

Dopo ALIM 2: i futuri possibili

Paolo Monella, Roberto Rosselli Del Turco

Paolo Monella
Università di Palermo
Università Ca' Foscari - Venice Centre
for Digital and Public Humanities



[ve]dph

Venice Centre for
Digital and Public
Humanities



Roberto Rosselli Del Turco
Università di Torino



Dopo ALIM 2: i futuri possibili

- La situazione attuale
 - Struttura e composizione della biblioteca digitale ALIM 2
- Bilancio
 - Sito ALIM 2
 - Codifica del testo
 - Metadati
- Arche di Noè
 - Testi e metadati
 - ALIM 2, Corpus Corporum, GitHub
- Futuri possibili
 - Nuovi finanziamenti?
 - Nuove collaborazioni?
 - Progetto Arca di Noè (strategia GitHub/Zenodo)
 - Progetto ALIM 2+ / ALIM 3

La questione

Sopravvivenza dell'Archivio
nel medio-lungo periodo

Situazione attuale

Situazione attuale: ALIM 2

- ALIM 2: una **biblioteca digitale** di testi di edizioni critiche
- ALIM 1 → ALIM 2: da archivio (?) a biblioteca digitale
 - sito web separato
 - macro gestione dei contenuti (collezioni tematiche)
 - filtri
 - analisi linguistica dei testi
- Struttura complessa costruita su layer di software diversi
 - sito Web: WP + PHP
 - software DL: Muruca + Solr + MySQL + OxGarage
 - strategia *mix and match*

Situazione attuale: ALIM 2



Situazione attuale: ALIM 2

- In teoria questa struttura potrebbe essere implementata in un unico software
 - uno dei primi: Greenstone <http://www.greenstone.org/>
 - più recentemente: Blacklight <http://projectblacklight.org/>
 - in alternativa: eXist-db <http://exist-db.org/exist/apps/homepage/index.html>
- Vantaggi
 - *all in one*: creazione DL, gestione e manutenzione semplificati
- Svantaggi
 - impossibilità di fare *mix and match* delle tecnologie migliori
 - minor flessibilità nella gestione del *front-end*
 - possibile assenza di funzionalità desiderate

Situazione attuale: ALIM 2

- L'approccio attuale (*mix and match*) è quello di molte DL
 - digilibLT <http://www.digiliblt.unipmn.it/>
 - CC <http://mlat.uzh.ch/MLS/index.php?lang=0>
 - SAWS <http://www.ancientwisdoms.ac.uk/>
 - Perseus <http://www.perseus.tufts.edu/hopper/>
 - Scaife Viewer <https://scaife.perseus.org/>
- Vantaggi
 - massima flessibilità, sia per le caratteristiche sia per l'UI/UX
- Svantaggi
 - possibili problemi di compatibilità fra componenti
 - problemi di manutenzione nel lungo periodo (punto di equilibrio instabile)
 - l'open source non è una panacea

Bilancio

Bilancio: Sito ALIM 2

Obiettivi	Stato	Problemi
Gestione XML e visualizzazione / esportazione in HTML, PDF, txt	✓	• Esportazione XML vuoti
Browse	✓	• Ordinamento autori e testi
Ricerca testuale	✓	• A volte poco intuitiva • Allografi non gestiti
Analisi del testo (Lexicon)	✓	
Metadati	✓	• Gestione livello 0, non livello 1 • Rapporto tra teiHeader e DB sito
Open source	✓ ...	• Memento mori ○ aggiornamenti e sicurezza ○ documentazione codice ○ fondi per la manutenzione

Bilancio: Codifica del testo

Obiettivi	Stato	Problemi
Livello 0 (export da ALIM 1)	✓	• Codifica semplificata
Livello 1 • Markup strutturale (paragrafi, capitoli, versi, battute teatrali...) • Per i vecchi testi (da ALIM 1) • Vedi Manuale di codifica	✓ ...	• J. Ctibor ha adeguato i vecchi testi, ma senza i testi a stampa • I testi ctiboriani sono su GitHub ma non ancora su ALIM2, perché? ○ Caricamento di massa ○ Esportazione XML vuoti
Livello 2 • Disambiguazione corsivo (citazioni, dialoghi...) • Per i nuovi testi	✓ ...	• Non tutti i testi nuovi sono arrivati • Alcuni ci sono giunti al livello 1 ○ Molto corsivo
Livello 3 (edizioni critiche born-digital)	✓	• Codifica non standardizzabile

Bilancio: Metadati

Obiettivi	Stato	Problemi
Metadati nel teiHeader, livello 0 (export da ALIM 1)	✓	● Visualizzati dal sito ALIM 2, ma non strutturati ○ <notesStmt> / <note> o <sourceDesc> / <bibl>
Verifica e integrazione metadati bibliografici	✓	● C. Casali, ma nei “Metadata custom fields” del sito ALIM2
Metadati nel teiHeader, livello 1 (schema TEI più completo e processabile)	✓ ...	● Testi nuovi: OK ● Testi vecchi: J. Ctibor ha importato nei teiHeader dei file XML i metadati verificati/integrati ○ I testi ctiboriani sono su GitHub ma non ancora su ALIM 2 (vd. sopra)
Integrazione tra metadati del teiHeader e sito ALIM2	✗	● ALIM 2 non importa i metadati dal teiHeader/livello 1 ○ Ne importa alcuni che erano già al livello 0 ○ Altri vanno inseriti sul sito al caricamento ○ Metadati duplicati (teiHeader / sito)

Arca di Noè

Arca di Noè

- Arche
 - ALIM 2
 - Farlo funzionare al meglio
 - Manutenzione nel lungo periodo
 - Corpus Corporum
 - GitHub
- Colomba
 - I metadati verificati e/o aggiornati sono ora nel `teiHeader` dei file TEI XML `ctiboriani`



Futuri possibili

Futuri possibili

- Nuovi finanziamenti
 - Ulteriore sviluppo DL
- Collaborazioni con altre DL
 - CLARIN
 - Corpus Corporum
- Progetto “Arca di Noè”
 - GitHub/Zenodo
- Progettare ALIM 3
 - non domani, e nemmeno dopodomani



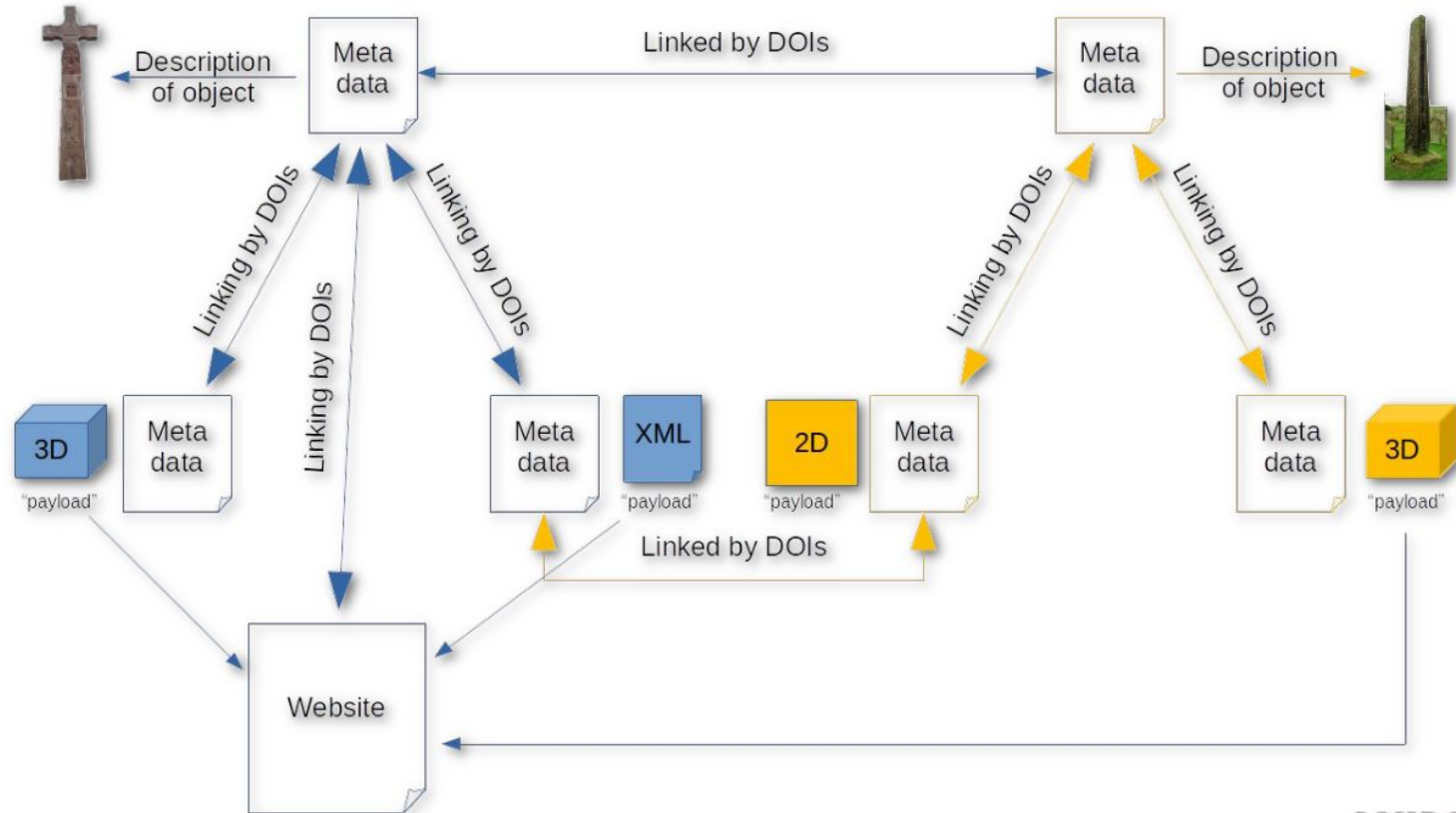
Futuri possibili: nuovi finanziamenti

- Ottima base di partenza
 - risorsa utile e utilizzata
 - ri-finanziamenti PRIN
- Stabilire obiettivi chiari per evoluzione progressiva
 - richiesta di “innovazione”: “it is how it is” (cit.)
 - niente effetti speciali, ma un’espansione di funzionalità etc.
- Oltre alle collaborazioni
 - edizioni critiche ALIM
 - edizioni diplomatiche / integrate (*Leges Langobardorum*)
 - progetto Arca di Noè come metodo supplementare di pubblicazione
 - rendere ALIM una risorsa LOD (*Orso Beneventano*)

Futuri possibili: oltre l'Arca di Noè

- Progetto “Good things come in small packages” (AIUCD 2020)
 - Un'evoluzione del progetto Visionary Cross
 - The *Visionary Cross* project on Zenodo:
https://zenodo.org/communities/the_visionary_cross/
 - O'Donnell et al. (2019) ‘Publishing (and Forgetting) the Small or Medium-sized Scholarly Edition or Cultural Heritage Collection as Linked Open Data’ <https://zenodo.org/record/3586007>
 - Idea di base: mettere a disposizione di altri studiosi tutti i dati dell'edizione (uso diretto senza intermediazione)
 - Non un approccio LOD, ma rientra nell'idea di edizione distribuita
 - Richiede collaborazione con il CERN e sperimentazione (EVT)

Futuri possibili: oltre l'Arca di Noè



Futuri possibili: ALIM 2+

- Un'evoluzione di ALIM 2 con duplice obiettivo
 - semplificare la manutenzione nel lungo periodo
 - migliorare la funzionalità esistente
- Step 1: scindere sito web dal front-end della DL
 - usare una versione WP senza plugin specifici ALIM (solo link diretti)
 - affidare la manutenzione e l'aggiornamento ai sistemisti di Siena
 - col tempo affinare l'aspetto
- Step 2: migrare su nuova versione Muruca
 - necessarie info da Net7 (in particolare per capire vantaggi)
- Step 3 (optional): migrare su nuovo motore di ricerca
 - parte delle limitazioni attuali (es. la gestione degli allografi nella ricerca) sono dovute alla versione di Solr utilizzata

Futuri possibili: progettare un futuro ALIM 3

- Ricerca sullo stato dell'arte
- Stabilire le priorità
 - stabilità > flessibilità: software unico
 - flessibilità > stabilità: mix and match
- Definire precisamente gli obiettivi
 - struttura della DL (front-end e back-end)
 - requisiti minimi (es. carico di lavoro, n. contatti, UI)
 - funzionalità desiderate (es. collezioni tematiche, allografi, filtri, ricerche)
 - piano di gestione e manutenzione nel lungo periodo
 - piano di espansione nel medio-lungo periodo
- Allocare le risorse necessarie
 - non solo monetarie, anche umane (gruppo di lavoro dedicato)