

# METADATA GENERATION FOR CULTURAL HERITAGE

**Creative Histories – The Josefsplatz Experience**

**Brigitte Krenn, Gregor Sieber, Hans Petschar**

{brigitte.krenn, gregor.sieber}@ofai.at

hans.petschar@onb.ac.at



Österreichische  
Nationalbibliothek

# Talk Outline

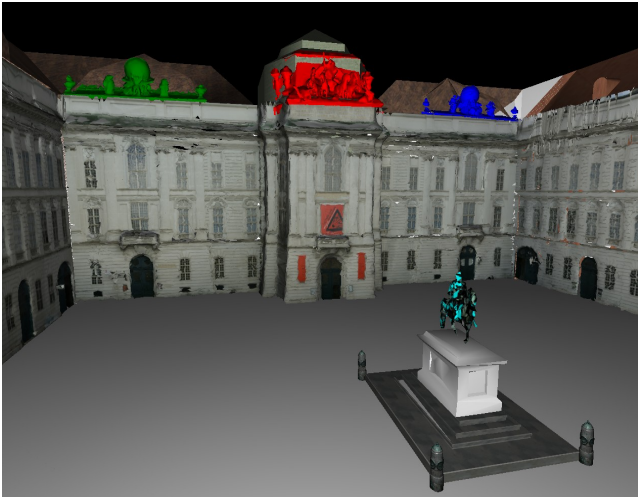
- introduction
- project overview
- metadata extraction
- metadata processing
- speech synthesis
- automatic tour generation
- summary

# Creative Histories



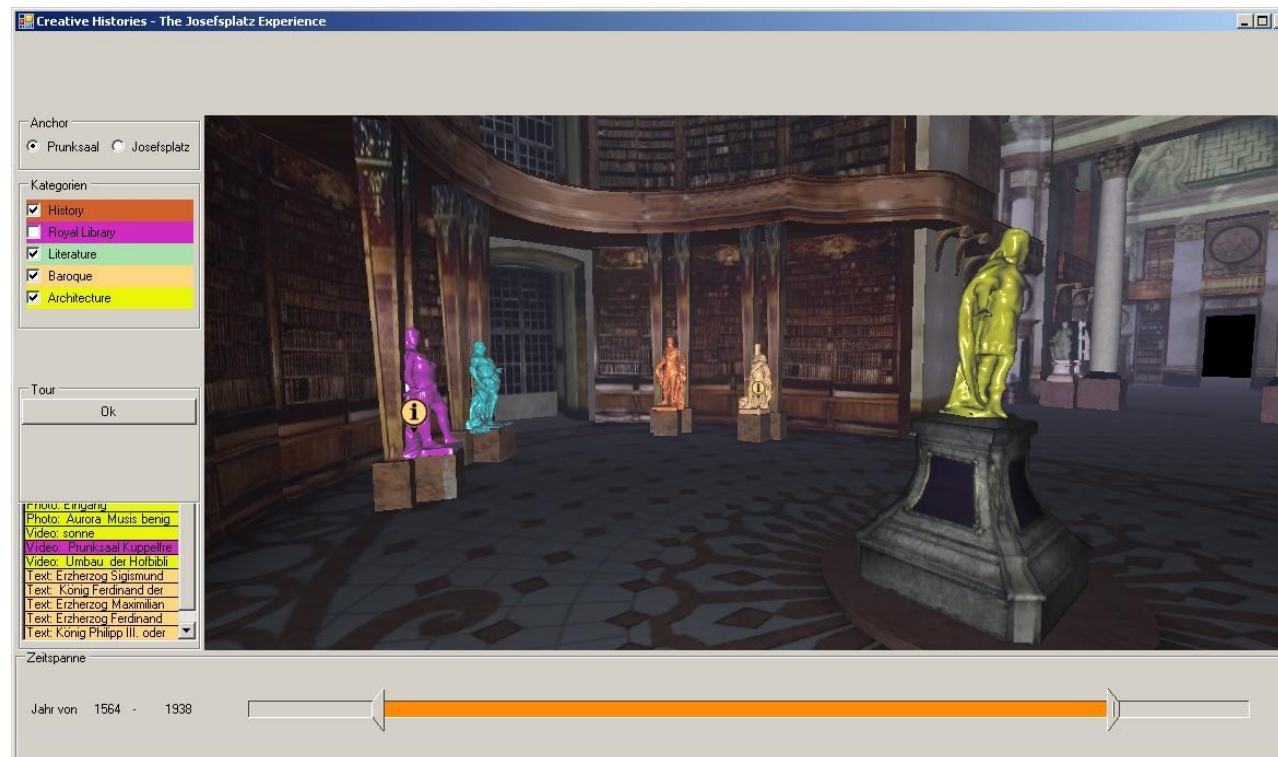
# Creative Histories

3d model of the Josefsplatz & Prunksaal  
reconstructed from (historic) photographs,  
paintings, maps



# Creative Histories

navigation of 3d space and embedded meta-information across different epochs on PC and mobile phones



# Conceptual Model

depending on

- spatial and temporal position in the model
- user interests (categories, level of detail)
- already accessed information (history)
- device used (mobile phone, desktop)

the 3d model will change and

different information will be accessible to the user

# Metadata Component

## **Main Tasks:**

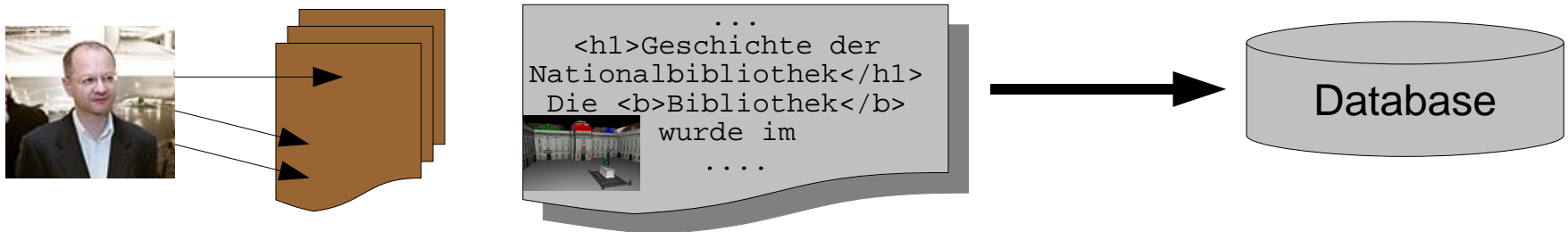
extraction of metadata from legacy documents

semantic representation & querying of  
metadata

automatic generation of presentations  
adapted to context, content and device

# Metadata Extraction

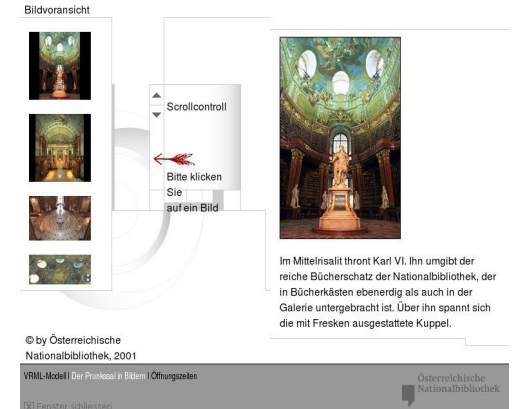
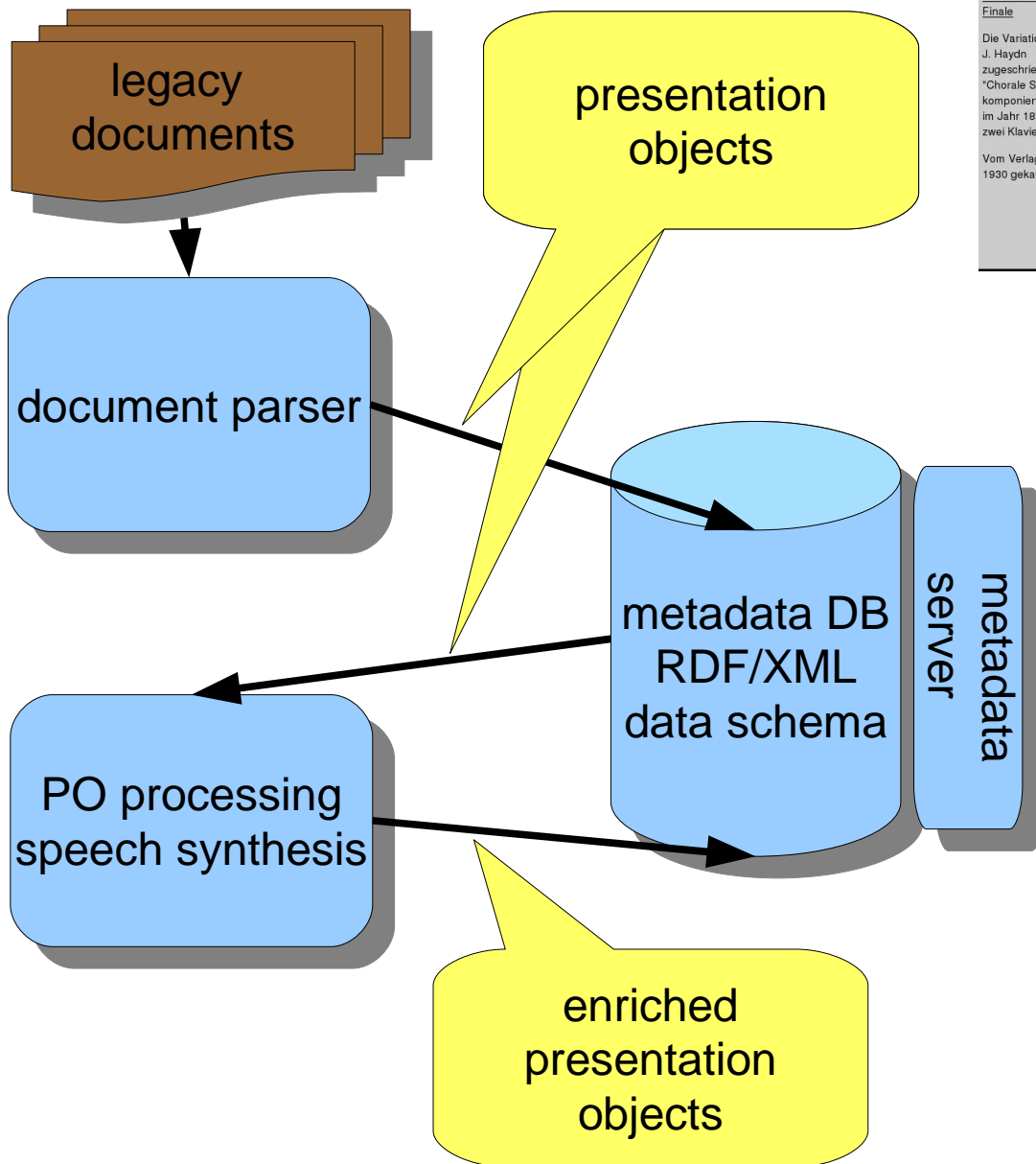
- selection of relevant documents by a human expert
- documents courtesy of the ONB include html documentations, encyclopaedias, exhibition materials
- automatic extraction of metadata using document structure and mark-up (citations)





# Metadata Processing

- interlinked 'presentation objects' (PO) of types *image, text, video, audio, vrml model* are stored in a RDF/XML framework
- PO are enriched with meta-information such as temporal and spatial anchors, category, file size, image dimensions, duration
- textual data are stored sentencewise



*3<sup>rd</sup> party tools used:*

Database & Modeling:

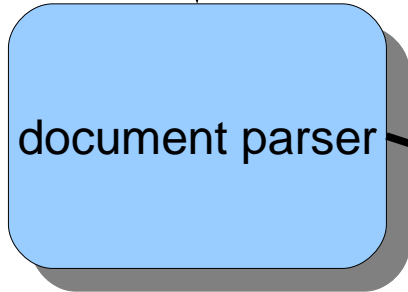
- Sesame RDF Database
- Apache Tomcat
- Protege Ontology editor

Text-to-Speech Synthesis:

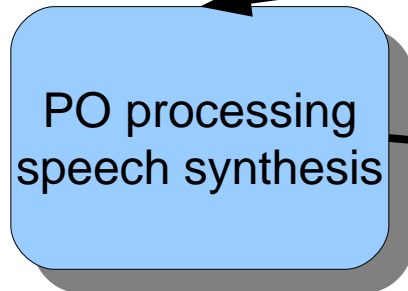
- MARY (diphone)
- Loquendo (unit selection)



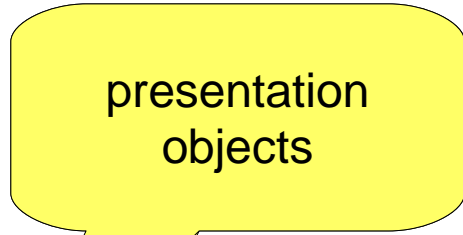
legacy documents



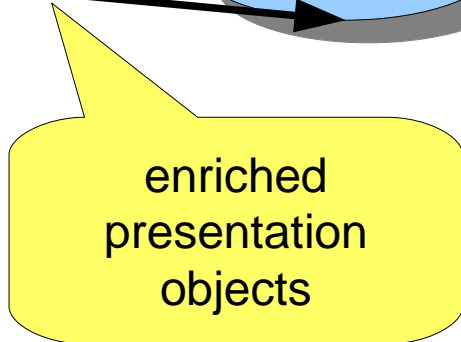
document parser



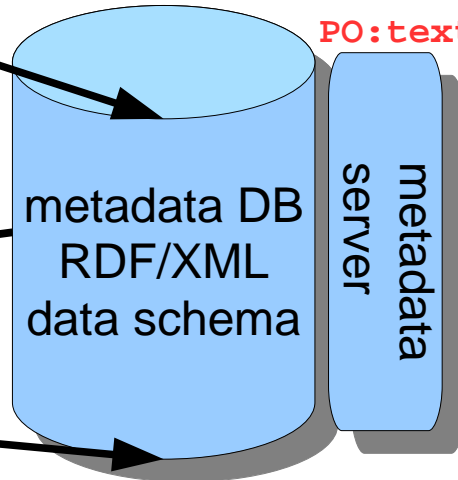
PO processing  
speech synthesis



presentation objects



enriched presentation objects



metadata DB  
RDF/XML  
data schema

PO:text

metadata server



Komponisten  
Johannes Brahms 1833-1897

Bildansicht: Komponist

Haydn-Variationen für  
Orchester op.56a  
Finale  
Die Variationen über den  
J. Haydn  
zugeschrieben  
"Chorale St. Antonii"  
komponierte J. Brahms  
im Jahr 1873 zuerst für  
zwei Klaviere.  
Vom Verlag Simrock  
1930 gekauf.

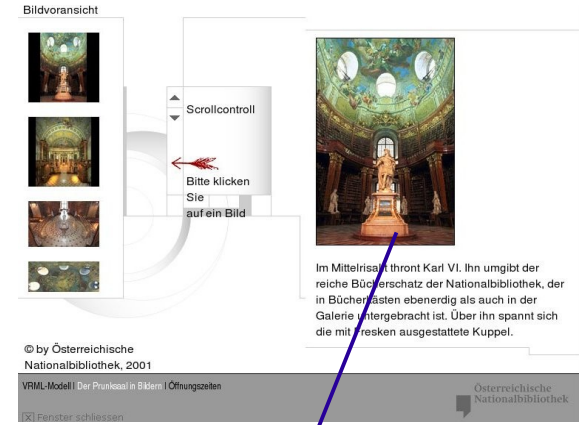


Autograph

PO:audio

links

PO:image

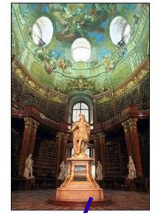


Bildvoransicht



Scrollcontroll

Bitte klicken  
Sie  
auf ein Bild



Im Mittelrisalit thront Karl VI. Ihn umgibt der  
reiche Bücherschatz der Nationalbibliothek, der  
in Bücherkästen ebenerdig als auch in der  
Galerie untergebracht ist. Über ihn spannt sich  
die mit Fresken ausgestattete Kuppel.

© by Österreichische  
Nationalbibliothek, 2001

VRML-Modell | Der Prozess | In Bildern | Öffnungszeiten

Österreichische  
Nationalbibliothek

[X] Fenster schließen

PO:image

legacy documents

document parser

PO processing  
speech synthesis

presentation objects

metadata DB  
RDF/XML  
data schema

enriched presentation objects

**Komponisten**  
Johannes Brahms 1833-1897

◀ Zurück ▶ Blättern

Bildansicht: Komponist

Haydn-Variationen für Orchester op.56a  
Finale


Die Variationen über den J. Haydn zugeschriebenen "Chorale St. Antonii" komponierte J. Brahms im Jahr 1873 zuerst für zwei Klaviere.

Vom Verlag Simrock 1930 gekauft.



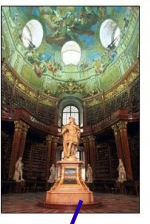
Autograph

Bildvoransicht



Scrollcontroll

Bitte klicken Sie auf ein Bild



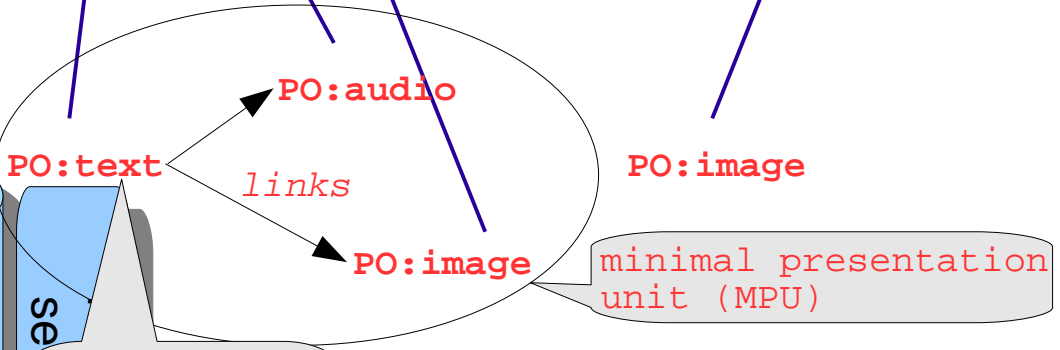
Im Mittelraum thront Karl VI. Ihn umgibt der reiche Bücherschatz der Nationalbibliothek, der in Bücherkästen ebenerdig als auch in der Galerie untergebracht ist. Über ihn spannt sich die mit Fresken ausgestattete Kuppel.

© by Österreichische Nationalbibliothek, 2001

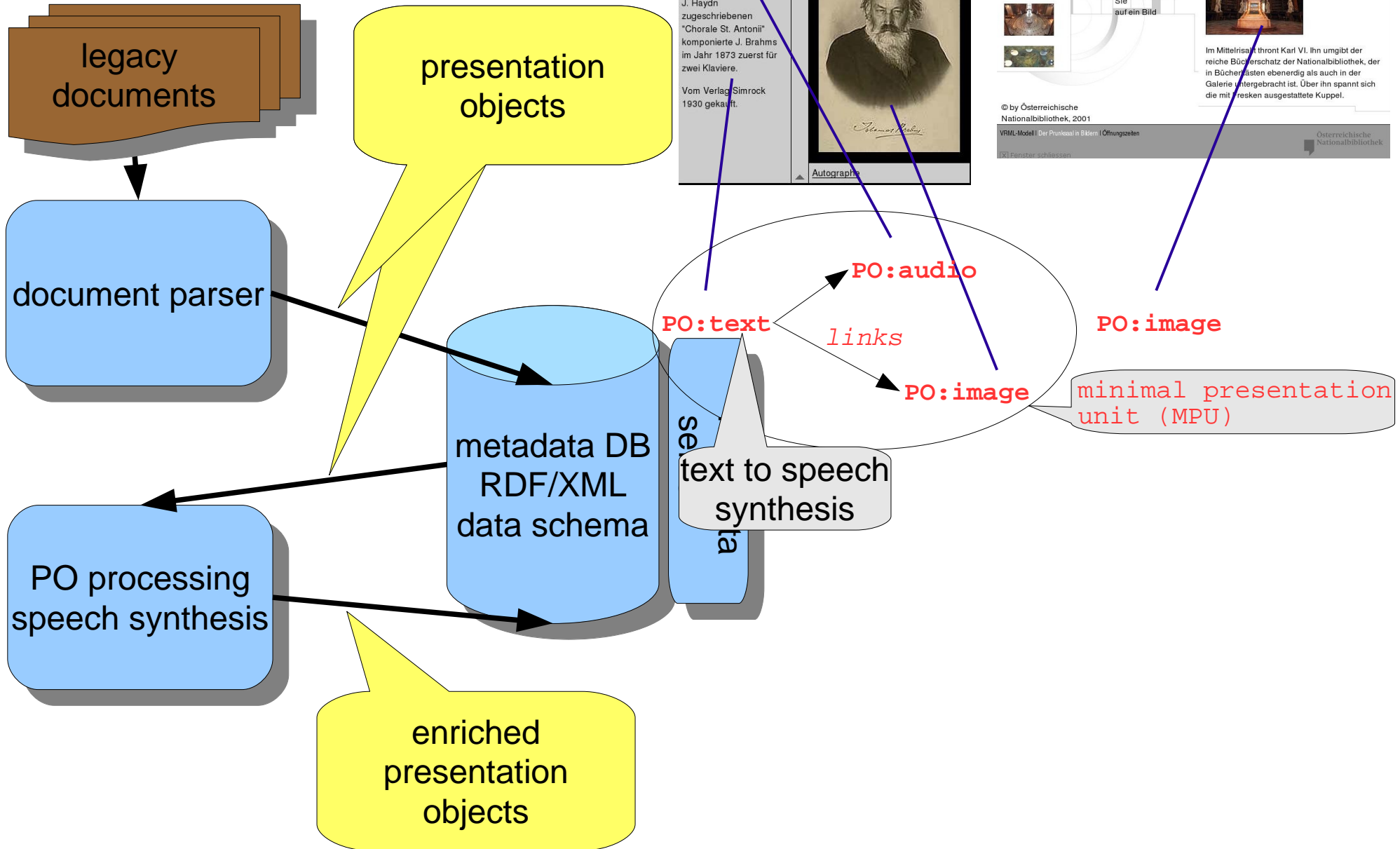
VRML-Modell | Der Prozess | In Bildern | Öffnungszeiten

Österreichische Nationalbibliothek

[X] Fenster schließen



se  
text to speech  
synthesis



# Speech Synthesis

## motivation:

restrictions on display of text on mobile devices,  
flexibility of presentation

## tools:

- automatic preprocessing
- computer-aided compilation of domain lexicon
- Loquendo (unit selection)
- MARY (diphone)

**Loquendo**  
VOCAL TECHNOLOGY AND SERVICES

**MARY** **T**ext **T**o **S**peech

# Speech Synthesis

unit selection quality low for domain vocabulary  
diphone synthesis better adaptable

... den Schatz Friedrichs III. nach ...



... den Schatz Friedrichs *des dritten* nach ...

The screenshot shows a software interface for speech synthesis. At the top, a text box contains the sentence: "Neukonstituierung der Hofmusikkapelle <say-as interpret-as='date' format='y'>1498</say-as>". Below this is a control panel with buttons for "P.Seen", "Prev", "Play", "Next", "N.Unseen", "Stop", and "Goto". Further down are input fields for "Abbrev:", "Text:", "SAMPR:", and "Method:", each with a corresponding "Insert" or "Play" button. Below these are buttons for "Add 2 lexicon", "Add 2 phrases", "Add 2 abbrev", "Clear 'seen words'", and "Clear 'seen sents'". At the bottom, there is an "Input file" section with "File set:" and "Single f:" fields, each with a "Browse" button. Finally, there are "Save config", "HELP", "verbose", and "Exit" buttons at the very bottom.

Pallavicinipalais, Redoutensäle,  
Mittelrisalit,...

# Tour Generation

online generation of interactive tours  
adapted to

- device requirements,
- context (user interest, position, history),
- content (media types, categories)

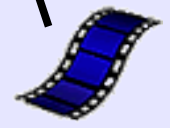
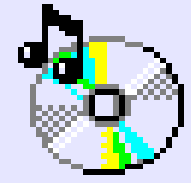
minimal presentation unit

minimal presentation unit

composed of:  
(representations of)  
presentation objects,  
internal connectors,  
external connectors;  
player control commands  
(position, orientation)

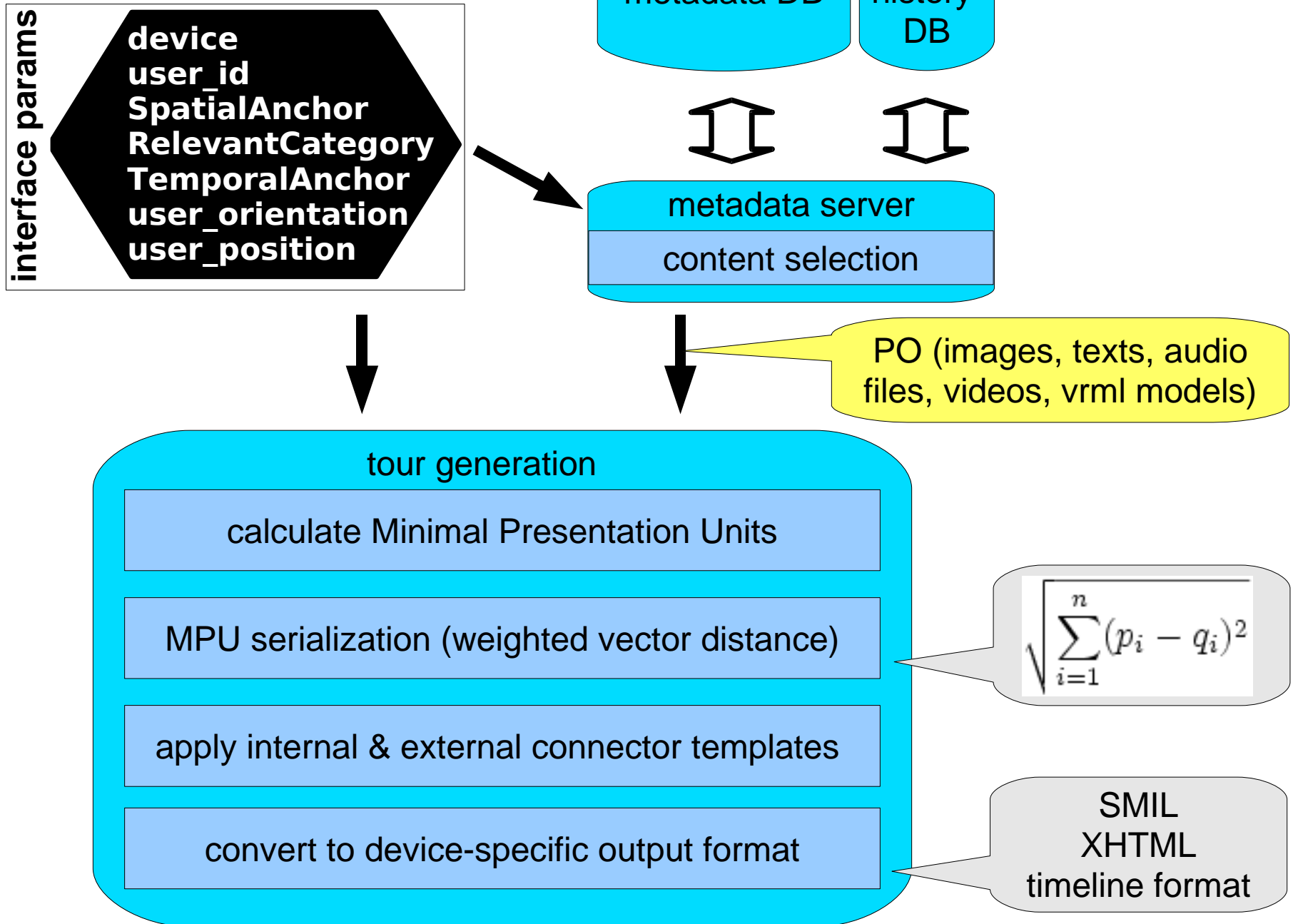
minimal presentation unit

user interaction:  
navigate forwards/backwards  
view/skip optional material  
return to 3d model view



**TOUR**





# Tour Generation: Alignment

images are temporally aligned with speech

- temporal position of pictures determined by link position and clause structure
- extraction of timing information using MARY



<break>

<break>

<break>

<break>

... zeigen das Haus Österreich, die Steiermark, Tirol, und Kärnten, die Länder in denen ...



# Summary

## **major characteristics revisited:**

automatic extraction and processing of presentation objects from legacy documents

storage of presentation objects along media type and interface dimensions (such as category, temporal anchor) in RDF/XML format

online generation of interactive tours adapted to context, content and device

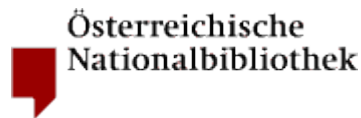
# Project Partners



**Creative Histories – The Josefsplatz Experience** [www.josefsplatz.info](http://www.josefsplatz.info)



**Austrian Research Institute for Artificial Intelligence (OFAI)**  
[www.ofai.at](http://www.ofai.at)



**Austrian National Library (ÖNB)**  
[www.onb.ac.at](http://www.onb.ac.at)



**Zentrum für Virtual Reality und Visualisierung (vrvis)** [www.vrvis.at](http://www.vrvis.at)



**Forschungszentrum Telekommunikation Wien (ftw)**  
[www.ftw.at](http://www.ftw.at)