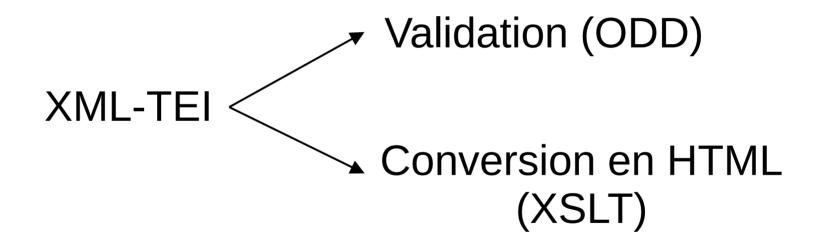
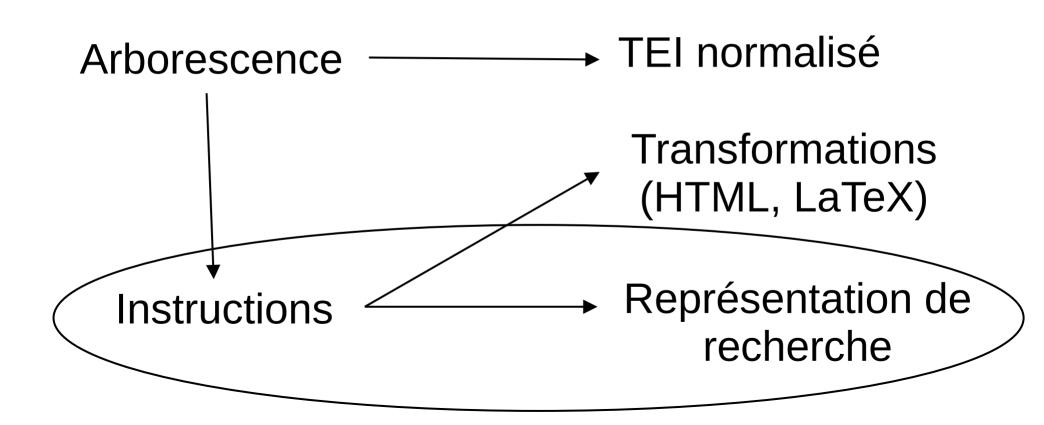
### État actuel



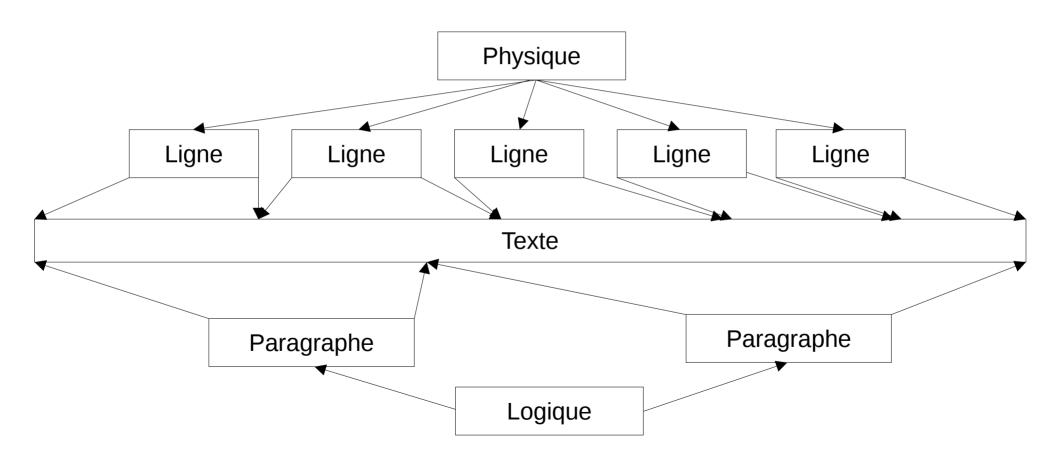
## Rationalisation (1)

Validation (grammaire) XML-TEI → Arborescence

# Rationalisation (2)



#### Arborescence



### Instructions

start-para lang English start-line text "Hello" end-line start-line text "world!" end-line end-para

(Stocké dans la BDD)

### Représentation de recherche

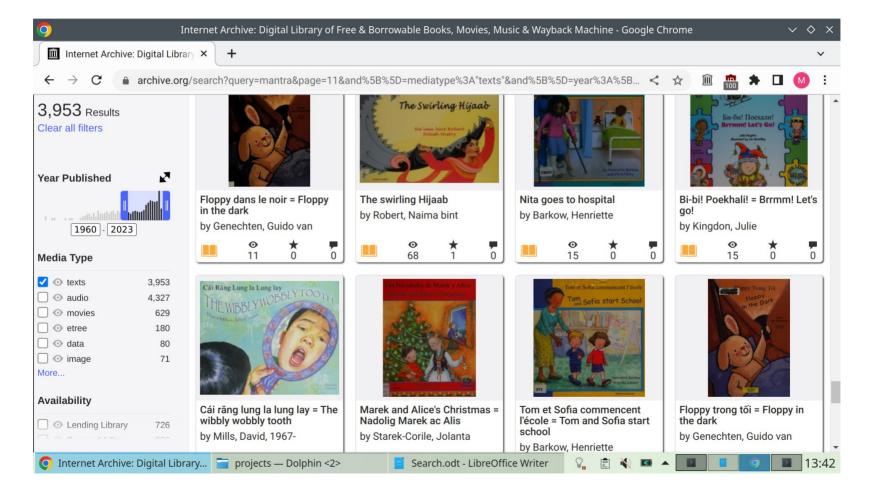
#### Série de champs

 Texte normalisé (regex, recherche approximative) dharma → dh, a, r, m, a

 $\rightarrow$  C, V, C, C, V

Valeurs exactes (comparaison, facettes)

#### **Facettes**



### Quelle base de données ?





Depuis 2000

Depuis 2006

Omniprésent, usages multiples Niche (humanités numériques)

# Popularité



SQLite: 93 700 eXist-db: 534 0.57 %

SQL: 661 000 XQuery: 5300 **0.80** %

#### Standards





C (ISO 1989)

Java (no standard)

SQL (ISO 1986)

XQuery (W3C 2007)

## Objectifs





Seulement une base de données (1 Mb code) Plateforme d'applications (163 Mb code)

Librairie

Framework

#### Structuration





Relationnel

**Document** 

Schéma explicite

Pas de schéma

### Extensions souhaitées

 Expressions régulières approximatives Modification de TRE

Support des facettes
Code ad hoc