

A edición dixital de textos

Potenciais e problemas

Victor Millet

(Universidade de Santiago de Compostela)

Qué é unha edición dixital?

Texto en arquivo dixital nativo

Normalmente en linguaxe dixital XML

Recoméndase seguir directrices TEI

Interactuable: non é pdf en web nin html estático

Transformación de texto en datos

Datos deben ser usables para análises de todo tipo

Perde moita información cando se imprime o texto

Capacidade de presentar moita información contextual

Tamaño

Condicións

- A edición dixital non é para documentos pequenos
- Os datos dixitais son interesantes a partir de certo número
- O esforzo en tempo e recursos xustifícase a partir dun umbral

Dificultades

- Moitos documentos e moi extensos
- Diversidade de material
- Procesos complexos que esixen xestión profesionalizada

Codificación XML/TEI

- O arquivo XML/TEI contén toda a información textual: transcripción, edición e metadatos
- Compre:
 - coñecer ben os textos e os seus fenómenos
 - Estabrecer con detalle o manual de codificación
 - Coñecer a visualización final
 - Saber qué uso a maiores se le pode querer dar ao material

Exemplos de codificación /1

<p>En tanto que de rofa i açucena</p>

<p>fe muestra la color en vuestro gesto,</p>

<p>i que vuestro mirar ardiente, onefto</p>

<p>enciende al coraçon i lo refrena;</p>

<p>I en tanto qu'el cabello, qu'en la vena</p>

<p>del oro f'escogio, con vuelo prefto</p>

<p>por el hermofo cuello blanco, enhiesto</p>

<p>el viento mueve, esparze i defordena;</p>

<p>Coged de vuestra alegre primavera</p>

<p>el dulce fruto, antes q̄ el tiempo airado</p>

<p>cubra de nieve la hermoſa cumbre.</p>

<p>Marchitarà la rofa el viento elado,</p>

<p>todo lo mudarà la edad ligera,</p>

<p>por no hazer mudança en fu costumbre.</p>

Exemplos de codificación /2

```
</lg>
  <| n="1">En tanto que de rofa i açucena</|>
  <| n="2">fe muestra la color en vuestro gesto </|>
  <| n="3">i que vuestro mirar ardiente, onefto</|>
  <| n="4">enciende al coraçon i lo refrena</|>
</lg>
</lg>
</lg>
  <| n="5">I en tanto qu'el cabello, qu'en la vena</|>
  <| n="6">del oro f'escogio, con vuelo prefto</|>
  <| n="7">por el hermofo cuello blanco, enhiefto</|>
  <| n="8">el viento mueve, esparze i defordena</|>
</lg>
</lg>
  <| n="9">Coged de vuestra alegre primavera</|>
  <| n="10">el dulce fruto, antes q̄ el tiempo airado</|>
  <| n="11">cubra de nieve la hermofo cumbre</|>
</lg>
</lg>
  <| n="12"> Marchitarà la rofa el viento elado</|>
  <| n="13"> todo lo mudarà la edad ligera</|>
  <| n="14"> por no hazer mudança en fu costumbre.</|>
</lg>
```

Exemplos de codificación /3

```
<lg ana="cuarteto" n="1">
```

```
<l n="1"><w>En</w> <w>tanto</w> <w>que</w> <w>de</w> <w>rofa</w> <w>i</w> <w>açucena</w></l>
```

```
<l n="2"><w>fe</w> <w>muestra</w> <w>la</w> <w>color</w> <w>en</w> <w>vueftro</w> <w>gefto</w><pc>,</pc></l>
```

```
<l n="3"><w>i</w> <w>que</w> <w>vueftro</w> <w>mirar</w> <w>ardiente</w><pc>,</pc> <w>onefto</w></l>
```

```
<l n="4"><w>enciende</w> <w>al</w> <w>coraçon</w> <w>i</w> <w>lo</w> <w>refrena</w><pc>;</pc></l>
```

```
</lg>
```

```
<lg ana="cuarteto" n="2">
```

```
<l n="5"><w>l</w> <w>en</w> <w>tanto</w> <w>qu</w><pc>'</pc><w>el</w> <w>cabello</w><pc>,</pc> <w>qu</w><pc>'</pc>
```

```
<w>en</w> <w>la</w> <w>vena</w></l>
```

```
<l n="6"><w>del</w> <w>oro</w> <w>f</w><pc>'</pc><w>efcogio</w><pc>,</pc> <w>con</w> <w>vuelo</w> <w>prefto</w></l>
```

```
<l n="7"><w>por</w> <w>el</w> <w>hermofo</w> <w>cuello</w> <w>blanco</w><pc>,</pc> <w>enhiefto</w></l>
```

```
<l n="8"><w>el</w> <w>viento</w> <w>mueve</w><pc>,</pc> <w>elparze</w> <w>i</w> <w>defordena</w><pc>;</pc></l>
```

```
</lg>
```

Exemplos de codificación /4

```
<pb n="1"/>
  <lg ana="cuarteto" n="1">
    <lb n="1"/>
      <l n="1">
        <w>En</w>
        <w>tanto</w>
        <w>que</w>
        <w>de</w>
        <w>rofa</w>
        <w><choice><orig>i</orig><reg>y</reg></choice></w>
        <w>a<choice><orig>ç</orig><reg>z</reg></choice>ucena</w>
      </l>
    <lb n="2"/>
      <l n="2">
        <w>fe</w>
        <w>muestra</w>
        <w>la</w>
        <w>color</w>
        <w>en</w>
        <w>vueftro</w>
        <w>gefeto</w>
        <pc>,</pc>
      </l>
    </lb>
  </lg>
</pb>
```


Exemplos de codificación /5

```
<pb n="1"/>
<lg ana="hc:Quatrain" n="1">
  <lb n="1"/>
  <l n="1" xml:id="l_1">
    <w xml:id="w_1_1">En</w>
    <c> </c>
    <w xml:id="w_1_2">tanto</w>
    <c> </c>
    <w xml:id="w_1_3">que</w>
    <c> </c>
    <w xml:id="w_1_4">de</w>
    <c> </c>
    <w xml:id="w_1_5">ro<choice><orig>f</orig><reg>s</reg></choice>a</w>
    <c> </c>
    <w xml:id="w_1_6"><choice><orig>i</orig><reg>y</reg></choice></w>
    <c> </c>
    <w xml:id="w_1_7">a<choice><orig>ç</orig><reg>z</reg></choice>uc<rhyme type="e-a" n="a">ena
      </rhyme></w>
  </l>
```

Exemplos de codificación /5

```
<l n="6152" xml:id="l_6152" hei:altN="6144">
  <w xml:id="w_6152_1"><choice><orig>l</orig><reg>i</reg></choice>r</w>
  <c> </c>
  <w xml:id="w_6152_2">vin<metamark ana="hc:Hyphen">=</metamark><lb n="3" break="no"/>det</w>
  <c> </c>
  <w xml:id="w_6152_3">no<choice><orig>&combdblac;</orig><reg>&combuml;</reg></choice>ch</w>
  <choice><orig/><reg ana="hc:PunctuationRegularization"><pc>!</pc></reg></choice>
  <c> </c>
  <w xml:id="w_6152_4">da<choice><sic>n</sic><corr>s</corr></choice></w>
  <c> </c>
  <w xml:id="w_6152_5"><hi rendition="hc:RedStroke">i<hi>ft</w>
  <c> </c>
  <w xml:id="w_6152_6">mei<choice><am>&bar;</am><ex>n</ex></choice></w>
  <c> </c>
  <w xml:id="w_6152_8">rat</w>
  <choice><orig><pc> </pc></orig><reg ana="hc:PunctuationRegularization"><pc>,</pc></reg></choice>
  <c> </c>
</l>
```

<note ana="hc:TextConstitutionNote" target="#l_6152">Die ersten beiden Worte ließen sich auch, in Anlehnung an die übrigen Handschriften, korrigieren in <quote>irwindet</quote>.

Ferramentas associadas 1:

Ferramentas HTR

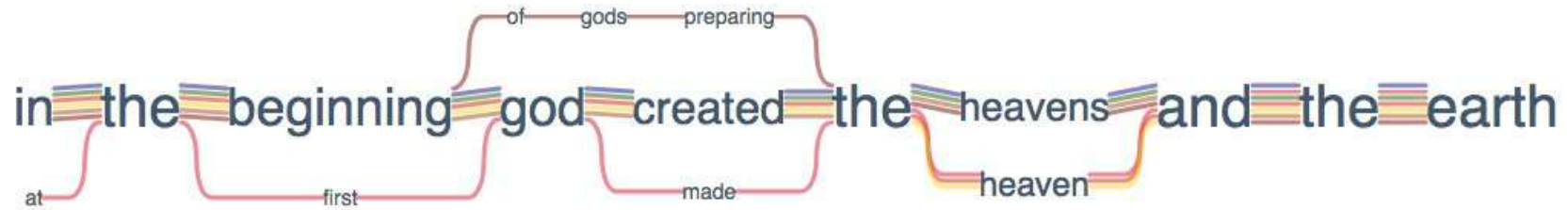
The screenshot displays the eScriptorium web application interface. At the top, there is a navigation bar with 'Home' and 'Contact' links, and a user profile section for 'Hello Demo'. Below this is a secondary navigation bar with tabs for 'Description', 'Ontology', 'Images', 'Edit', 'Models', and 'Reports'. The main content area is split into two panels. The left panel shows a scanned image of a handwritten document in French, with a large '3' and 'Ona' written at the top. The right panel shows the transcribed text of the document, with line numbers 1 through 25 on the left and the corresponding text on the right. The text is a poem by Guillaume Apollinaire titled 'Alcools'. The interface also includes a search bar with the word 'manual' entered, and various utility icons for zooming, panning, and other document navigation functions.

escriptorium Home Contact My Projects My Models Hello Demo

Description Ontology Images Edit Models Reports Guillaume Apollinaire, Alcools. Page 5 - 4_caf4b2_default.jpg - (2616x3308) - 660.91 KB manual

1 Bergère ô tour Eiffel le troupeau des ponts bêle ce matin
2 Tu en a assez de vivre dans l'antiquité grecque et romaine
3 Ici même les automobiles ont l'air d'être anciennes
4 La Religion seule est restée toute neuve la Religion
5 Est restée simple comme les langues de Port-Aviation
6 Seul en Europe tu n'es pas antique ô Christianisme
7 L'Européen le plus moderne c'est vous Pape Pie X
8 Et toi que les fenêtres observent la honte te retient
9 D'entrer dans une église et de t'y confesser ce matin
10 Tu lis les prospectus, les catalogues les affiches qui chantent tout haut
11 Voilà la poésie ce matin et pour la prose il y a les journaux
12 Il y a les livraisons à vingt cin 25 centimes pleines d'aventures policières
13 O portraits des grands hommes et mille titres divers
14 J'ai vu ce matin une jolie rue dont j'ai oublié le nom
15 Neuve et propre du soleil elle est le clairon
16 Les directeurs les ouvriers et les belles sténo-dactylographes
17 Du lundi matin au samedi soir quatre fois par jour y passent
18 Le matin par trois fois la sirène y gémit
19 Une cloche rageuse y aboie vers midi
20 Les inscriptions des enseignes et des murailles
21 Les plaques les avis à la façon des perroquets crient
22 J'aime la grâce de cette rue industrielle
23 Située à Paris entre la rue Aumont-Thiéville et l'avenue des Ternes
24 A la fin tu es las de ce monde ancien
25 Zone

Ferramentas asociadas 1: Colación automatizada



Shared By: Dana Wheelles
Description: From Stuart Curran's digital edition at Romantic Circles...

Mary Shelley's Frankenstein (full)

juxta commons beta

Need Help? Click here

Sign in to your account | Sign up for account

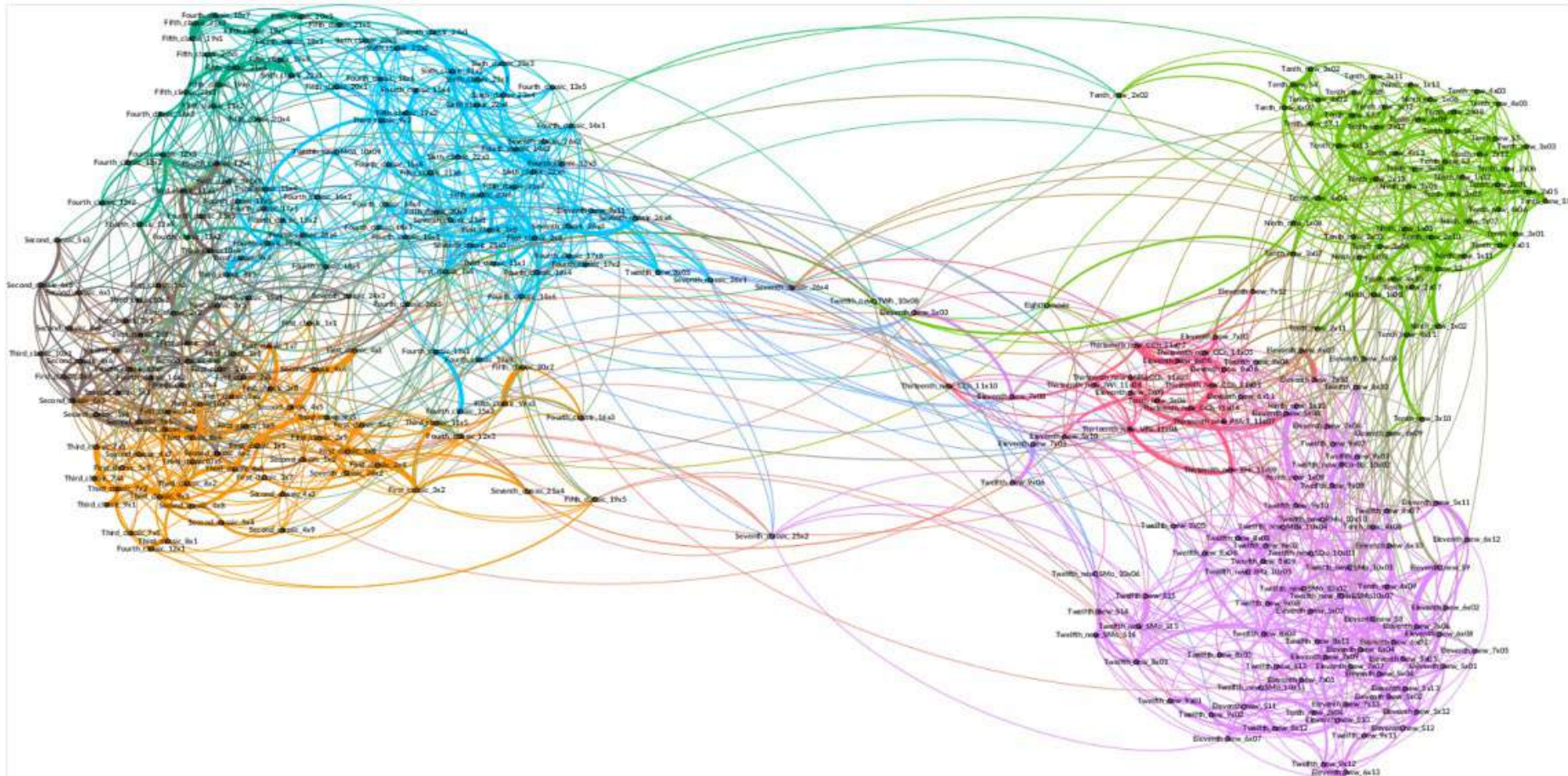
Witness List	Sort by	Frankenstein (1818)
Frankenstein (1818)	[base]	overwhelmed by disappointments; yet when he has retired into himself, he will be like a celestial spirit, that has a halo around him, within whose circle no grief or folly ventures.
Frankenstein (1831)	Difference from base	Will you laugh at the enthusiasm I express concerning this divine wanderer? If you do, you must have certainly lost that simplicity which was once your characteristic charm. Yet, if you will, smile at the warmth of my expressions, while I find every day new causes for repeating them.

Witness Differences

▲ Frankenstein (1831)

... merits of this wonderful man. Sometimes I have endeavoured to discover what quality it is which he possesses, that elevates him so...

Ferramentas associadas 3: Ferramentas de estatística



Exemplo de edición con notas e variantes

Ein ritter so gelert was,
daz er an den buchen las;
swan er sine stunde
nicht bas bewenden chunde,
25 das er tihtens phlach:
daz man gerne horen mach,
sinen vliz leit er dar an.
er was geheizen Hartman
und was ein Ouware.
30 der tihte ditz mære.
Ez het der chunch Artus
ze Charidol in sinem hus
ze einen phingsten geleit
nach siner gewonheit
35 ein so riche hochzeit,
daz er vor des noch seit
so riche nie gewan.
destwar, da was ein swacher man
in liehtem werde,
40 wan sich besament uf der erde
weder da noch anderswa
so manich gut ritter als da.
ouch ward in von hove gegeben
in alle wis ein wunscheleben:
45 in liebt hof und den lip
manch schone magt und wip,
die besten von den richen.
mich jamert warlichen
und hulf ez iht, ich wolde ez chlagen,
50 daz nu bei unsern tagen
solch vraude nicht werden mach,
der man bei sinen ziten phlach.

Content Notes Imprint Text display Annotations

V. 1
B,A,D,J,b,c,l,p,u,z,P: swer an rehte gv̄te
d: wer an zerte guete
r: So wer an rechte gv̄te

V. 10
B,D,J,b,c,l,p,r,u,z,P: daz er der eren chrone
A: dazer der erden crone
d: daz Er der Eeren ein kronc

V. 11
B,A,D,J,b,c,d,u: do truc und noch sin nam treit
l: Do trug vnd noch trayt
p,r: Drug vnd nach sin nāmen dreit
z: trug vnd noch sie nieman tret

V. 14
B,A,J,l,p,u,z: Si (p: Die) ichent er lebe noch hīfte
D,r: sprechen er lebe noch hute
b,c: Vnd wenne er lebe (c: lebt) noch hūte
d: S̄y ichent Er lej̄he noch heūte

V. 12–14 verstehen wir als ‚apo koinú‘-Konstruktion: *sine lantlute* gehört sowohl zum vorigen wie auch zum nachfolgenden Vers. Eindeutige Fälle dieser Art brechen wir nicht durch neuzeitliche Satzzeichen.

V. 29
B,A,D,b,c,r,u,z: v̄n was eín öwære
J: Vnd ist ain auwer
d: vnd was auch wære

Das *das* der Handschrift ist eigentlich grammatisch korrekt, ergibt aber im Satz keinen richtigen Sinn, daher emendieren wir zu *da*.

V. 45
B,A,J,u: In liebet den hof v̄n den lip
D: Liebt hof vnd den lip

Complexidade técnica e obxectivo filolóxico

- Se etiqueta ou 'codifica' o que se quere amosar...
- ...ou o que técnicamente se pode amosar.
- Pero también se pode codificar máis e amosar menos
⇒ maior riqueza de datos

Usabilidade dos datos

Principios FAIR da ciencia:

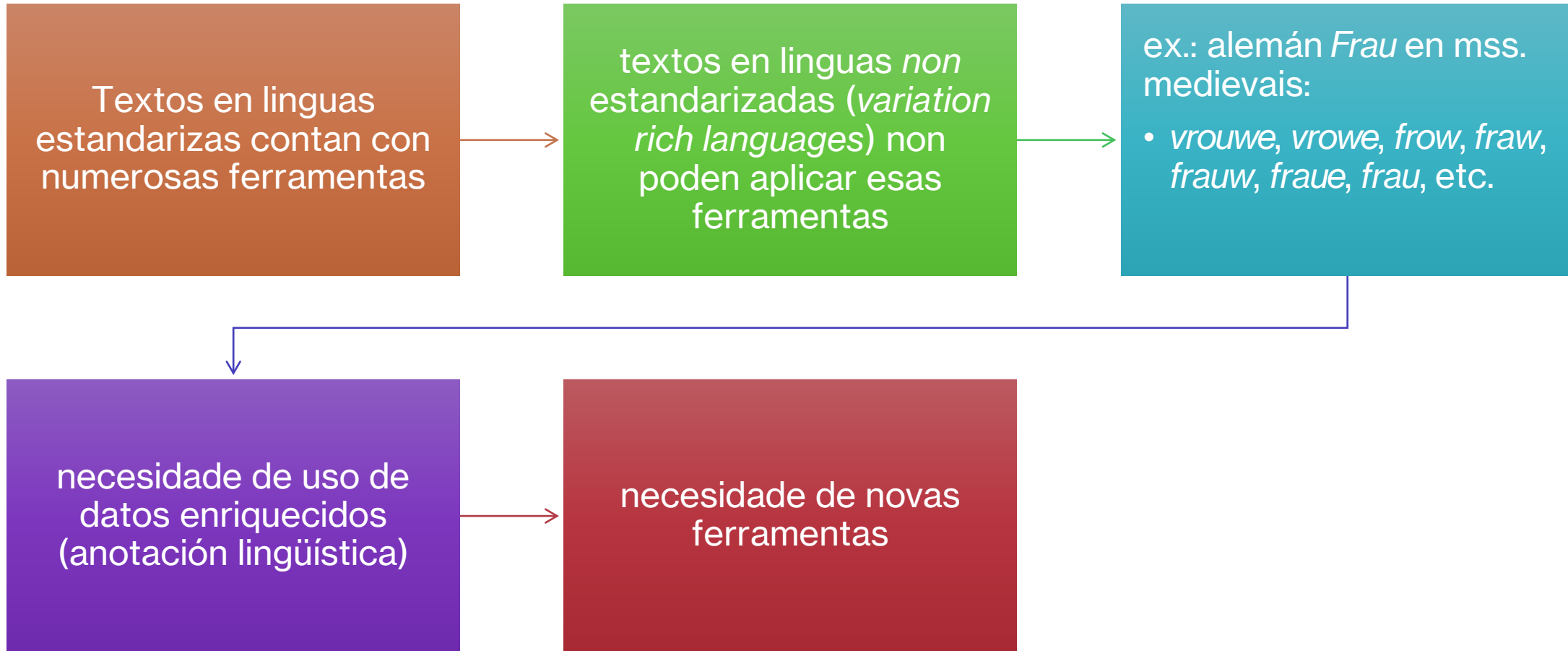
findable, accessible, interoperable, re-usable

Uso dos datos para análises da propia edición

Xustifica o maior esforzo e os maiores custos

Recomendable para recoñecemento e financiación

Dificultades lingüísticas



Transformación e visualización

Visor de imágenes

Visor de texto editado

Transformación de XML en HTML

Visualización interactiva

Visualización de variantes, notas,
materiais, etc.



Lasciate ogne speranza, voi ch'intrate

- ou: Dificultades para facer boas edicións dixitais en España:
 - Ausencia de asistencia técnica para a publicación en web (e para a creación da edición).
 - Ausencia de editoriais
 - Falta de espazos de aloxamento institucional
 - Ausencia de rexistros biblioteconómicos → moitas edicións non son *findable*
 - Non hai garantía para a perseverancia dos datos científicos (!!)

Moitas grazas!

Victor Millet

victor.millet@usc.es

<https://investigacion.usc.gal/investigadores/59667/detalle>

<https://literaturaalemanamedieval.usc.gal>

<https://www.youtube.com/@cuentosMedievales>

This presentation was financed by research grant PID2022-137490NB-I00 by the Spanish Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033/)