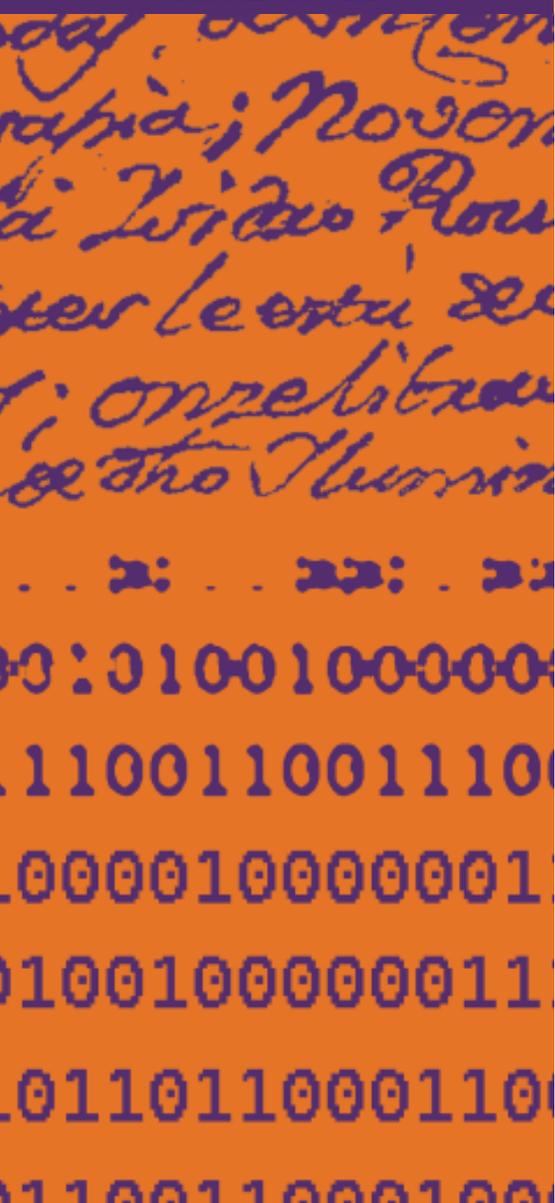


DEL TRAÇ AL BYTE

Eines digitals per a l'estudi històric
del canvi social



Explorando los cambios sociales silenciosos: la explotación digital de una mina de datos históricos

Rosa Congost, Jordi Regincós, Rosa Ros y Enric Saguer

Universitat de Girona

Un proyecto en colaboración

Pattern Recognition and
Human Language Technology
Research Center



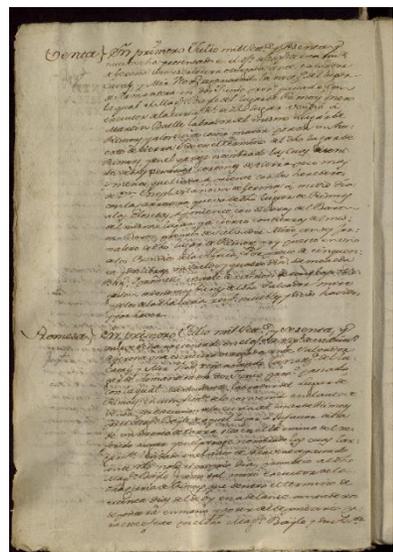
OBJETIVO

aplicar las herramientas y procesos de reconocimiento de texto manuscrito a las imágenes digitalizadas de una gran serie documental

los libros de las Oficinas de Hipotecas (1768–1861) de la región de Girona

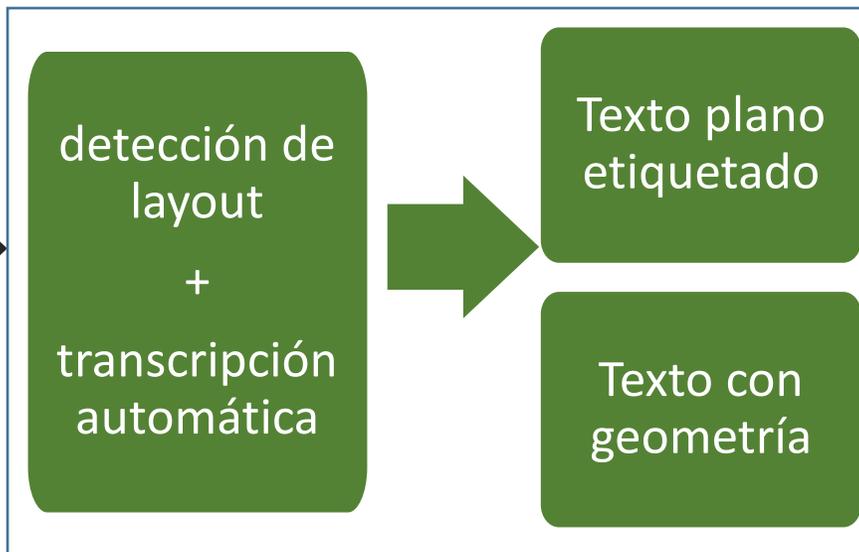
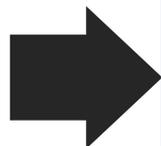
para analizar los procesos de cambio social desde mediados del siglo XVIII hasta mediados del siglo XIX

Etapas del proceso



Fase 0
Digitalización
AHG & CRHR

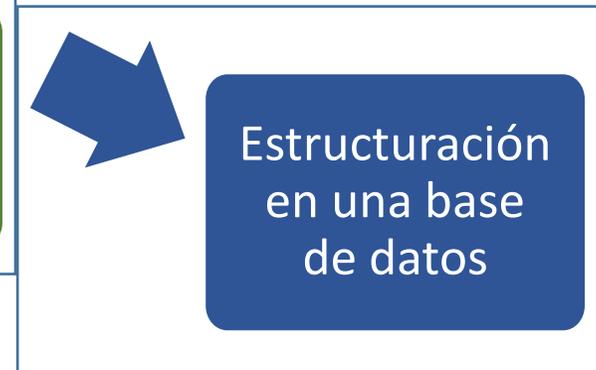
103.300 imágenes



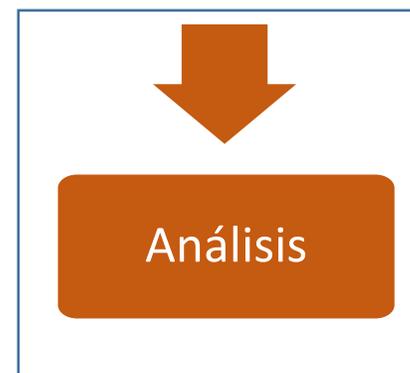
Fase 1
Transcripción
PRHLT & CRHR



Exploración
de
resultados



Fase 2
J. Regincós
EPS UdG



Fase 3
CRHR

Esquema intervención

1. Contenido y potencialidades de los libros de la Oficina de Hipotecas
2. Estado actual del proyecto
3. ¿Qué hacemos con el texto transcrito y etiquetado?
4. Conversión del texto plano en una base de datos



En proceso

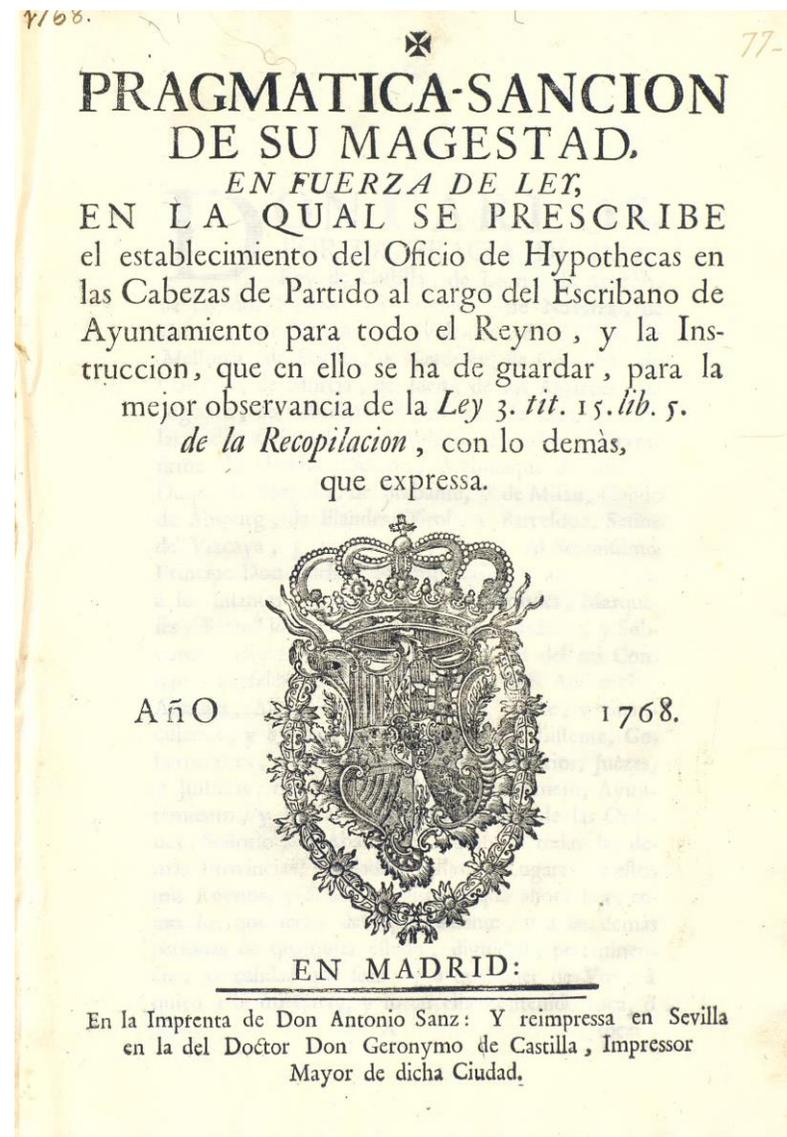
El Registro de Hipotecas

- Institución de publicidad registral
- Creación mediante Real Pragmática 5 febrero 1768
- Cambios en la denominación:

1768 – Oficio de Hipotecas

1829 – Contaduría de Hipotecas

1845 – Registro de Hipotecas



El Registro de Hipotecas

Pragmática
31 enero 1768

1768

1829

derecho de
inscripción (0,5%)

1845

reorganización de
las contadurías

orientación hacia las
transmisiones de
propiedad

1862

Sustitución por
el Registro de
la Propiedad

El contenido del Registro de Hipotecas de Girona

- contratos con cargas e hipotecas sobre bienes
- criterio de aplicación no homogéneo
- Cataluña = inclusión de un amplio número de actas notariales (garantías hipotecarias generales)

↑↑↑ colaboración notarios

para que en ellos precisamente se tome la Razon de todos los Instrumentos de imposiciones, ventas, y redenciones de Censos, ò Tributos, ventas de bienes raices, ò considerados por tales, que constare estar gravados con alguna carga, fianzas, en que se hypothecaren especialmente tales bienes, Escrituras de Mayorazgos, ò Obra Pia, y generalmente todos los que tengan especial, y expresa Hypoteca, ò gravamen, con expresion de ellos, ò su liberacion, y redencion. Que

El contenido del Registro de Hipotecas de Girona

Tipologías documentales básicas en los Oficios de Hipotecas de la región de Girona, 1768-1770

	n	%
operaciones de crédito	3.262	29%
compraventas	2.472	22%
capítulos matrimoniales, testamentos	1.452	13%
establecimientos	1.145	10%
arrendamientos	901	8%
ápocas, cartas de pago	416	4%
donaciones, cesiones	333	3%
concordias, sentencias	215	2%
caución, indemnidad, garantía	136	1%
inventarios, subastas	135	1%
modificaciones de precio, reducciones	66	1%
definiciones	54	0,5%
cabrevaciones, confesiones	40	0,4%
poderes	8	0,1%
otros	468	4%
no identificada	133	1%
	11.236	100%

nota: Estimación realizada a partir de los 4 primeros volúmenes

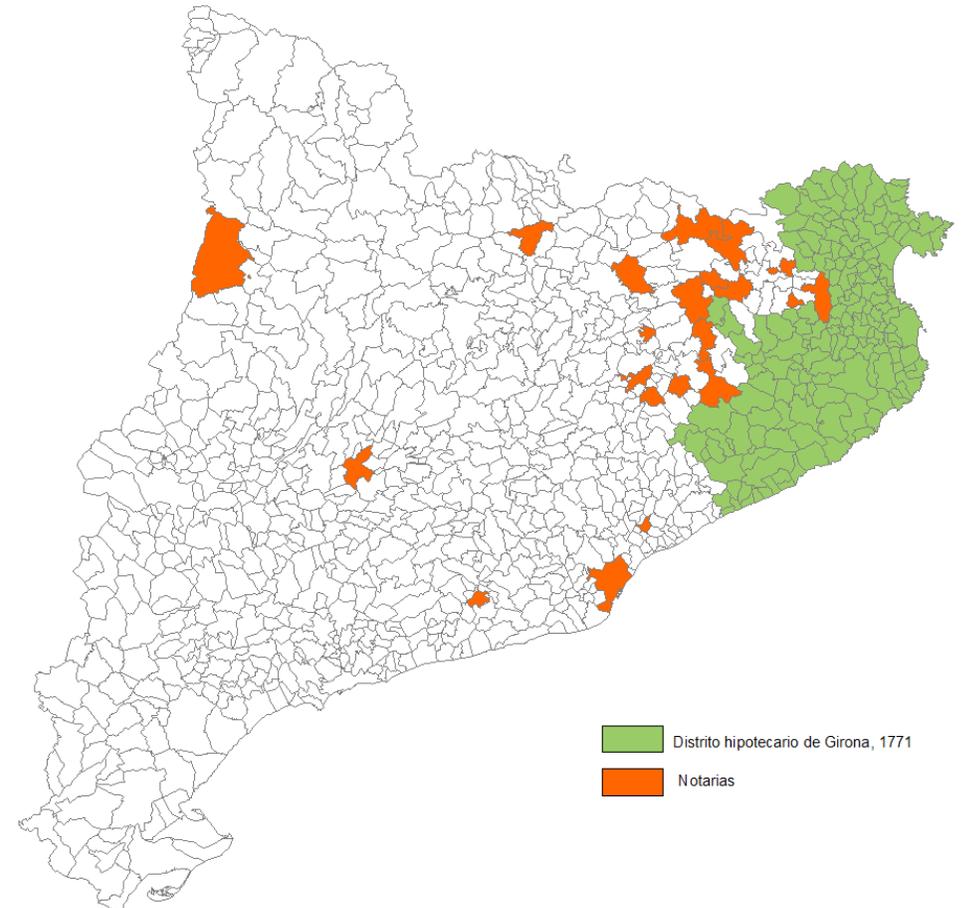
Contenido del Registro de Hipotecas de Girona

- amplia extensión territorial
 - 1768-1773 = 3.884 km²
- Obligación de presentar la escritura en la oficina donde se hallan los bienes afectados



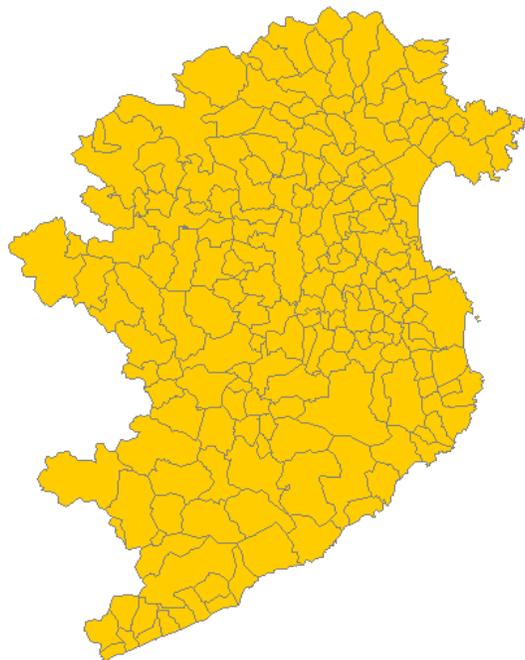
Inclusión de todas las escrituras referidas a fincas del territorio de cada oficina

Notarías de procedencia de las escrituras registradas en el Oficio de Hipotecas de Girona, 1771



Modificaciones territoriales

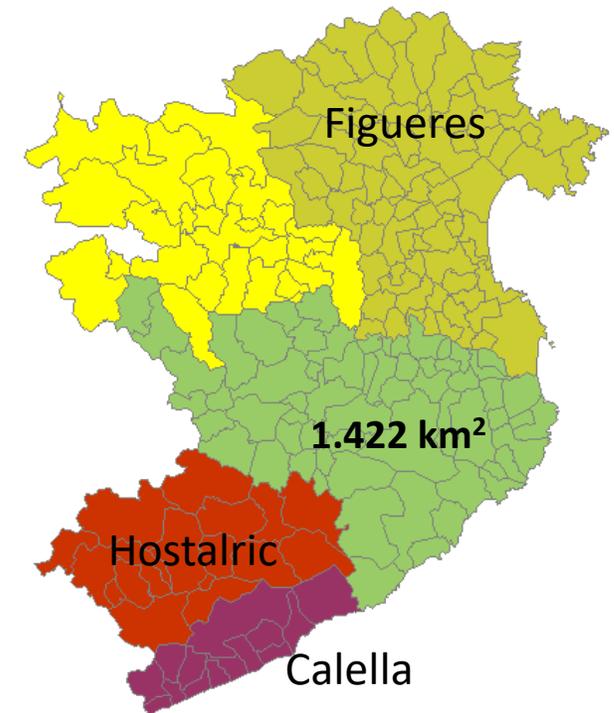
- Pragmática 1768 → 1 oficina / cabeza de partido
 - Inexistencia de Partidos en Cataluña → Corregimientos (7)
 - Traslación Pragmática al Corregimiento de Girona: 2 oficinas (Girona + Besalú → Alcaldía Mayor)
 - 1774 → Segregación de los Oficios de Figueres i de Hostalric (1780 Calella)



Corregimiento

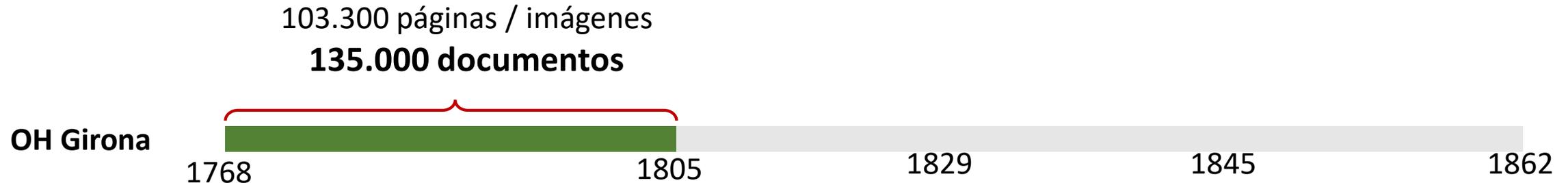


Distritos hipotecarios 1768



Distritos hipotecarios 1780

Objetivo operativo a medio plazo: OH Girona, 1768-1805



Por qué hasta 1805?

- relativa homogeneidad geográfica (excepto 1768-1773)
>1806 - recuperación libros OH de Figueres
- mantenimiento de una única tipología libros (registros generales)
- estimación vaciado manual: 25 años x 6 estudiantes x 15 horas/semana
- relativa continuidad en el tipo de letra manuscrita (≈)

Cambios en el tipo de letra

entrenamiento sistema de transcripción → letra 1769

1768

Decreto Juizial } En once Mayo mil sett. setenta y ocho, se ha pintado
en el officio de Hipotecas de esta Ciudad de Leonora

1786

Cap. Mat. } En quatro Julio mil sett. ochenta y seis, se ha
pintado en el off. de Hipot. de esta Ciudad de Leonora, una casa de
la calle de la Llamada

1790

Debitorio y } En tres Julio mil sett. y noventa, se presentó en el off. de
novo caso Venta } de esta Ciudad de Leonora una escritura otorgada ante Antonio Arroyo

1792

Realcédula } En tres Julio mil sett. noventa y dos, se presentó en el off.
de esta Ciudad de Leonora una escritura otorgada ante Antonio Arroyo

1805

Ariendo } En cinco Enero mil ochocientos CINCO, se presentó en el off.
de esta Ciudad de Leonora una escritura otorgada ante Juan de
la villa de la Sibola en veinte y ocho de

¿respuesta del sistema de transcripción ante los cambios de letra?

Estado actual

mayo 2016

Fase de contacto inicial y entrenamiento del sistema RTM

volumen 1769 (1179 folios), en lotes de 50 a 400 imágenes

enero 2019

inicio del proceso de transcripción masiva

revisión del layout  herramienta = TRANSKRIBUS

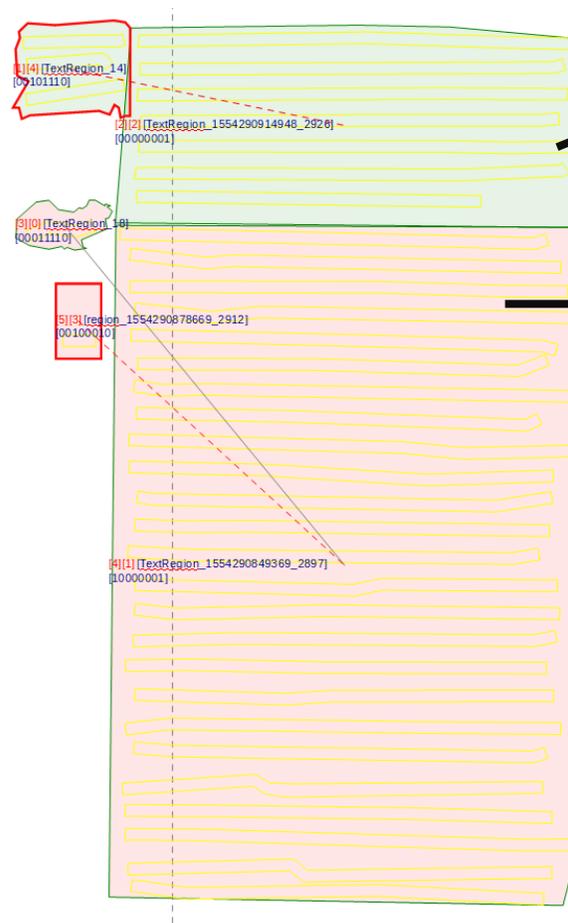
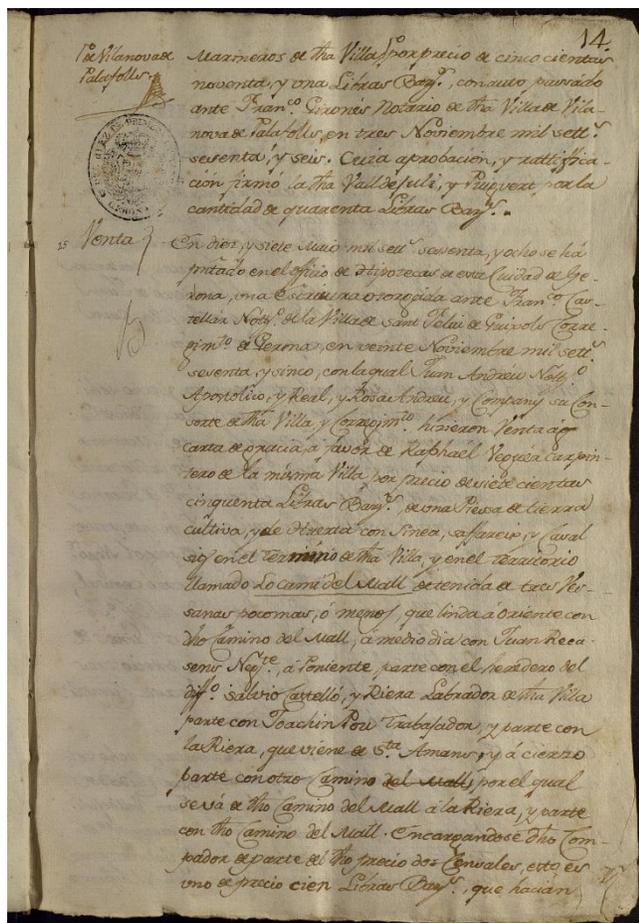
revisión muestra de 50 imágenes  herramienta = CATTI

Estado actual

		Fase 0	Fase 1		Fase 2	
libro	año	digitalizado	revisión layout	transcripción	inclusión en BD	num imágenes
1	1768					2091
2	1769					2373
3	1769					1785
4	1770					2307
5	1770					1934
6	1771					2459
7	1771					2652
8	1772					2170
9	1772					1774
10	1772					1676

¿Qué obtenemos?

1. Archivo de texto plano para cada imagen, con distinción de las regiones de texto



```
<TextRegion>
  $pag:54
</TextRegion>
<TextRegion>
  <ofi>Marineros</ofi> de dicha Villa; Por precio de cinco cientos
  noventa y una libras barcelonesa, con auto pasado
  ante <ant>Francisco Girons</ant> <ofi>Notario</ofi> de dicha villa de <top>Panova</top>
  <top>de Palafoalls,</top> en tres Noviembre mil Setecientos
  sesenta, y seis; cuio aprobacion, ratificacion
  firmó la dicho <ant>Valldejuli y Puig</ant> de la
  cantidad de quarenta libras barcelonasas
</TextRegion>
<TextRegion>
  de <top>Vilanova</top>
  <top>Palafoalls</top>
  Figueras
</TextRegion>
<tip> Venta </tip>
</TextRegion>
<TextRegion>
  En diez y Siete Maio mil Setecientos sesenta, y ocho su
  en el Oficio de Hipotecas de esta ciudad de <top>Gerona</top>
  una escritura otorgada ante <ant>Francisco Ca-rer</ant>
  Notario de la villa de <top>San Feliu de Guixols</top> Ferre-mino
  de <top>Gerona</top> en veinte Noviembre mil Seteciento
  sesenta y cinco, con la qual <ant>Juan Andreu</ant> Notario
  <ant>Antonio y Real,</ant> y <ant>Rosa Andreu y Fontana</ant> su conte
  de dicha villa y termino hizieron censal
  carta de gracia á favor de <ant>Raphael Saguer</ant> <ofi>Carpintero</ofi>
  de la misma Villa, por precio de siete cientos
  cinquenta libras barcelonasas, de una piessa de tierra
  cultivada, y huerto con <ant>Serra</ant> salario, y assol
  sito en el termino de dicha Villa, y en el territorio
  Camp del Moli detenida de tres vessanas
  poco mas, ó, menos, que linda á oriente con
  dicho Camino de Mal, á medio dia con <ant>Juan Re-ca-sen</ant>
  Negociante, á poniente parte con el heredero de
  difuncto <ant>Salvio Castelló y Riera</ant> <ofi>Labrador</ofi> de dicha villa
  parte con <ant>Joachin Pou</ant> <ofi>trabajador,</ofi> y parte con
  la Riera, que se de <top>Sant Amans,</top> y á cierto
  parte con otro camino del Manso por qual
  se de dicho Camino del Mal la Riera, y parte
  con dicho camino de <top>Mar</top> encargado de dicho comvador
  de parte de dicho precio de censales, esto es
  de precio cien libras barcelonasas, que hizo
</TextRegion>
<TextRegion>
  Bisbal
</TextRegion>
```

transcripción enriquecida con etiquetas xml + abreviaturas expandidas + particiones palabras

Región de texto

topónimo

antropónimo

oficio

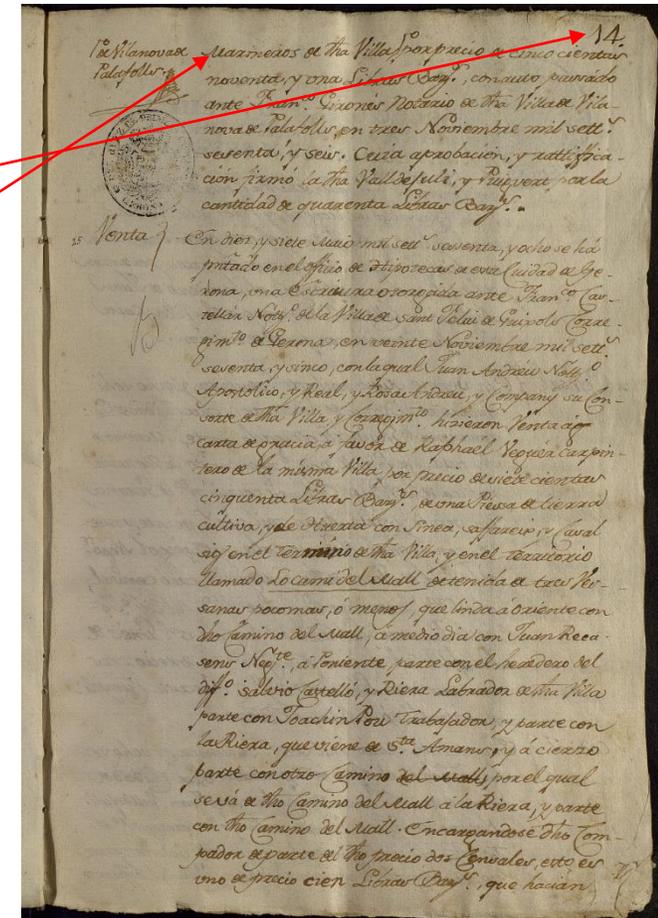
```
<TextRegion>
  <tip> Venta </tip>
</TextRegion>
<TextRegion>
  En diez y Siete Maio mil Setecientos sesenta, y ocho su
  en el Oficio de Hipotecas de esta ciudad de <top>Gerona</top>
  una escritura otorgada ante <ant>Francisco Ca-rer</ant>
  Notario de la villa de <top>San Feliu de Guixols</top> Ferre-mino
  de <top>Gerona</top> en veinte Noviembre mil Seteciento
  sesenta y cinco, con la qual <ant>Juan Andreu</ant> Notario
  <ant>Antonio y Real,</ant> y <ant>Rosa Andreu y Fontana</ant> su conte
  de dicha villa y termino hizieron censal
  carta de gracia á favor de <ant>Raphael Saguer</ant> <ofi>Carpintero</ofi>
  de la misma Villa, por precio de siete cientas
  cinquenta libras barcelonesas, de una piessa de tierra
  cultivada, y huerto con <ant>Serra</ant> salario, y assol
  sito en el termino de dicha Villa, y en el territorio
  Camp del Moli detenida de tres vessanas
  poco mas, ó, menos, que linda á oriente con
  dicho Camino de Mal, á medio dia con <ant>Juan Reca-sen</ant>
  Negociante, á poniente parte con el heredero de
  diffuncto <ant>Salvio Castelló y Riera</ant> <ofi>Labrador</ofi> de dicha villa
  parte con <ant>Joachin Pou</ant> <ofi>trabajador,</ofi> y parte con
  la Riera, que se de <top>Sant Amans,</top> y á cierzo
  parte con otro camino del Manso por qual
  se de dicho Camino del Mal la Riera, y parte
  con dicho camino de <top>Mar</top> encargado de dicho comvador
  de parte de dicho precio de censales, esto es
  de precio cien libras barcelonesas, que hizo
</TextRegion>
```

Qué obtenemos?

2. Archivo page con la geometría y las etiquetas sin desarrollar

→ generación de los documentos de la base de datos

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<PcGts xmlns="http://schema.primaresearch.org/PAGE/gts/pagecontent/2013-07-15" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="
http://schema.primaresearch.org/PAGE/gts/pagecontent/2013-07-15 http://schema.primaresearch.org/PAGE/gts/pagecontent/2013-07-15/pagecontent.xsd">
  <Metadata>
    <Creator>P2PaLa-PRHLT</Creator>
    <Created>2019-03-28T18:37:06</Created>
    <LastChange>2019-04-03T13:36:29.508+02:00</LastChange>
  </Metadata>
  <Page imageFilename="17002512000001,0036.tif" imageWidth="2790" imageHeight="4027">
    <ReadingOrder>
      <OrderedGroup id="ro_1554291391601" caption="Regions reading order">
        <RegionRefIndexed index="0" regionRef="TextRegion_1"/>
        <RegionRefIndexed index="1" regionRef="TextRegion_1554290914948_2926"/>
        <RegionRefIndexed index="2" regionRef="TextRegion_14"/>
        <RegionRefIndexed index="3" regionRef="TextRegion_18"/>
        <RegionRefIndexed index="4" regionRef="TextRegion_1554290849369_2897"/>
        <RegionRefIndexed index="5" regionRef="region_1554290878669_2912"/>
      </OrderedGroup>
    </ReadingOrder>
    <TextRegion id="TextRegion_1" custom="readingOrder {index:0;} structure {type:$pag;}">
      <Coords points="2666,102 2659,82 2619,58 2579,66 2448,70 2437,78 2430,153 2430,173 2452,188 2561,184 2611,196 2659,196 2666,169"/>
      <TextLine id="TextLine_0_1" custom="readingOrder {index:0;} structure {type:$pag;}">
        <Coords points="2631,145 2562,117 2488,141 2503,188 2560,170 2563,171 2631,200 2644,148"/>
        <Baseline points="2500,177 2561,157 2623,183 2632,146"/>
        <TextEquiv>
          <Unicode>$pag:54</Unicode>
        </TextEquiv>
      </TextLine>
    </TextRegion>
    <TextRegion id="TextRegion_1554290914948_2926" custom="readingOrder {index:1;} structure {type:$pag;}">
      <Coords points="755,157 1718,145 2107,153 2662,204 2630,1008 688,988"/>
      <TextLine id="TextLine_7_23" custom="readingOrder {index:0;} structure {type:$pag;}">
        <Coords points="2641,200 2371,194 1756,174 789,188 784,238 1753,224 2369,244 2640,250"/>
        <Baseline points="796,227 1754,212 2370,232 2641,238"/>
        <TextEquiv>
          <Unicode>$ofi:Marineros de dha$dicha Villa; Por precio de cinco cientos</Unicode>
        </TextEquiv>
      </TextLine>
      <TextLine id="TextLine_5_23" custom="readingOrder {index:1;} structure {type:$pag;}">
        <Coords points="2584,289 2539,301 2396,302 792,308 793,358 2397,352 2545,351 2597,338"/>
        <Baseline points="815,349 2397,340 2544,339 2594,326"/>
        <TextEquiv>
          <Unicode>noventa y una libras bar.$^s$.barcelonesa, con auto passado</Unicode>
        </TextEquiv>
      </TextLine>
    </TextRegion>
  </Page>
</PcGts>
```



Una transcripción con cierto margen de error

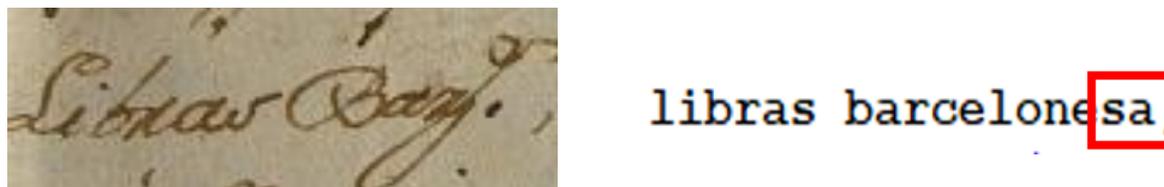
```
<TextRegion>
  $pag:54
</TextRegion>
<TextRegion>
  <ofi>Marineros</ofi> de dicha Villa; Por precio de cinco cientos
noventa y una libras barcelonesas con auto passado
ante <ant>Francisco Girons</ant> <ofi>Notario</ofi> de dicha villa de <top>Panova</top>
de Palafoills,</top> en tres Noviembre mil Setecientos
sesenta, y seis; cuio aprobacion, ratificacion
firmó la dicho <ant>Valdejuli y Puig</ant> de la
cantidad de quarenta libras barcelonesas
</TextRegion>
<TextRegion>
  de <top>Vilanova</top>
  <top>Palafoills</top>
  Figueras
</TextRegion>
<TextRegion>
  <tip> Venta </tip>
</TextRegion>
<TextRegion>
  En diez y siete Maio mil Setecientos sesenta, y ocho su
en el Oficio de Hipotecas de esta ciudad de <top>Gerona</top>
una escritura otorgada ante <ant>Francisco Ca-rrer</ant>
Notario de la villa de <top>San Feliu de Guixols</top> <ant>Ferre-mino
de <top>Gerona</top> en veinte Noviembre mil Seteciento
sesenta y cinco, con la qual <ant>Juan Andreu</ant> Notario
<ant>Antonio y Real,</ant> y <ant>Rosa Andreu y Pontana</ant> su conte
de dicha villa y termino hizieron <ant>censal
carta de gracia á favor de <ant>Raphael Saquer</ant> <ofi>Carpintero</ofi>
de la misma Villa, por precio de siete cientas
cinquenta libras barcelonesas, de una piessa de tierra
cultiva, y huerto con <ant>Serra</ant> salario, y assol
sito en el termino de dicha Villa, y en el territorio
Camp del Moli detenida de tres vessanas
poco mas, ó, menos, que linda á oriente con
dicho Camino de Ma á medio dia con <ant>Juan Reca-sen</ant>
Negociante, á poniente parte con el heredero de
diffuncto <ant>Salvio Castelló y Riera</ant> <ofi>Labrador</ofi> de dicha villa
parte con <ant>Joachin Pou</ant> <ofi>trabajador,</ofi> y parte con
la Riera, que se de <top>Sant Amans,</top> y á cierzo
parte con otro camino del Manso por qual
se de dicho Camino del Ma la Riera, y parte
con dicho camino de <top>Mar</top> encargado de dicho conador
de parte de dicho precio de censales, esto es
de precio cien libras barcelonesas, que hizo
</TextRegion>
<TextRegion>
  Bisbal
</TextRegion>
```

1. errores irrelevantes

- numeración página



- concordancia de número o de genero

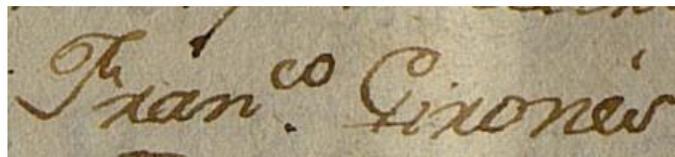


Una transcripción con cierto margen de error

```
<TextRegion>
  Spag: 54
</TextRegion>
<TextRegion>
  <ofi>Marineros</ofi> de dicha Villa; Por precio de cinco cientos
noventa y una libras barcelonesas con auto passado
ante <ant>Francisco Girons</ant> <ofi>Notario</ofi> de dicha villa de <top>Panova</top>
de Palafoflls, en tres Noviembre mil Setecientos
sesenta, y seis; cuio aprobacion, ratificacion
firmó la dicho <ant>Valledejuli y Puig</ant> de la
cantidad de quarenta libras barcelonesas
</TextRegion>
<TextRegion>
  de <top>Vilanova</top>
  <top>Palafoflls</top>
  Figueras
</TextRegion>
<TextRegion>
  <tip> Venta </tip>
</TextRegion>
<TextRegion>
  En diez y siete Maio mil Setecientos sesenta, y ocho su
en el Oficio de Hipotecas de esta ciudad de <top>Gerona</top>
una escritura otorgada ante <ant>Francisco Ca-rer</ant>
Notario de la villa de <top>San Feliu de Guixols</top> ferre-mino
de <top>Gerona</top> en veinte Noviembre mil Seteciento
sesenta y cinco, con la qual <ant>Juan Andreu</ant> Notario
<ant>Antonio y Real,</ant> y <ant>Rosa Andreu y Pontana</ant> su conte
de dicha villa y termino hizieron censal
carta de gracia á favor de <ant>Raphael Saquer</ant> <ofi>Carpintero</ofi>
de la misma Villa, por precio de siete cientas
cinquenta libras barcelonesas, de una piessa de tierra
cultiva, y huerto con <ant>Serra</ant> salario, y assol
sito en el termino de dicha Villa, y en el territorio
Camp del Moli detenida de tres vessanas
poco mas, ó, menos, que linda á oriente con
dicho Camino de Mal á medio dia con <ant>Juan Re-ca-sen</ant>
Negociante, á poniente parte con el heredero de
diffuncto <ant>Salvio Castelló y Riera</ant> <ofi>Labrador</ofi> de dicha villa
parte con <ant>Joachin Pou</ant> <ofi>trabajador,</ofi> y parte con
la Riera, que se de <top>Sant Amans,</top> y á cierto
parte con otro camino del Manso por qual
se de dicho Camino del Mal la Riera, y parte
con dicho camino de <top>Mar</top> encargado de dicho conrador
de parte de dicho precio de censales, esto es
de precio cien libras barcelonesas, que hizo
</TextRegion>
<TextRegion>
  Bisbal
</TextRegion>
```

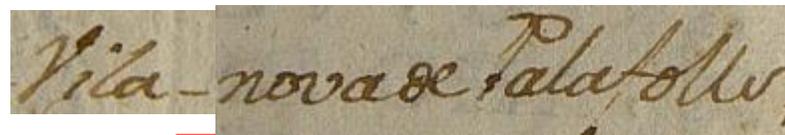
2. errores interpretables y enmendables

- nombre del notario

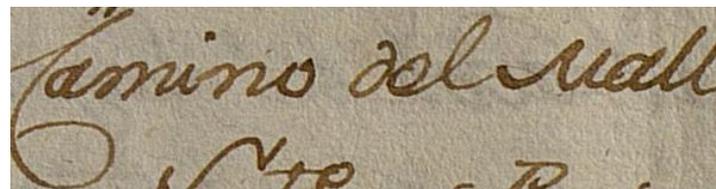


Francisco Girons

- topónimos



<top>Panova</top> <top>de Palafoflls



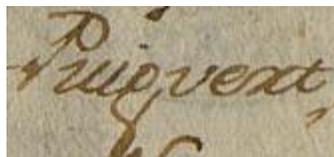
Camino del Mal
camino del Manso
camino de <top>Mar

Una transcripción con cierto margen de error

```
<TextRegion>
  Spag:54
</TextRegion>
<TextRegion>
  <ofi>Marineros</ofi> de dicha Villa; Por precio de cinco cientos
noventa y una libras barcelonesas con auto pasado
ante <ant>Francisco Girons</ant> <ofi>Notario</ofi> de dicha villa de <top>Panova</top>
<top>de Palafoflls,</top> en tres Noviembre mil Setecientos
sesenta, y seis; cuio aprobacion, ratificacion
firmó la dicho <ant>Valledejuli y Puig</ant> de la
cantidad de quarenta libras barcelonesas
</TextRegion>
<TextRegion>
  de <top>Vilanova</top>
  <top>Palafoflls</top>
  Figueras
</TextRegion>
<TextRegion>
  <tip> Venta </tip>
</TextRegion>
<TextRegion>
  En diez y Siete Maio mil Setecientos sesenta, y ocho su
en el Oficio de Hipotecas de esta ciudad de <top>Gerona</top>
una escritura otorgada ante <ant>Francisco Ca-rrer</ant>
Notario de la villa de <top>San Feliu de Guixols</top> <ant>Ferre-mino
de <top>Gerona</top> en veinte Noviembre mil Seteciento
sesenta y cinco, con la qual <ant>Juan Andreu</ant> Notario
<ant>Antonio y Real,</ant> y <ant>Rosa Andreu y Pontana</ant> su conte
de dicha villa y termino hizieron <ensal
carta de gracia á favor de <ant>Raphael Saquer</ant> <ofi>Carpintero</ofi>
de la misma Villa, por precio de siete cientas
cinquenta libras barcelonesas, de una piessa de tierra
cultiva, y huerto con <ant>Serra</ant> salario, y assol
sito en el termino de dicha Villa, y en el territorio
Camp del Moli detenida de tres vessanas
poco mas, ó, menos, que linda á oriente con
dicho Camino de Ma á medio dia con <ant>Juan Re-ca-sen</ant>
Negociante, á poniente parte con el heredero de
diffuncto <ant>Salvio Castelló y Riera</ant> <ofi>Labrador</ofi> de dicha villa
parte con <ant>Joachin Pou</ant> <ofi>trabajador,</ofi> y parte con
la Riera, que se de <top>Sant Amans,</top> y á cierto
parte con otro camino del Manso por qual
se de dicho Camino del Ma la Riera, y parte
con dicho camino de <top>Mar</top> encargado de dicho conador
de parte de dicho precio de censales, esto es
de precio cien libras barcelonesas, que hizo
</TextRegion>
<TextRegion>
  Bisbal
</TextRegion>
```

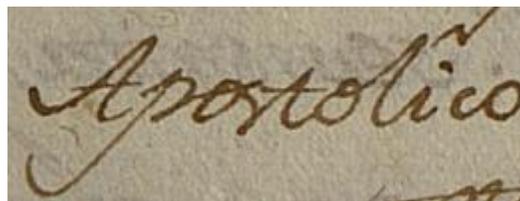
3. errores difíciles de enmendar

- antropónimos



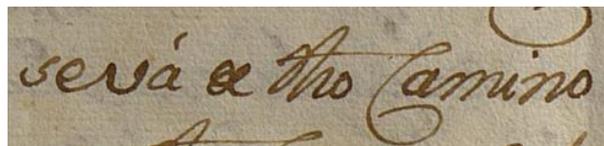
Puig</ant>

- etiquetado y transcripción incorrectos



<ant>Antonio

- palabras no transcritas



se de dicho Camino

**¿Qué queremos hacer
con el texto transcrito y etiquetado?**

Objetivo general

- Estudio de los procesos de cambio social desde mediados del siglo XVIII hasta mediados del siglo XIX

existencia de procesos de cambio social previos a la crisis del antiguo régimen



Emergencia de un grupo de trabajadores 'relativamente' enriquecidos

MENESTRALS



Consolidación de un grupo de propietarios de masos, con orígenes campesinos, que se convierte en clase dominante

HISENDATS

Objetivos analíticos

1. Análisis de trayectorias individuales (enfoque prosopográfico)

- ejemplo 1: Hipótesis sobre la pérdida de cualificación (∇ skill premium, ∇ cuota catastral) y posible descenso social de los **albañiles**
 - identificar y aislar a los albañiles
 - seguir sus pautas de movilidad intergeneracional (capítulos matrimoniales)
 - seguir sus movimientos en el mercado inmobiliario y del crédito
- ejemplo 2: Hipótesis sobre la acumulación patrimonial de los **grandes arrendatarios (masovers)**
 - identificarlos y aislarlos
 - reconstruir su actividad en el mercado inmobiliario
 - analizar su movilidad

Objetivos analíticos

2. Análisis del funcionamiento del **mercado** (de la tierra, del crédito, matrimonial,...)
 - ejemplo: el mercado formal del **crédito**
 - aislar las tipologías documentales referidas a fórmulas de préstamo (censal, deudor, obligación, venta a carta de gracia,...)
 - identificar a acreedores y deudores (a escala individual y social)
 - medir los flujos de crédito: volumen, geografía,...

Objetivos analíticos

3. Análisis de procesos colectivos de **cambio social**

- ejemplo: Hipótesis sobre la **emergencia de un nuevo grupo social** (*menestrals*) surgido de las filas de los *treballadors*
 - identificar los orígenes sociales de los que, entre 1795 y 1805, se autoidentifican como *menestrals*
 - delimitar sus condiciones de vida material (mercado inmobiliario) y cronología de los procesos de acumulación
 - analizar sus estrategias matrimoniales (con quién se casan, importe dote recibida)
 - ...

¿Qué necesitamos para empezar a trabajar “en serio”?

para observar cambios sociales se necesita **tiempo**

= masa documental suficiente para seguir trayectorias durante un período algo largo

1768

1805

En proceso

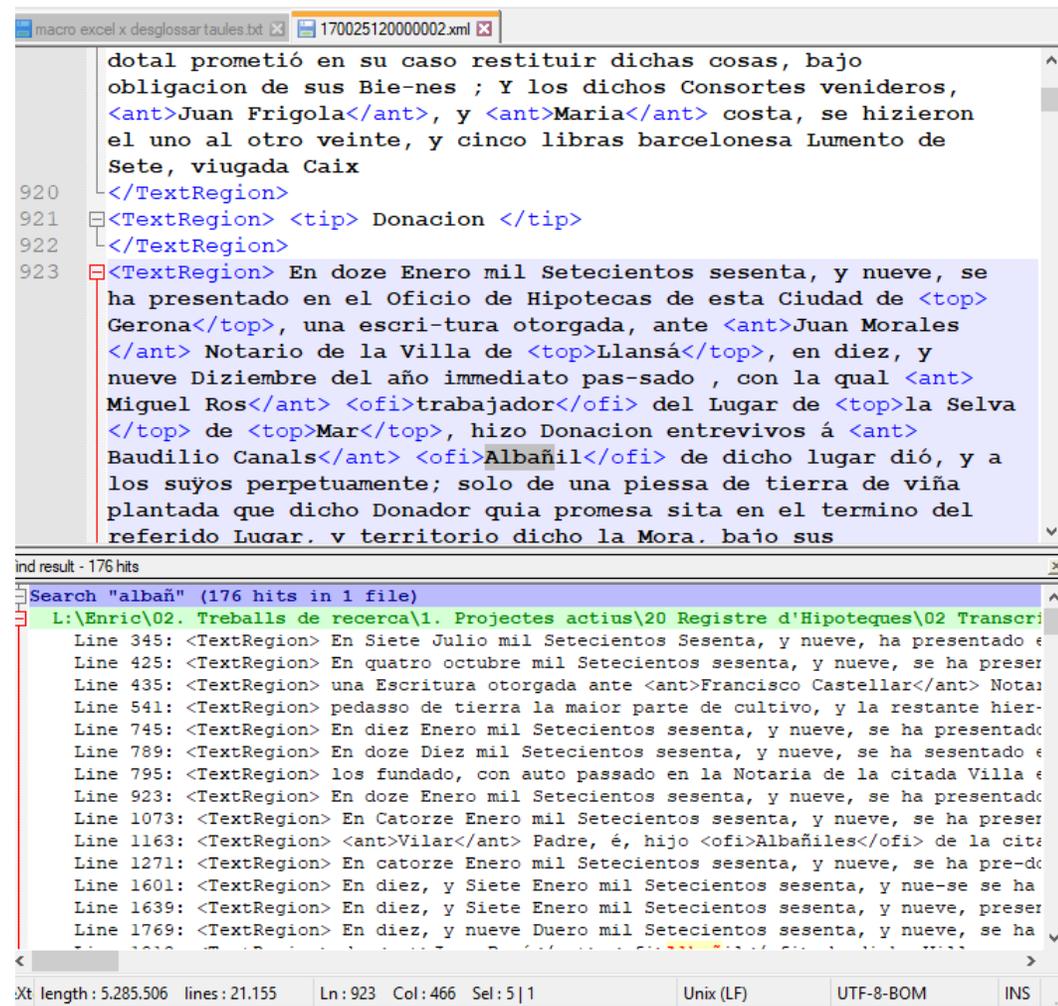
¿Cómo explotar las transcripciones?

Búsqueda del termino
"albañ" (albañil) con
Notepad++

Opción 1: Exploración directa de los archivos de texto plano

búsqueda directa a través de palabras clave

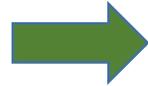
- simplicidad
- uso primario
- accesible a cualquier usuario (interno o externo)



```
macro excel x desglossar taules.txt 170025120000002.xml
dotal prometió en su caso restituir dichas cosas, bajo
obligacion de sus Bie-nes ; Y los dichos Consortes venideros,
<ant>Juan Frigola</ant>, y <ant>Maria</ant> costa, se hizieron
el uno al otro veinte, y cinco libras barcelonesa Lumento de
Sete, viugada Caix
920 </TextRegion>
921 <TextRegion> <tip> Donacion </tip>
922 </TextRegion>
923 <TextRegion> En doze Enero mil Setecientos sesenta, y nueve, se
ha presentado en el Oficio de Hipotecas de esta Ciudad de <top>
Gerona</top>, una escri-tura otorgada, ante <ant>Juan Morales
</ant> Notario de la Villa de <top>Llansá</top>, en diez, y
nueve Diziembre del año immediato pas-sado , con la qual <ant>
Miguel Ros</ant> <ofi>trabajador</ofi> del Lugar de <top>la Selva
</top> de <top>Mar</top>, hizo Donacion entrevivos á <ant>
Baudilio Canals</ant> <ofi>Albañil</ofi> de dicho lugar dió, y a
los suyos perpetuamente; solo de una piessa de tierra de viña
plantada que dicho Donador quia promesa sita en el termino del
referido Lugar. v territorio dicho la Mora. baño sus
Find result - 176 hits
Search "albañ" (176 hits in 1 file)
L:\Enric\02. Treballs de recerca\1. Projectes actius\20 Registre d'Hipoteques\02 Transcri
Line 345: <TextRegion> En Siete Julio mil Setecientos Sesenta, y nueve, ha presentado e
Line 425: <TextRegion> En quatro octubre mil Setecientos sesenta, y nueve, se ha preser
Line 435: <TextRegion> una Escritura otorgada ante <ant>Francisco Castellar</ant> Notar
Line 541: <TextRegion> pedasso de tierra la maior parte de cultivo, y la restante hier
Line 745: <TextRegion> En diez Enero mil Setecientos sesenta, y nueve, se ha presentad
Line 789: <TextRegion> En doze Diez mil Setecientos sesenta, y nueve, se ha sesentado e
Line 795: <TextRegion> los fundado, con auto passado en la Notaria de la citada Villa e
Line 923: <TextRegion> En doze Enero mil Setecientos sesenta, y nueve, se ha presentad
Line 1073: <TextRegion> En Catorze Enero mil Setecientos sesenta, y nueve, se ha preser
Line 1163: <TextRegion> <ant>Vilar</ant> Padre, é, hijo <ofi>Albañiles</ofi> de la cita
Line 1271: <TextRegion> En catorze Enero mil Setecientos sesenta, y nueve, se ha pre-de
Line 1601: <TextRegion> En diez, y Siete Enero mil Setecientos sesenta, y nue-se se ha
Line 1639: <TextRegion> En diez, y Siete Enero mil Setecientos sesenta, y nueve, preser
Line 1769: <TextRegion> En diez, y nueve Duero mil Setecientos sesenta, y nueve, se ha
Xt length: 5.285.506 lines: 21.155 Ln: 923 Col: 466 Sel: 5 | 1 Unix (LF) UTF-8-BOM INS
```


¿Cómo explotar las transcripciones?

Opción 2: generación de una base de datos estructurada



estructura básica:

- referencia notarial (notario y notaría)
- tipo de operación
- fecha
- actores (con oficio y residencia)
- lugares
- valor o precio de la operación

mayor potencialidad analítica



facilita la normalización

consultas complejas,
facilidad de agrupación
facilidad de recuento

mayores riesgos de pérdida de información



derivada de



errores de transcripción

errores de etiquetaje

errores de estructuración

Un ejemplo de consulta en base de datos: Oficios con más de 100 referencias

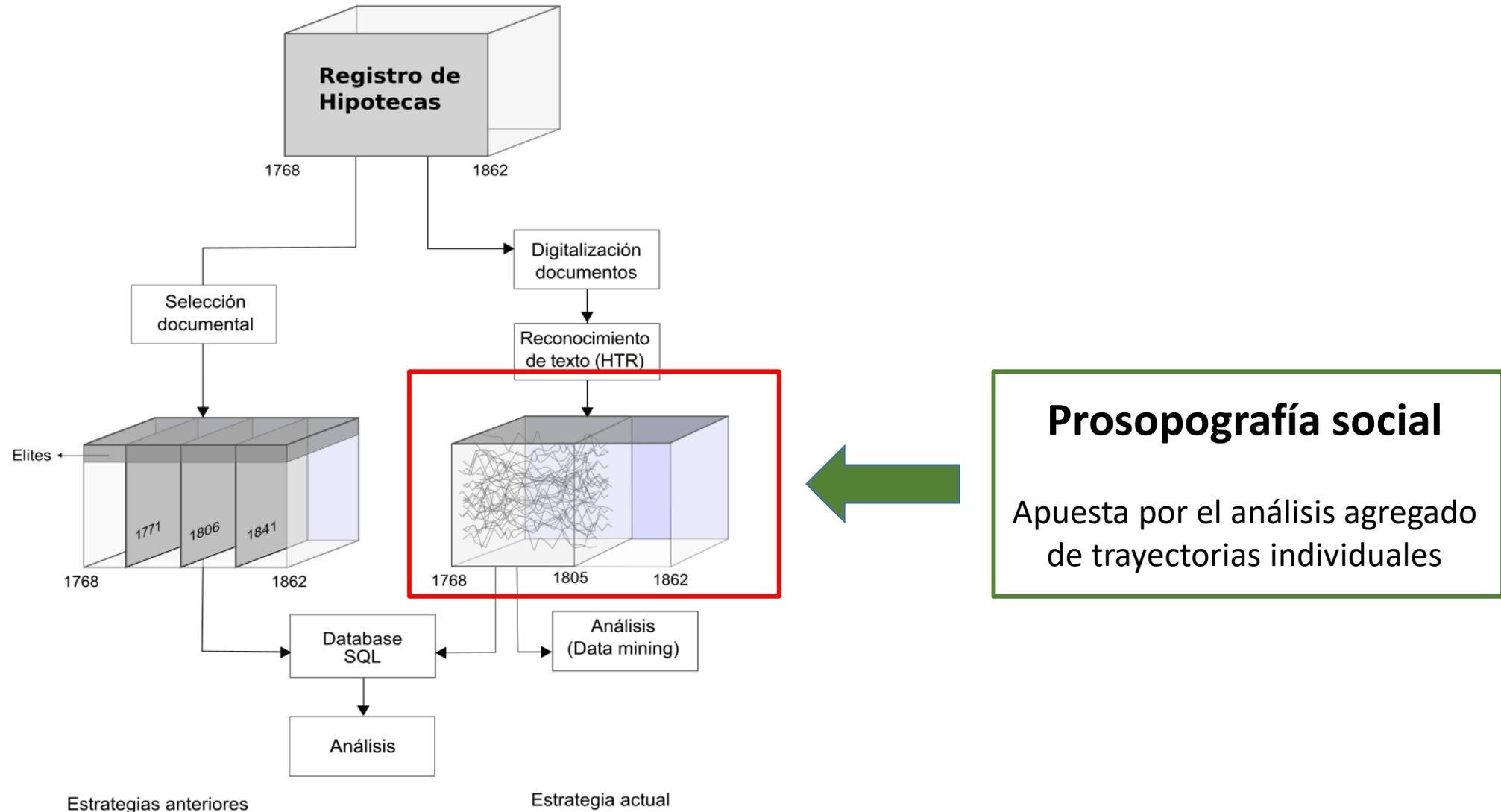
3 libros
(junio 1768-1769)

Notario	8319
labrador	6900
trabajador	4394
jornalero	1104
Negociante	618
presbítero	596
Albañil	463
Marinero	417
tejedor	368
Herrero	336
Carpintero	324
Sastre	321
Ciudadano	307
zapatero	281
Pescador	242
Cirujano	223
Paraire	217
Boticario	193
Soguero	169
Droguero	166
tendero	153
comerciante	131
noble	117
Arriero	114
carretero	114
Molinero	110



objetivo prioritario

Otro ejemplo: la reconstrucción de trayectorias individuales

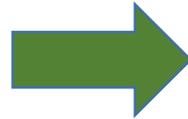


¿Es posible reconstruir trayectorias individuales?

Extracción de antropónimos,
oficios y topónimos

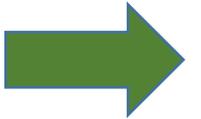
+

posición dentro del documento



Reconstrucción de la vinculación
antropónimo - oficio - lugar

criterio de proximidad



	text	tipusCamp	document	ordre
▶	Gerona	top	1	25
	Pedro Pages	ant	1	30
	Notario,	ofi	1	33
	Barcelona	top	1	41
	Amerio Camps y Pages	ant	1	54
	Bracero	ofi	1	59
	Torroella de Mondel	top	1	64
	Doctor	ofi	1	86
	Juan Cortada y Torras	ant	1	90
	Barcelona	top	1	99
	Gerona	top	2	25
	Antonio Virell	ant	2	35
	Canet	top	2	48
	Mar	top	2	50
	Felix Llauger	ant	2	67
	Marinero	ofi	2	70
	Canet de Mar	top	2	75
	Gerona	top	2	81
	Cathalina Llauger,	ant	2	84
	Cruañas	ant	2	88
	Patron	ofi	2	92
	Felix Llauger	ant	2	93
	Cathalina Cruañas y Gispert	ant	2	104
	Pedro Planas	ant	2	113
	Cathalina	ant	2	119
	Raymundo Juana	ant	2	125
	Gerona	top	2	81
	Patron	ofi	2	132

DATOS ORIGINALES				EXTRACCIÓN				
text	tipusCamp	document	ordre	nom	ofici	residència	document	ordre
Gerona	top	1	25					
Pedro Pages	ant	1	30	Pedro Pages	Notario,	Barcelona	1	1
Notario,	ofi	1	33					
Barcelona	top	1	41					
Amerio Camps y Pages	ant	1	54	Amerio Camps y Pages	Bracero	orroella de Monde	1	2
Bracero	ofi	1	59					
Torroella de Mondel	top	1	64					
Doctor	ofi	1	86					
Juan Cortada y Torras	ant	1	90	Juan Cortada y Torras		Barcelona	1	3
Barcelona	top	1	99					
Gerona	top	2	25					
Antonio Virell	ant	2	35	Antonio Virell		Canet	2	1
Canet	top	2	48					
Mar	top	2	50					
Felix Llauger	ant	2	67	Felix Llauger	Marinero	Canet de Mar	2	2
Marinero	ofi	2	70					
Canet de Mar	top	2	75					
Gerona	top	2	81					
Cathalina Llauger,	ant	2	84	Cathalina Llauger,			2	3
Cruañas	ant	2	88	Cruañas	Patron		2	4
Patron	ofi	2	92					
Felix Llauger	ant	2	93	Felix Llauger			2	5
Cathalina Cruañas y Gispert	ant	2	104	Cathalina Cruañas y Gispert			2	6
Pedro Planas	ant	2	113	Pedro Planas			2	7
Cathalina	ant	2	119	Cathalina			2	8
Raymundo Juana	ant	2	125	Raymundo Juana	Patron		2	9
Patron	ofi	2	132					

¿Es posible reconstruir trayectorias individuales?

Normalización de oficios y topónimos

tablas de equivalencia

ofici	ofici normalitzat
Albanil	Albañil
Albanyil	Albañil
Albanyiles,	Albañil
Albanyill	Albañil
Albañá,	
Albañil	Albañil
Albañil,	Albañil
Albañil;	Albañil
Albañiles	Albañil
Albardera	Albardo
Albardera,	Albardo
Albardo	Albardo
Albardo,	Albardo
Albaro	
Albaxador	Abaixador
Albeitre	Albeitre
Alcalde	Alcalde
Alcon	
Alde	
Alfarero	Alfarero
Alferero	Alfarero
Alferrero	Alfarero
AlSester	
Aluna	
Alvila	

Identificación nominal (nominal record linking)

criterios

Opción A: similitud nombre y apellido + oficio + residencia

Opción B: similitud nombre y 2 apellidos + oficio + no contradicción residencia

criterios muy restrictivos

dificultad

gran variabilidad en la escritura + errores de transcripción

escasa utilidad algoritmos de codificación (soundex,...)

identificación manual

¿Es posible reconstruir trayectorias individuales?

Identificación de trabajadores

- selección:
- trabajador
 - jornalero
 - bracero

nom	ofici normalitzat	residència	id persona	document	ordre
Agustin Costa Mora Francisco	trabajador			2199	102
Agustin Gibert	trabajador		Agustin Gibert	3143	2
Agustin Gibert	trabajador		Agustin Gibert	3164	3
Agustin Gimpera	trabajador			2678	2
Agustin Matheu	trabajador			5412	1
Agustin Moliner	jornalero	Albons		1269	2
Agustín Ribas y Serra	trabajador			5721	1
Agustin Sabell	trabajador		Agustin Sabell	1529	4
Agustin Sabell	trabajador		Agustin Sabell	1529	7
Agustin Serra	trabajador	Sant Dalmas,		2336	4
Agustin vilanova	trabajador	Vilallonga,	Agustin vilanova	5883	1
Agustin vilanova	trabajador	vilallonga,	Agustin vilanova	5882	1
Ajats	trabajador	Terrades,		6274	1
Aldon Casas	trabajador			3776	1
Alemany	trabajador			2111	5
Alexo Carbonell	trabajador			1437	7
Alexo Carreras	trabajador	Cassá de la Selva		6253	1
Alexo Carreras	trabajador	Santa Pallaya;		6313	1
Alexo Delforn	jornalero			1264	2
Alexo Llauger y Lareba	trabajador	San Jayme de San Pol	Alexo Llauger y Lareba	7163	1
Alexo Llauger y lareba	trabajador	San Pol	Alexo Llauger y Lareba	3008	2
Alexo Llauger y Llazeba	trabajador	San Pol	Alexo Llauger y Lareba	7007	1
Alexo Ros	trabajador	Gerona		1823	13
Alexos Carreras	trabajador			6251	1
Alexos Carreras	trabajador	Gerona		6252	1
Aloy	trabajador			3518	1

código único asignado manualmente

Revisable

No cumplen criterios de identificación

Resultados del proceso de vinculación de trabajadores

**Identificación de
trabajadores**

**3 libros
(1768-1769)**

58.244 registros
nominales
(28.872 con oficio)



5.582 registros nominales
referidos a trabajadores



**Vinculación de
registros
nominales**



862 vinculados



4.720 no vinculados



361 trabajadores

núm. vinculaciones	núm. trabajadores
18	1
9	2
8	1
5	6
4	19
3	57
2	277
	363

Próximo paso

determinar el rol de los trabajadores
en las operaciones en que participan