Xeral da Universidade de Santiago de Compostela, Instituto de Estudos Galegos Padre Sarmiento, Museo de Pontevedra, Fundación Penzol, Biblioteca da Deputación Provincial de Ourense, Biblioteca do Real Consulado da Coruña, Círculo das Artes de Lugo, Biblioteca da Fundación Manuel María...), sen deixar de lado a consulta de toda a bibliografía e documentación que se está a publicar en datas recentes sobre este tema.

4. O TRABALLO INFORMÁTICO NO PROXECTO

O traballo informático no proxecto estivo centrado en catro tarefas:

1. Construción da base de datos (BD): A partir das prioridades establecidas polos obxectivos do proxecto, deseñouse o modelo conceptual de datos, documentado nun diagrama Entidade-Relación, que recompila toda a información que se precisa na actualidade, dá cabida a ampliacións futuras e asegura a súa explotación eficiente. Este modelo foi finalmente implementado no Sistema Xestor de Bases de Datos (SXBD) PostgreSQL, un software libre que se ten consolidado nos últimos anos como unha das mellores alternativas para a implementación profesional de BD de envergadura media.
2. Desenvolvemento da aplicación web de xestión de datos: Deseñouse unha aplicación web a medida para o proxecto e implementouse con tecnoloxía J2EE. A ferramenta permite que os membros do equipo de traballo xestionen, de forma colaborativa, toda a información almacenada na BD. A aplicación dá soporte á alimentación manual e mais aos procesos de revisión; conta, tamén, cun buscador avanzado que permite construír filtros para o acceso á información almacenada combinando varios criterios.
3. Desenvolvemento das ferramentas de importación de datos: Desenvolveuse un conxunto de ferramentas que permitiron automatizar, completa ou parcialmente (dependendo do caso), o proceso de alimentación da BD a partir de fontes *MARC* (tanto *MARC21 Tagged Display*, como *MARC21 XML*). Foi preciso localizar e analizar os campos e subcampos de interese nos arquivos *MARC* e almacenalos nos campos correspondentes da BD propia.
4. Desenvolvemento da aplicación web de consulta e explotación de datos: Na actualidade, estamos a traballar nesta última tarefa de desenvolvemento da aplicación web de consulta e explotación de datos, definindo as funcionalidades que proporcionará ao público xeral e deseñando unha interface de usuario que resulte atractiva, áxil e sinxela de manexar. Esta permitirá acceder a toda a información froito do traballo de catalogación e análise realizado.

A seguir, descríbese o resultado do traballo realizado nas tres primeiras tarefas, que son as que están completadas a día de hoxe.

4.1. Construción da base de datos

Na figura 1 preséntase o Diagrama Entidade-Relación que describe a información que se recolle na BD.

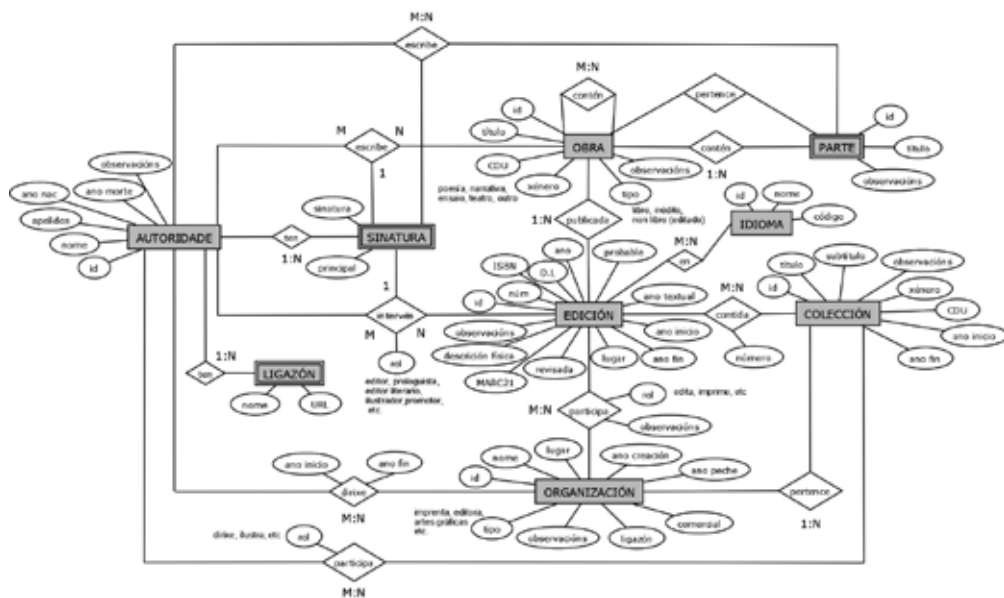


Fig. 1. Diagrama Entidade-Relación.

O centro do modelo é a entidade EDICIÓN que reflicte cada unha das “tiraxes” (edicións, reproducións, reimpresións, edicións críticas...) que se fixeron dunha obra/libro, recompilando a información específica de cada unha e deixando os datos comúns e que non varían en todas elas, co fin de os recompilar unha única vez na entidade OBRA (libro). Os atributos da entidade, que se representan cun nome dentro dunha elipse, corresponden ás características propias das edicións que se recollen na BD. Á parte, recóllese a ORGANIZACIÓN (editora, imprenta, concello, universidade...) que editou e/ou imprimiu dita edición, as/os editores, prologuistas, ilustradores, promotores, etc., isto é, as AUTORIDADES que interviñeron e a COLECCIÓN ou serie en que está esa edición, de ser o caso. Esta información descríbese a través das relacións (marcadas como rombos no diagrama) que existen entre a entidade EDICIÓN e o resto de entidades do modelo.

Do resto do modelo, cómpre destacar as estruturas que permiten almacenar información de natureza máis complexa. Referímonos, principalmente, a tres:

- a) A que permite establecer unha relación entre unha obra con calquera outra das almacenadas na BD, necesaria para contemplar, por exemplo, o caso das antoloxías.
- b) A que fai posíbel distinguir partes específicas do contido dunha obra. Esta permite, por exemplo, recuperar as diferentes edicións en que figura un poema, tanto a orixinal onde se publicou por vez primeira, como as posteriores edicións ou aquelas recompilacións en que dito poema apareza.

A que describe o feito de que algúns autores/as asinan as súas obras co seu nome real e/ou con outros pseudónimos. Permite recuperar toda a produción dun autor/a, xa se busque polo nome real ou por calquera dos seus pseudónimos.

4.2. Aplicación web de xestión de datos

A aplicación web de xestión de datos está publicada en <<http://ediciongaliciafranquista.udc.es/admin/>> mais o acceso está permitido exclusivamente ás persoas do equipo do proxecto.

Unha das interfaces deseñadas para esta aplicación é a que permite presentar un conxunto de edicións nunha listaxe tabulada (ver Fig. 2). Esta táboa deseñouse de xeito que en cada fila se amosan os datos máis representativos dunha edición: os necesarios para a permitir identificar, para acceder aos datos completos da súa catalogación (premendo no campo "Número") e mais os que permiten navegar e acceder aos datos específicos da información relacionada (a correspondente á obra, ás organizacións que editan e/ou imprimen e ás autoridades que interveñen).

Esta interface úsase na aplicación para presentar o conxunto completo de edicións da BD, ou ben un subconxunto que responde a unhas características concretas expresadas a través dos campos do buscador avanzado que ofrece a ferramenta, ou as establecidas a partir de clicks de navegación que permiten chegar ás edicións en que interveñen unha autoridade determinada, as edicións dunha colección específica, ou ás editadas/impresas por unha organización concreta.

No exemplo concreto da figura 2, pódese ver a colección dos 3.577 rexistros que existen na actualidade correspondentes a edicións de libros.

Caso de premer no "Número" da edición, preséntase a interface que permite xestionar os seus datos de catalogación (ver Fig. 3). Como se pode comprobar, para cada atributo e para cada relación da entidade EDICIÓN co resto de entida-

Fig. 2. Listaxe tabulada de edicións.

Data	Número	Obra	Autoría	Editora	Lugar	Descripción
✓ 1939	2 ^a ed. dec.	Arriba Española	Pérez de Cabo, J.	La Región	Ourense	126 p., 28 cm
✓ 1939	1 ^a ed.	Basada de a muller asturi e otras cosas	Deba, Enciso de	s.n.	A Coruña	[86] p., ...
✓ 19407	1 ^a ed.	Bando	Mosquera, F.	Monst. Imprenta: editora	A Coruña	
✓ 1939	1 ^a ed.	Colocación obrera		s.n.	Lugo	124 p., 15 cm
✓ 1939	1 ^a ed.	Con Jasevito, modo hispico de seguir la misa	Yáñez Núñez, J. Demian	Imprenta e Papelería La Industrial	Ourense	80 p., 18 cm
✓ 1939	1 ^a ed.	Del fente de Asturias al de Madrid pasando por el quibloco (por obra de un comerciante)	López Sánchez, Antonio	La Voz de la Verdad	Lugo	147 p., 21 cm
✓ 1939	1 ^a ed.	Discursos pronunciados por D. Ramón Bermúdez de Castro y Pío Alenda Comandante Militar de Asturias en las fiestas de oratoria en dicho punto para solemnizar la consigna de Madrid por el ejército español y el fracaso de la guerra de independencia hispana	Bermúdez de Castro y Pío, Ramón	s.n.	A Coruña	65 [1] p., ...
✓ 1939	1 ^a ed.	En el centenario de la aparición de la imprenta moderna del espíritu en la idea de la imprenta: discurso leído en la Academia de la Lengua Española el día 20 de octubre de 1939 a 1940	Bail del Castillo y Calvo de Osón, Carlos	Universidad de Santiago de Compostela	Santiago de Compostela	68 p., 27 cm
✓ 1939	1 ^a ed.	Estadístico de amor, de amistad	Flórez	Biblioteca Hogaña	Lugo	79 p., 18 cm
✓ 1939	1 ^a ed.	Isaapa para el libro de Cristóbal Colón	Ray Sánchez, Luciano	s.n.	A Coruña	417 p., 22 cm
✓ 1939	1 ^a ed.	Tranco ante el Adelantado	Calvo Mauro, José	Editora P.P.K.O.	Vigo	28 p., 28 cm
✓ 1939	1 ^a ed.	Una gran provincia de España según los libros producidos	López Otero, Enrique	s.n.	A Coruña	260 p., 19 cm
✓ 1939	1 ^a ed.	Dramático del año: curso sintético	Maggio, Luis	Biblioteca Hogaña	Lugo	72 p., 20 cm
✓ 1939	1 ^a ed.	Informe del inspector de sanidad de la provincia de Lugo para el año de 1938	s.n.	s.n.	Lugo	82 p., 24 cm

Fig. 3. Xestión de edicións.

des do modelo, existe na interface un cadro de texto para editar a información correspondente.

Aínda que está previsto que sexa a través da aplicación pública que na actualidade se está a desenvolver desde onde se poda realmente explotar a información recollida na BD, xa a ferramenta de administración conta cun buscador avanzado (ver Fig. 4) que permite extraer datos de interese.

Así, permite realizar buscas combinando campos como o título, o ano ou o tipo de publicación, as autoridades e o papel que xogaron na edición (autor/a, ilustrador/a, editor/a, prologuista...), a organización que editou e/ou imprimiu e o idioma en que está escrita. Deste xeito, podemos coñecer xa, por exemplo, as "edicións de libros nos que un Isaac Díaz Pardo fixo de ilustrador", ou as "edicións publicadas en galego pola editorial Celta entre os anos 1968 e 1975".

A resposta da BD a unha consulta realizada a través do buscador avanzado é a listaxe de edicións que encaixan coas características de busca, que se presenta nun formato similar ao da figura 2.

Ademais do corpus de edicións, na BD contamos con tesauros relacionados de autoridades (escritores, ilustradores, prologuistas...), organizacións que editaron ou imprimiron na época e de coleccións.

A aplicación permite visualizar e xestionar os anteditos tesauros e facilita o acceso ás edicións relacionadas. Por exemplo, na figura 5, preséntase a listaxe de autoridades que temos recollidas a día de hoxe na BD. Non só podemos acceder aos datos recompilados de cada unha delas, senón que tamén temos dispoñíbel a listaxe de edicións que produciron.

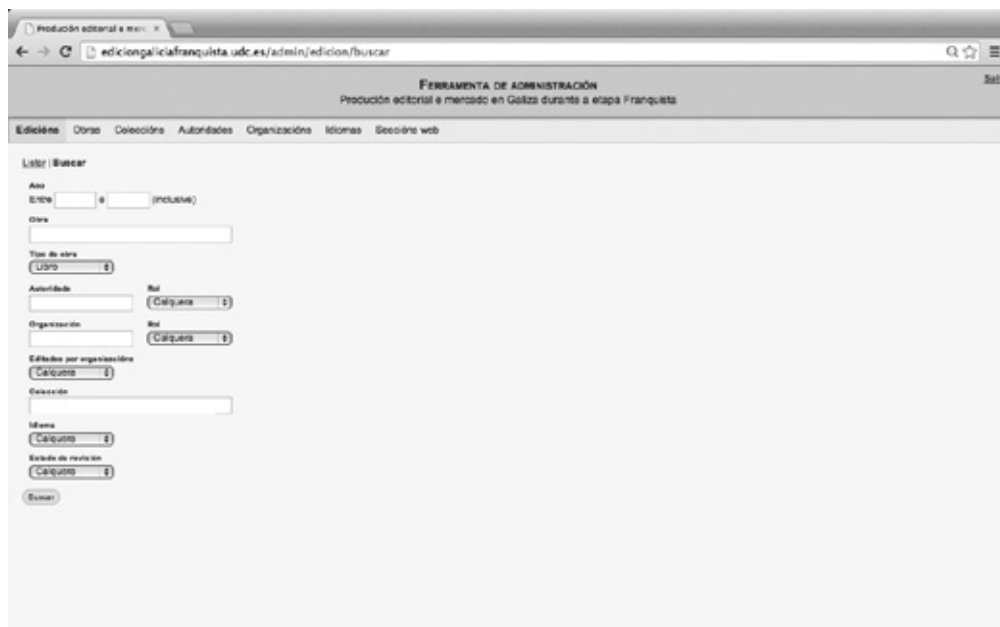


Fig. 4. Buscador avanzado.

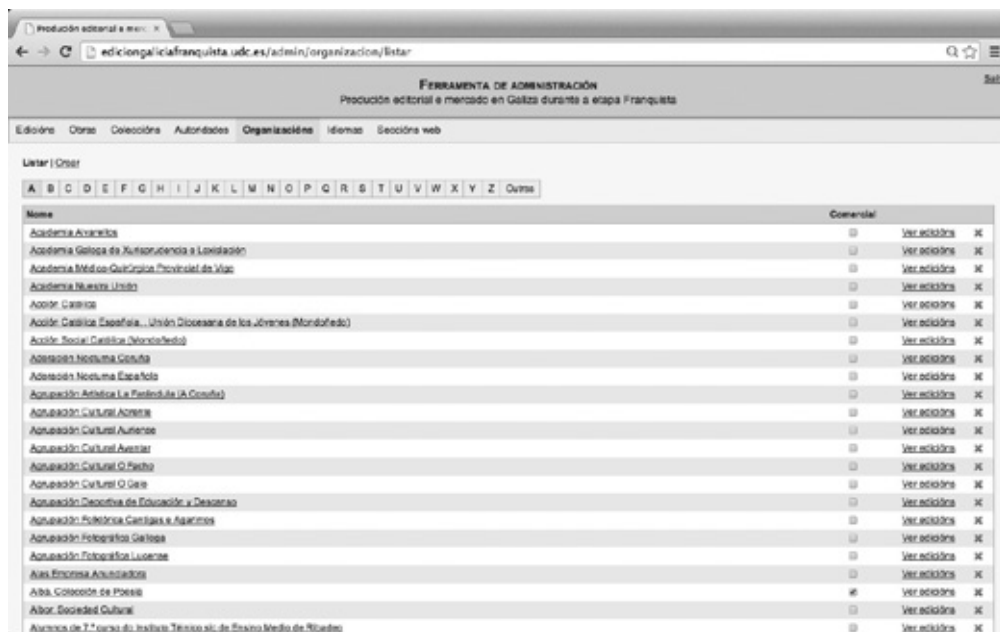


Fig. 5. Tesouro de organizacións: editoras, imprentas, etcétera.

4.3. Ferramentas de importación de datos

O obxectivo desta tarefa era importar de forma automática a información almacenada nas fontes *MARC* das que partimos, un total de 22.495 rexistros, dos cales 4.622 correspondían ao formato *MARC21 Tagged Display* (ver Fig. 6 superior) e 17.873 ao formato *MARC21 XML* (ver Fig. 6 inferior). Construíronse, polo tanto, dúas ferramentas (*parsers*), cada unha delas capaz de ler as fontes *MARC*, extraer a información dos campos relevantes para o noso contexto e almacenala nos campos correspondentes da BD.

Como xa se dixo, esperabamos que a primeira execución destas aplicacións revelase a existencia de información duplicada e, mesmo, triplicada ao pertenceren as fontes de información a diferentes bibliotecas, mais que puidese ser detectada e solucionada de forma automática, como tamén se fixo, ao afinar os *parsers* para que eliminasen signos de puntuación, espazos en branco, etc.; con todo, o maior problema veu das redundancias xeradas por culpa de non estar a información normalizada. Así, atopamos dúcias de casos en que o proceso de importación xeraba rexistros diferentes na BD para entidades ou atributos que claramente eran un só⁷. Este feito obrigou a realizar revisións exhaustivas da información das fontes *MARC*, antes e despois de as importar.

A estratexia que seguimos para evitar que o proceso de revisión fose exclusivamente manual foi extraer das fontes *MARC* listaxes de autoridades, organizacións, roles, etc. para arquivos Excel, de xeito que se puidesen detectar e resolver este tipo de redundancias. Xa con estas listaxes corrixidas, modificáronse os *parsers* con regras que permitiron almacenar as edicións asociadas ás entidades "unificadas" de forma automática.

Con todo, aínda ficou na BD moita outra información que houbo que revisar de forma completamente manual, xa que non só non era posíbel atopar patróns que permitisen construír un software que arranxase de vez todos os problemas, senón que se precisaba persoal experto que tratase e solucionase caso por caso. Porén, modificouse o esquema da BD incorporando novos atributos e desenvolvéronse funcionalidades específicas na aplicación web de xestión de datos para facilitar este traballo de revisión manual.

5. PROBLEMAS ATOPADOS E SOLUCIÓNS

Como xa foi sinalado, os maiores problemas cos que nos atopamos á hora de realizarmos o traballo de depuración e/ou revisión dos materiais importados foron os derivados da redundancia da información e da deficiente catalogación. No noso caso, esta tivo a súa orixe nos problemas preexistentes nos contidos

⁷ Isto ocorreu, por exemplo, con ORGANIZACIÓNS como "Acción Católica", para a que apareceron até 7 rexistros: "Acción Católica", "Accion Catolica", "Accion Católica", "Acción Catolica", "Acción Católica", "Acción católica; 1943]", "Acción Católica]". Tamén o detectamos con AUTORES, como "Enrique Fernández Villamil", almacenado como 5 autoridades diferentes: "Fernández Villamil y Alegre, Enrique", "Fernández Villamil, Enrique", "Fernández-Villamil Alegre, Enrique", "Fernández-Villamil y Alegre, Enrique" y "Fernndez-Villamil Alegre, Enrique". E atopamos o mesmo problema nas abreviaturas que describen os ROLES: "editor/a", "prologuista", etc., xa que variaban dun rexistro a outro con formas como "ed.", "edit." ou "edic" (para editor/a) e "pr.", "prol." ou "prolog." (para prologuista).

```

LEADER 00000nam 2200000 i 4500
001 2900515660
007 t
008 990415q1942 esp 00011 spa d
040 BUSC|bgao
090 860-1*19*
100 10 Orbe, Enrique de
245 00 Balada de la mujer estéril ; y otras poesías /|cde Enrique de Orbe ; comentadas por M. Yzquierdo y Vivas
260 1 [Galicia? :|bs.n.,|c1942?]|e(La Coruña :|fTip. El Ideal Gallego)
300 [36] p. ;|c24 cm
500 Información tomada da cuberta
700 21 Yzquierdo y Vivas, M.

```

```

<record type="Bibliographic">
  <leader>00000nam 2200000 i 4500</leader>
  <datafield tag="100" ind1="1" ind2=" " >
    <subfield code="a">Cunqueiro, Álvaro</subfield>
    <subfield code="d"> (1911-1901)</subfield>
  </datafield>
  <datafield tag="245" ind1="0" ind2="0">
    <subfield code="a">Xente de aquí e de acolá</subfield>
    <subfield code="c"> / Álvaro Cunqueiro</subfield>
  </datafield>
  <datafield tag="260" ind1=" " ind2=" " >
    <subfield code="a">Vigo</subfield>
    <subfield code="b"> : Galicia</subfield>
    <subfield code="c">, 1971</subfield>
  </datafield>
  <datafield tag="300" ind1=" " ind2=" " >
    <subfield code="a">176 p.</subfield>
    <subfield code="c"> ; 21 cm</subfield>
  </datafield>
  <datafield tag="655" ind1=" " ind2="7">
    <subfield code="a">Novela da vida cotiá</subfield>
  </datafield>
  ...
</record>

```

Fig. 6. Exemplos de Fontes MARC.

das fichas MARC 21 recibidas desde as bibliotecas que, dun modo esquemático, se poden resumir en:

- a) Falta de normalización existente nos rexistros bibliográficos dos catálogos (diferentes criterios nas denominacións, uso de abreviaturas distintas...).
- b) Erros ortográficos.
 - Duplicidades galego/español.
 - Información colocada incorrectamente nos campos MARC.
 - Ausencia de datos nas fichas MARC.

Para solucionar isto, foi preciso facer unha análise exhaustiva dos valores contidos nos distintos campos dos rexistros bibliográficos e, despois, segundo o caso, optar por diferentes tipos de procesos de depuración dos rexistros. Así, un número relativamente grande de redundancias puido ser identificado, como se indicou, de modo automático ou semiautomático mediante a comparación literal dos valores dalgúns campos.

Porén, atopamos moitos casos nos que a comparación literal de dous valores non resultou suficiente para discernir se se trataba da mesma información ou non. Por exemplo, se un rexistro contén o valor "Imprenta Moret" e outro contén "Imp. Moret", para a ferramenta informática trataríase de dúas organizacións totalmente alleas unha da outra. Dado que o período estudado é o suficiente-

mente extenso para ter que recoller mudanzas técnicas ou nominais na dirección das imprentas ou editoriais, houbo que reunificar distintas denominacións baixo unha soa: así, o que para a ferramenta informática sería unha nova organización "Gráficas do Castro-Moret", nós incluímos no mesmo rexistro "Imprenta Moret". Este mesmo problema de unificación deuse co nome de institucións (Caixas de Aforros, Deputacións Provinciais, Concellos...) que ao longo do tempo modificaron o nome e, tamén, coas variacións toponímicas que a ferramenta entendía como nomes diferentes: *Tuy/Tui*, *Pontearreas/Puentearreas*...

Outro problema consistiu na importación de rexistros que non coincidían co noso ámbito de investigación, tanto temporal (edicións posteriores a 1975, que só tiñan unha data probable de edición (197?) como xeográfica: libros con lugar de edición probable (Galiza?) ou descoñecido (s/l) que, logo de comprobar que foran editados fóra do noso marco espacial, foron borrados manualmente. Tamén se verificou e reasignou a organización editorial de boa parte dos numerosísimos rexistros que aparecían baixo a etiqueta "s.n." (sen mención editorial).

Esta exhaustiva depuración dos datos importados é a que explica a considerable redución de rexistros que se produciu desde o momento da importación dos datos ao día de hoxe (de 22.495 a 4.849 rexistros⁸). Aliás, cómpre subliñar que esta revisión só se puido facer consultando e cotexando un a un os rexistros recollidos na nosa ferramenta cos de diversas bases de datos en rede: catálogo das Bibliotecas das Universidades de Santiago de Compostela, Coruña e Vigo, catálogo da Biblioteca Nacional, do Depósito Legal, da Real Academia Galega, Biblioteca Dixital de Galicia, etc. ou, mesmo, co cotexo físico dos libros. Esta fase está a ser de enorme interese pois nos permite completar datos que non aparecen na ficha *MARC 21* da que partimos, como a información referida a coleccións ou series (que en moi poucos casos aparecían recollidas nas fichas catalográficas da época), así como outros datos complementares de grande interese (autoría de limiares, prólogos, capas, ilustracións...).

Ao longo deste proceso de revisión constatamos tamén a necesidade de crear as fichas de certas obras que, constándonos a súa existencia, non aparecían na importación dos datos das bibliotecas (isto é, non existía ningún exemplar das mesmas nas bibliotecas públicas galegas). Estas incorporacións manuais ao programa informático non levan incluída, como é lóxico, a ficha *MARC 21*.

Aínda que unha BD destas características sobre a edición en Galiza de 1939 a 1975 é un traballo que dificilmente estará de todo rematado (pensemos nas innumerábeis edicións de autor que aparecen en calquera biblioteca particular que revisemos), sabemos que a nosa recompilación de datos ofrece unha nova perspectiva sobre o panorama editorial da época até o de agora nunca enfrontado no seu conxunto.

CONCLUSIÓNS

A partir da nosa experiencia neste proxecto de investigación, podemos extraer dous tipos de conclusións, unhas relacionadas cos aspectos metodolóxicos e outras cos contidos do noso traballo.

⁸ Aínda que, como xa se explicou, este número de rexistros inclúe libros, folletos, separatas e revistas, o noso traballo centrase nos 3.300 títulos de libros que temos actualmente na base de datos.

Canto á metodoloxía, primeiro, parécenos imprescindible insistir en que a colaboración entre equipos interdisciplinares de Humanidades e de Informática é básica para poder afrontar estudos como o desenvolvido neste proxecto. En segundo lugar, aínda que a día de hoxe a dixitalización chegou a todas partes (sen ir máis lonxe, aos catálogos das bibliotecas) e que, por tanto, se pode facer un tratamento informático deses datos, a realidade é que estamos nun tempo de transición onde, por máis que teñamos ferramentas informáticas e catálogos dixitalizados, as metodoloxías de traballo son aínda, por así dicilo, "pre-informáticas". Faltan, pois, normalización, estandarización, protocolos de traballo, procedementos... liñas dun camiño que haberá que ir facendo.

Finalmente, canto aos resultados do noso proxecto, podemos afirmar que o nosa traballo de recompilación fornece unha visión novedosa do panorama editorial na Galiza da época franquista, cuxa análise en conxunto non fora feita até agora. Unha vez rematada, esta será unha Base de Datos de referencia para o estudo deste período.

BIBLIOGRAFÍA

- Alonso Montero, Xesús y Ramos de Castro, Epifanio, *Cen anos de literatura galega. Catálogo da exposición bibliográfica 9/XII/1963-29/XIII/1963*, Lugo, Círculo das Artes, 1963.
- Alvarez García, Isabel, *O Libro Galego, onte e hoxe 1993: catálogo da Exposición bibliográfica, Santiago, xaneiro 1995*, Santiago, Federación de Libreiros de Galicia, 1995.
- Bassave Roibal, Rosa M^a et al., *O Libro Galego, onte e hoxe 1988: catálogo da Exposición bibliográfica, Santiago, maio 1988*, Santiago, Federación de Libreiros de Galicia, 1989.
- Cuadernos de Estudios Gallegos*, apéndice Bibliografía de Galicia, tomos I-XI (1945-1956), XIII (1958), XIX-XXIII (1964-1968).
- Galaxia Almanaque: 1950-1975*, Vigo, Galaxia, 1975.
- Instituto Nacional del Libro Español (ed.), *Catálogo General de la Librería Española 1931-1950*, Madrid, Instituto Nacional del Libro Español, 1957-1961.
- Instituto Nacional del Libro Español (ed.), *Índice de la producción editorial española 1968-1972*, Madrid, Instituto Nacional del Libro Español, 1972.
- Instituto Nacional del Libro Español (ed.), *Catálogo de libros en gallego*, Madrid, Instituto Nacional del Libro Español, 1974.
- Instituto Nacional del Libro Español (ed.), *Catálogo de libros en gallego*, Madrid, Instituto Nacional del Libro Español, 1978.
- Palau y Dulcet, Antonio, *Manual del librero hispanoamericano: bibliografía general española e hispano-americana desde la invención de la imprenta hasta nuestros tiempos con el valor comercial de los impresos descritos*, Barcelona, Antonio Palau, 1948.
- Servicio Nacional de Información Bibliográfica (ed.), *Bibliografía Española*, Madrid, Dirección General de Archivos y Bibliotecas, 1958-1968.
- Vilariño Pintos, Daría y Pardo Gómez, M^a Virtudes, *O Libro galego onte e hoxe. Catálogo da exposición bibliográfica, Santiago, maio-xuño 1979*, Santiago, Secretariado de Publicacións da USC, 1981.



RESUMO

Explicación da metodoloxía e das ferramentas informáticas desenvolvidas para acometer un proxecto de investigación propio das Humanidades dixitais. En concreto, este proxecto ten como obxectivo fundamental estudar a produción editorial na Galiza durante

a etapa franquista (1939-1975), analizar se existiu unha industria cultural editora e en que medida esta funcionou como canle difusora, tanto dunha liña de resistencia política contra o franquismo, como noutra de aposta pola conservación e posta en valor da identidade cultural da Galiza.

Tras unha breve introdución ao proxecto, preséntase a metodoloxía seguida no proceso de recuperación e catalogación dos fondos, explícase como se desenvolveu o proceso de importación masiva dos catálogos das bibliotecas públicas máis relevantes da Galiza, así como as súas características técnicas e, a seguir, descríbense as ferramentas informáticas deseñadas. Finalmente, sinálanse os problemas metodolóxicos atopados e as solucións que se foron adoptando pois considérase que poden ser de interese para outros proxectos do mesmo tipo. Coidamos que o noso modo de enfrontalos e a experiencia obtida á hora de resolvelos constitúen, de seu, achegas relevantes deste proxecto ao mundo das Humanidades dixitais.

Palabras clave: Producción editorial, base de datos, clasificación bibliográfica, franquismo, Galiza.

ABSTRACT

Explanation of the methodology and the computational tools developed to undertake a research project within the domain of the Digital Humanities. In particular, the main goal of this project is to study the publishing industry in Galicia during the Franquist period (1939-1975), to analyze whether there was a culturally-concerned publishing industry and to determine the extent to which this acted as a means of public diffusion, both of political resistance against Franquism as well as the commitment to the preservation and enhancement of the Galician cultural identity.

After a brief introduction to the project, we present the methodology followed throughout the process of retrieval and cataloging of the lists of titles, we explain now the process of the massive importation of the catalogs of the most important Galician public libraries was carried out, as well as their technical features and, subsequently, we describe the computational tools designed. Finally, we point out the methodological problems found and the solutions adopted given their possible usefulness for similar projects. We consider that the way they were dealt with and the experience gained solving them constitute, in themselves, important contributions of this project to the area of digital humanities.

Keywords: Publisher production, Data bases, Bibliographic compilation, Franquism, Galiza.