Approaching a Digital Scholarly Edition through Metadata
Zacharias Topelius Skrifter

This poster is an overview of the database structure in the digital critical edition of Zacharias Topelius Skrifter. It specifies what kind of information is or will be available for the user, either directly through indexes or other related display models, or through a future API. It shows how the entity relations open a possibility for the user to approach the edition, by taking advantage of the information and descriptive metadata in indexing systems. Presented here are both already available features and features in progress, everything is open access.

Manuscript descriptions
Manuscript descriptions and descriptive metadata, related to each published letter in a correspondence, or to a facsimile of a letter.
Possibilities to show links to the people and the geographic index.

People index
Index based on annotated persons in published xml encoded texts.
Possibilities to show authors or writers linked from the bibliographic and letter database. Possibilities to show groups and families.

Timeline
Timeline linking to records in a corresponding database through dates. Written letters and manu-scripts, published books and articles, family members being born or died.
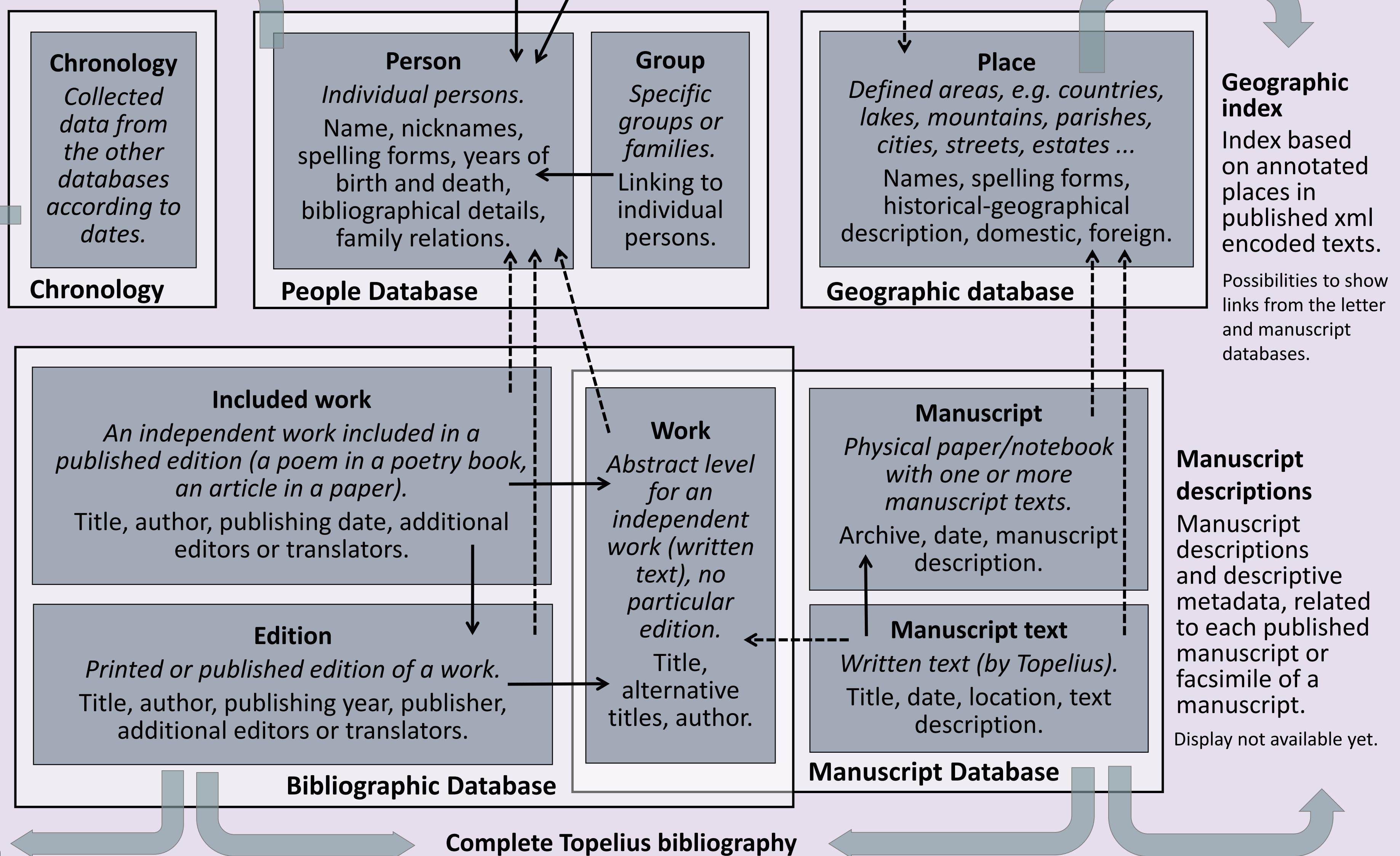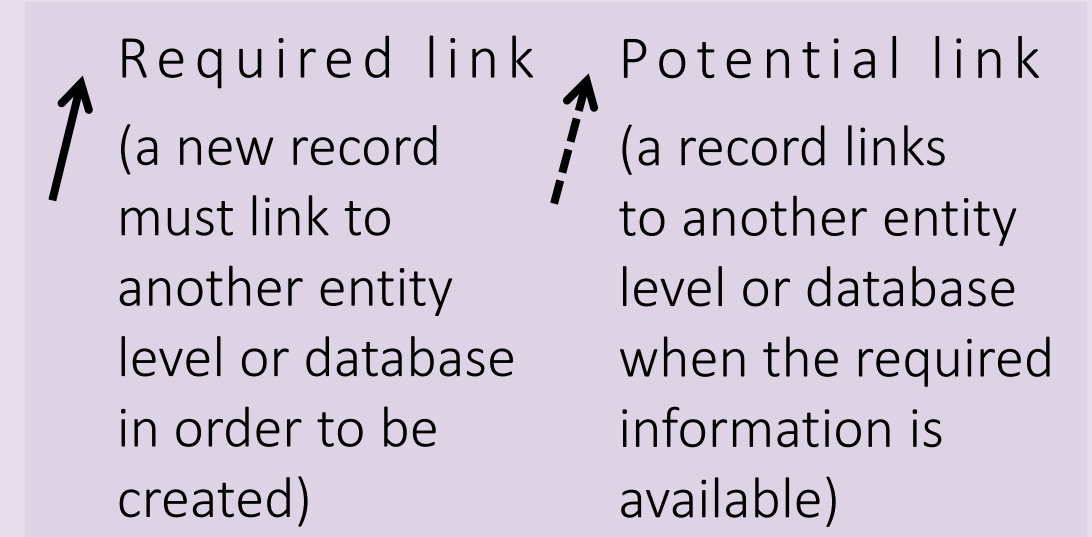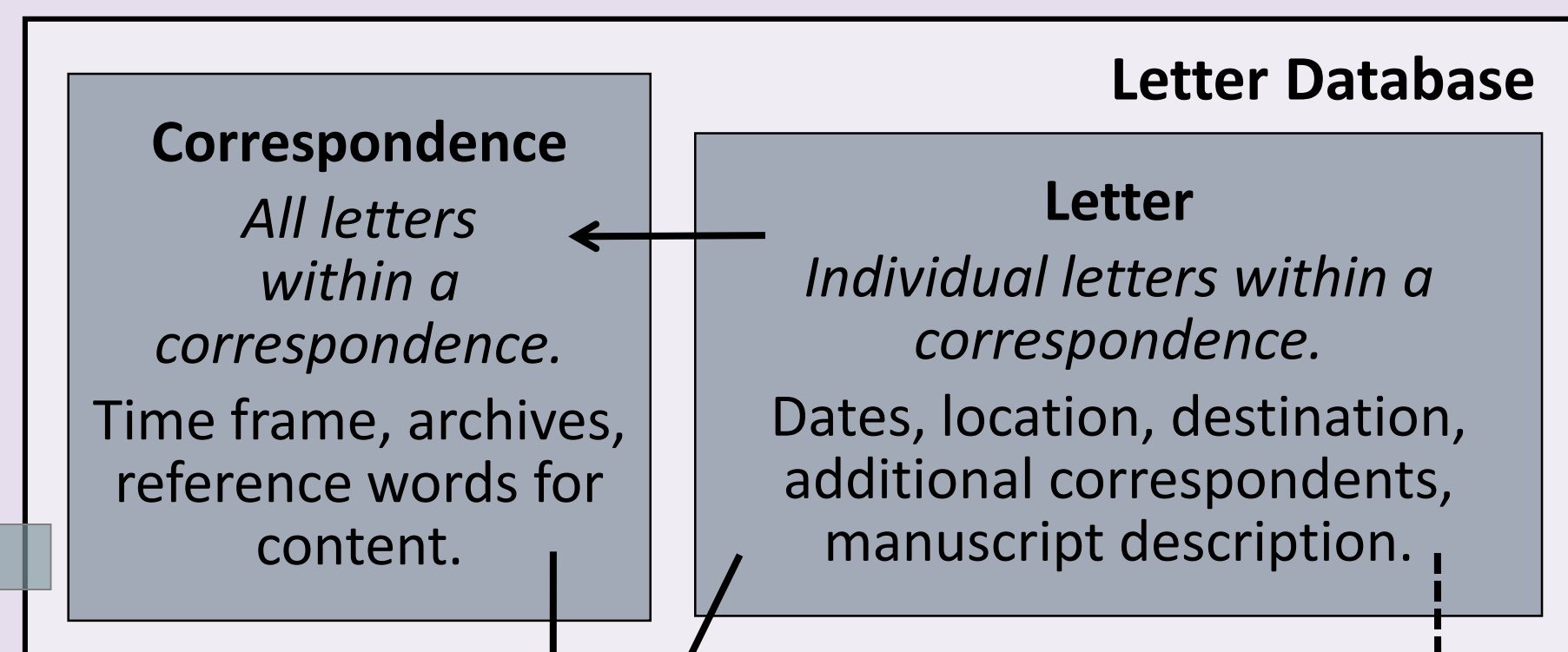Display not available yet.

Bibliographical index
Index based on annotated mentions of books or texts in published xml encoded texts. Authors, editors and translators link to the people index. The display takes both the abstract works and the defined editions in consideration.
Display available in 2019.

By 2022 *Zacharias Topelius Skrifter* will cover almost a dozen different genres in 21 volumes; altogether nearly 20 000 pages of text by Topelius.

Letter Database

Correspondence
All letters within a correspondence.
Time frame, archives, reference words for content.

Letter
Individual letters within a correspondence.
Dates, location, destination, additional correspondents, manuscript description.

Required link (a new record must link to another entity level or database in order to be created)
Potential link (a record links to another entity level or database when the required information is available)

Chronology
Collected data from the other databases according to dates.
Chronology

Person
Individual persons.
Name, nicknames, spelling forms, years of birth and death, bibliographical details, family relations.

Group
Specific groups or families.
Linking to individual persons.

People Database

Place
Defined areas, e.g. countries, lakes, mountains, parishes, cities, streets, estates ...
Names, spelling forms, historical-geographical description, domestic, foreign.
Geographic database

Geographic index
Index based on annotated places in published xml encoded texts.
Possibilities to show links from the letter and manuscript databases.

Included work
An independent work included in a published edition (a poem in a poetry book, an article in a paper).
Title, author, publishing date, additional editors or translators.

Work
Abstract level for an independent work (written text), no particular edition.
Title, alternative titles, author.

Manuscript
Physical paper/notebook with one or more manuscript texts.
Archive, date, manuscript description.

Edition
Printed or published edition of a work.
Title, author, publishing year, publisher, additional editors or translators.
Bibliographic Database

Manuscript text
Written text (by Topelius).
Title, date, location, text description.
Manuscript Database

Manuscript descriptions
Manuscript descriptions and descriptive metadata, related to each published manuscript or facsimile of a manuscript.
Display not available yet.

Complete Topelius bibliography
A complete bibliography over everything Topelius wrote, from every known manuscript to printed editions and translations. Possibilities for different display views.
Display available in 2019.

The **bibliographic database** is the most complex database. It follows the *Functional Requirements for Bibliographic Records* (FRBR) model, making a difference between the abstract work and published editions. The abstract level is significant when it is impossible to determine to which edition Topelius is referring. FRBR focuses on the content relationship and continuum between levels; anything regarded a separate work starts as a new abstract record, from where its own editions start.
Some texts can have their own independent editions as well as being included in editions (e.g. magazines, anthologies) with other texts. This requires two different manifestation levels subordinated the abstract level. Records of texts included in other editions must link to a record of regular editions. The **manuscript database** follows the same structure, sharing the same abstract level.

A high level of detail in database records can make searches in the edition more effective.
Topelius uses various spellings, name forms and abbreviations, and even an advanced search in the texts would seldom give a comprehensive result. Different spellings and name forms are registered and the information included in the index search function. The letter database also includes reference words describing the contents, beyond the wordings of the original texts.