



MetaMap Transfer (MMTx)

An Overview

Presented December 8, 2005



MMTx Outline

- What is it?
- What is it good for?
- Who uses it?
- What does it look like?
- Where can I get it?
- Requirements, Terms and Conditions
- Customization, Common Options
- Team Members and Points of Contact



MMTx

What is it?

- Tool for discovering controlled medical vocabulary terms (via UMLS Concepts) within documents
- Tokenizes documents into sentences, phrases, terms, words
- Matches phrases to closest UMLS Concepts
- Java Implementation of MetaMap
- It is Middleware. There is no GUI



MMTx

What is it Good for?

- Information Extraction tasks
- Classification/Categorization tasks
- Text Summarization/Question & Answer tasks
- Data-mining tasks
- Knowledge discovery tasks
- Text Understanding tasks
- UMLS Concept based indexing and retrieval
- NLP Tasks



MMTx

Who uses MMTx?

- Medical Problem Extraction from Clinical Documents
- Extraction of Concepts from Chief Complaints
- Extraction of Concepts from Clinical Narratives
- Extracting Diagnoses from Discharge Summaries
- Concept-Value Pair Extraction from Semi-Structured Echocardiogram Reports
- Enzyme class Annotation
- Term Identification in the Biomedical Literature
- Extracting Molecular Binding Relationships from Biomedical Text



MMTx

What does it do?

Extracts UMLS
concepts from text

Meta Mapping (1000):

C0496836

(Malignant neoplasm of eye, unspecified)

[Neoplastic Process]

Retinoblastoma

What is **retinoblastoma**?

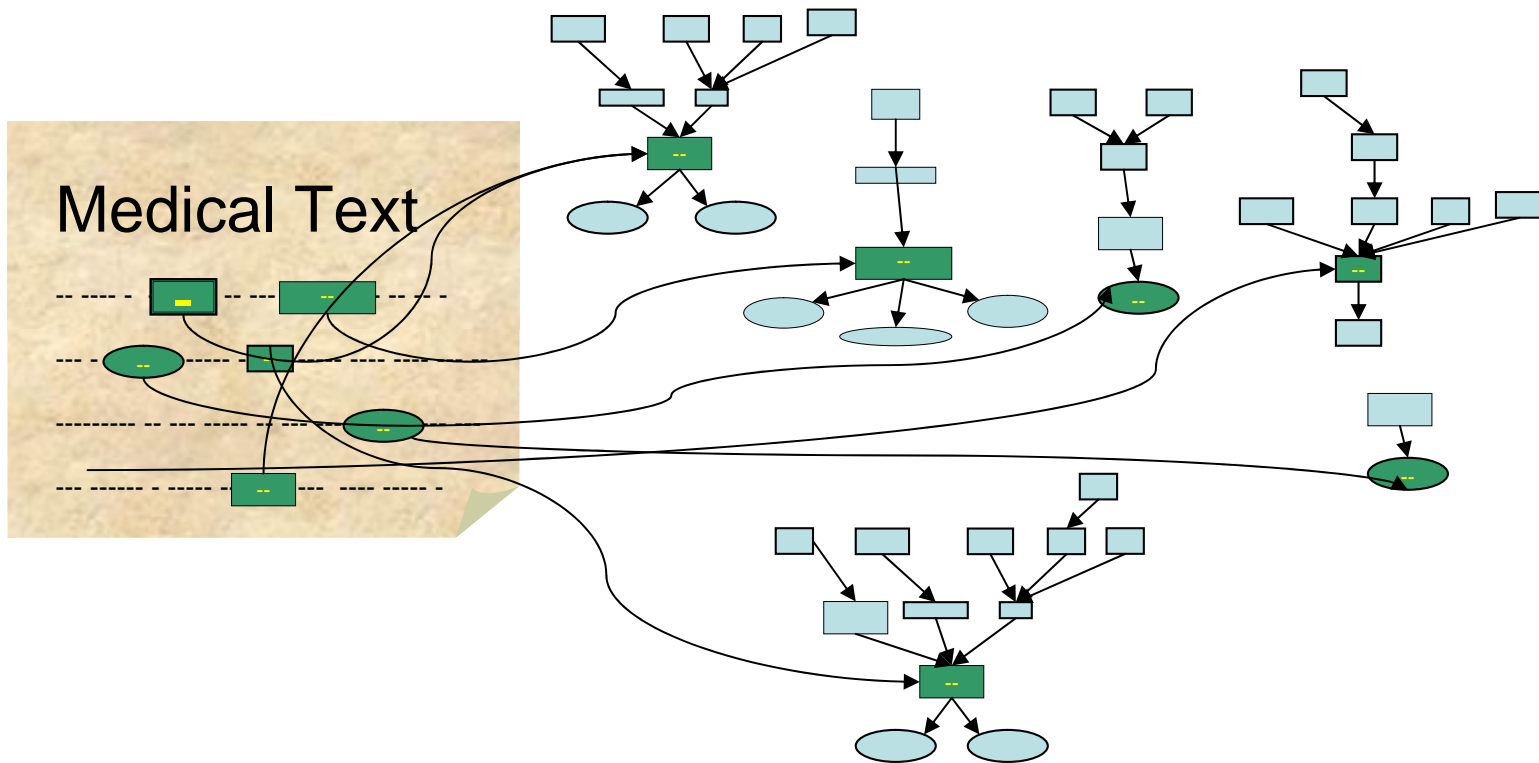
Retinoblastoma is a rare type of **eye cancer** that develops in the retina, which is the part of the eye that detects light and color.

Although this disorder can occur at any age, it usually develops in young children.



MMTx

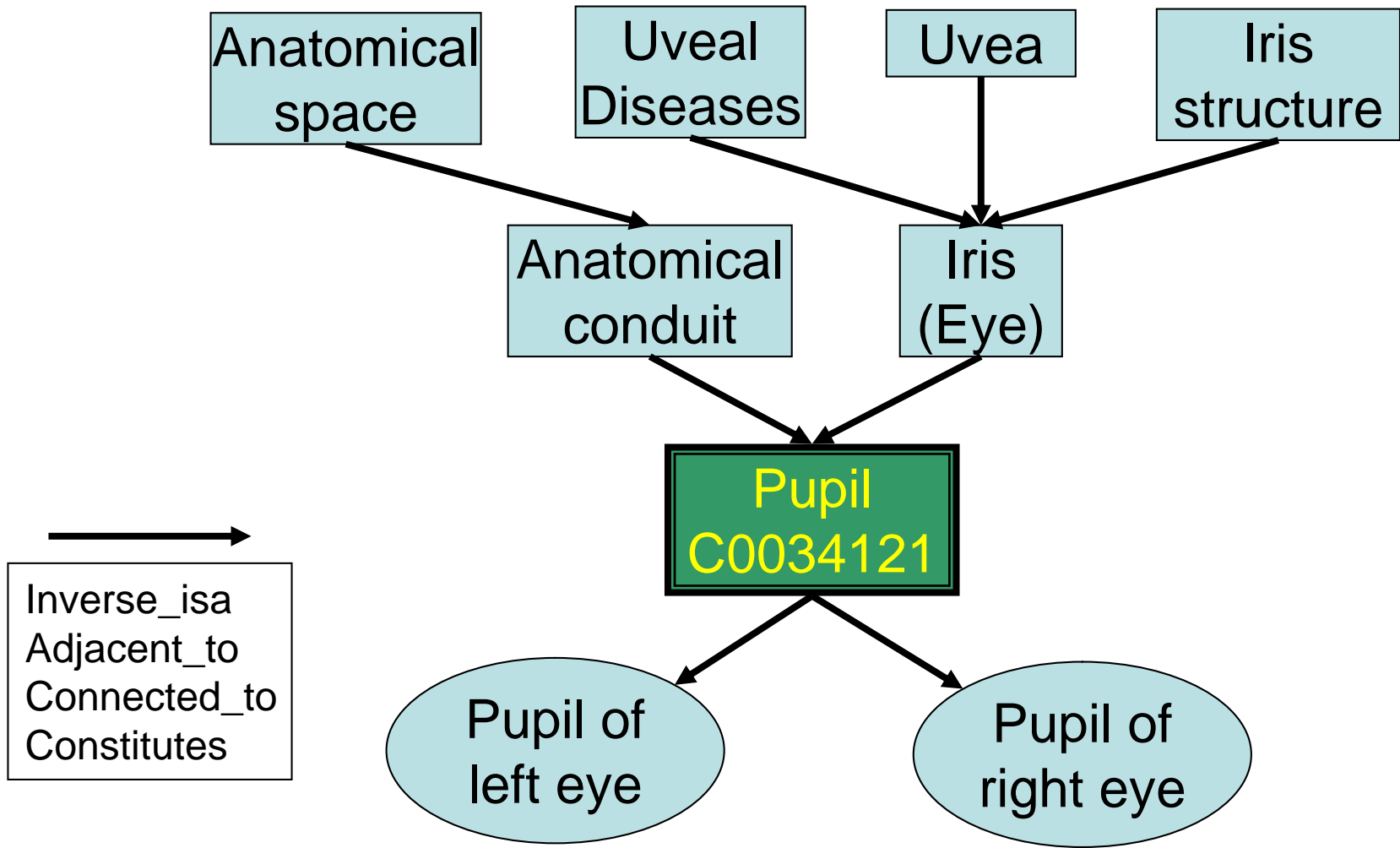
Why would you want to use it?





MMTx

Why would you want to use it? (2)





MMTx

What does it look like?

> **mmtx** --fileName=retnoblastoma.txt

Phrase: "*of eye cancer*"

Meta Candidates (7)

- 1000 Eye Cancer (Malignant neoplasm of eye, unspecified) [Neoplastic Process]
- 861 Cancer (Malignant Neoplasms) [Neoplastic Process]
- 861 Cancer (Malignant neoplasm, primary (morphologic abnormality)) [Neoplastic Process]
- 861 Cancer (Cancer Genus) [Invertebrate]
- 694 Eye [Body Part, Organ, or Organ Component]
- 694 Eye (Entire eye) [Body Part, Organ, or Organ Component]
- 638 Ophthalmic [Spatial Concept]

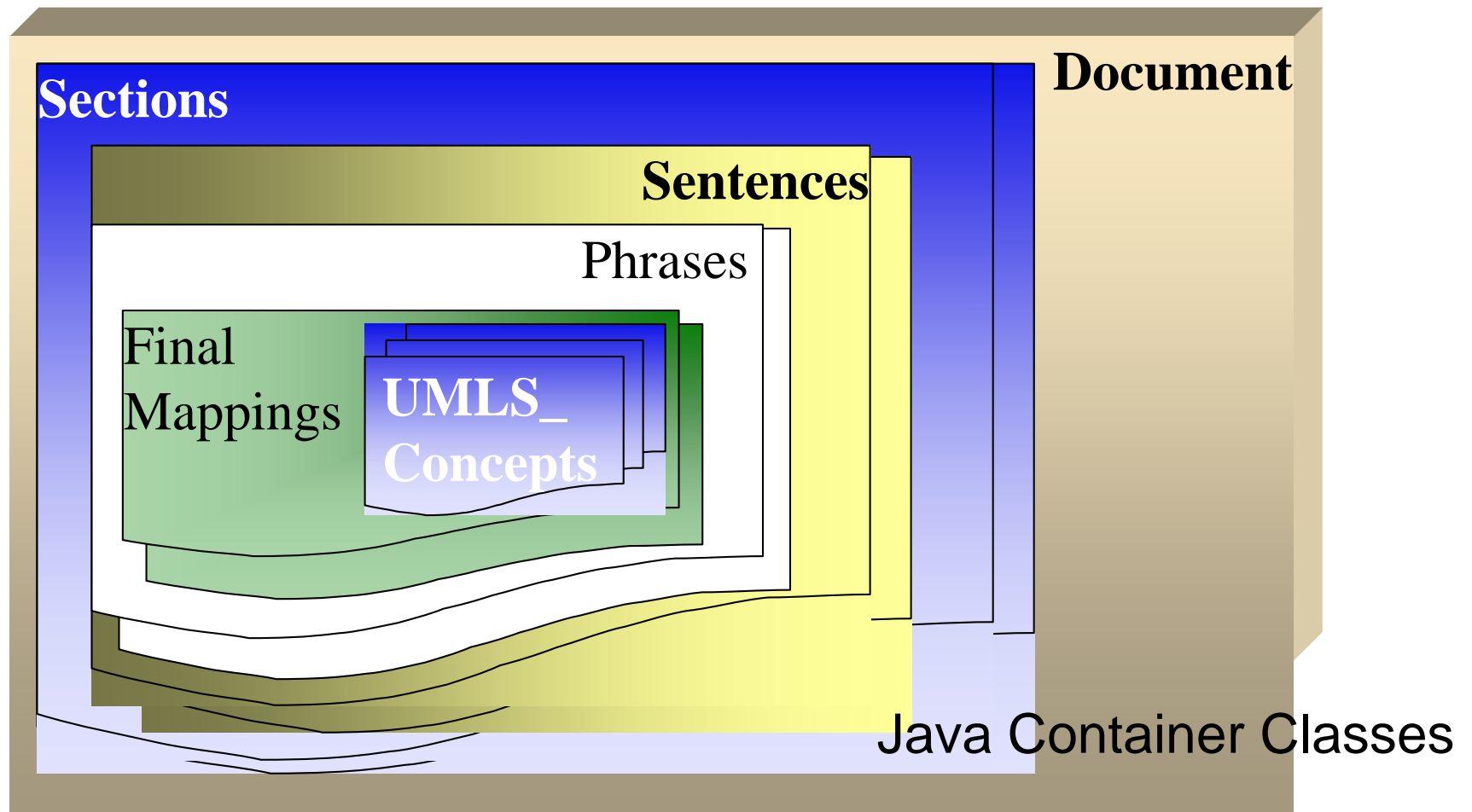
Meta Mapping (1000)

- 1000 Eye Cancer (Malignant neoplasm of eye, unspecified) [Neoplastic Process]



MMTx

What does the API look like?





MMTx

API Methods and Constructors

MMTxAPI

MMTxAPI()

MMTxAPI(String[] args)

Document **processDocument** (File pFile)

Document **processDocument** (String pDocumentText)

void **processSentence** (Sentence pSentence)

Sentence **processSentence** (String pSentenceText)

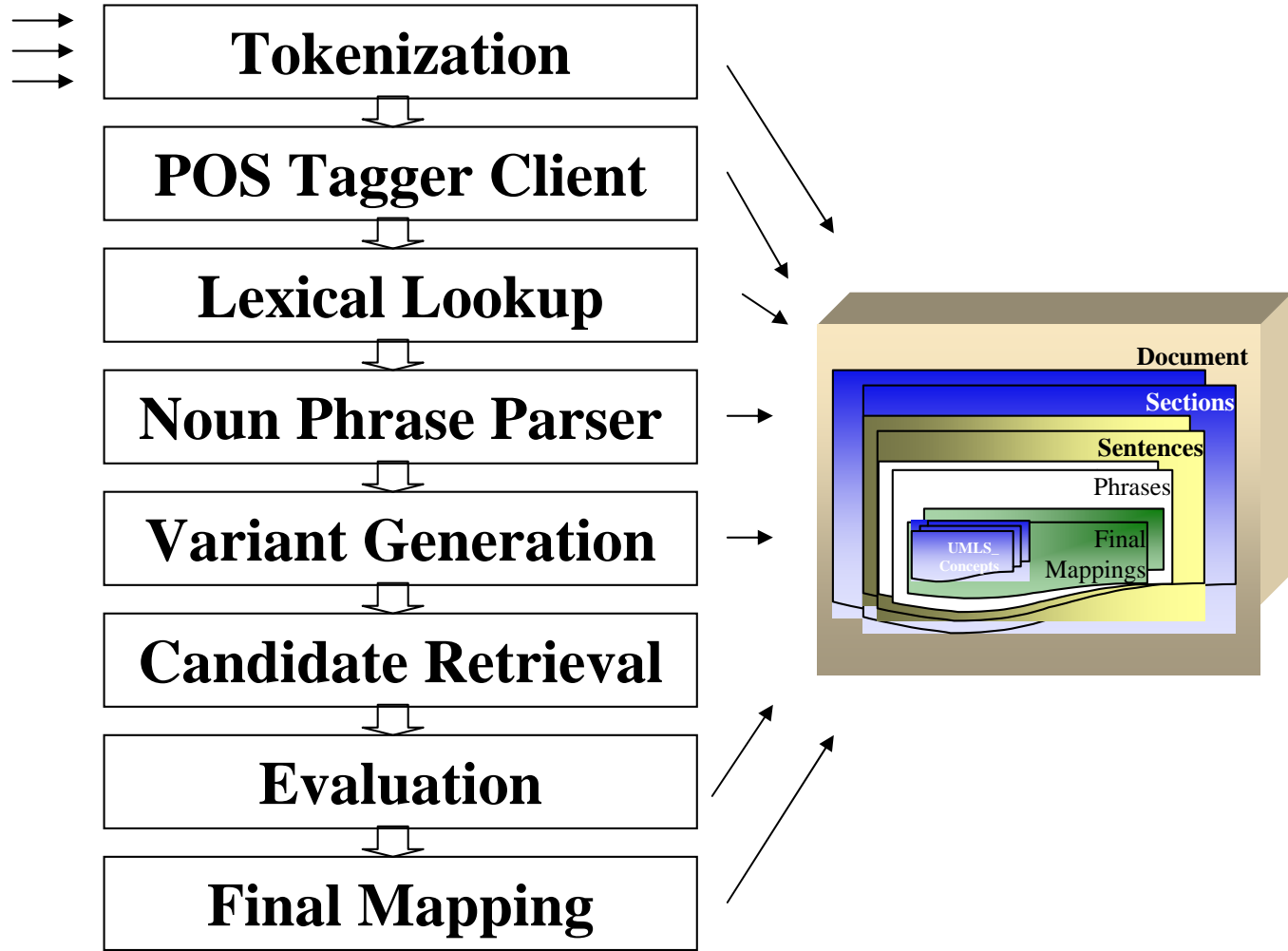
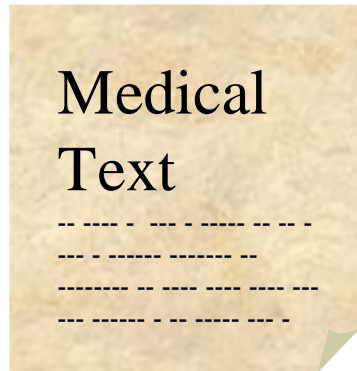
Sentence **processString** (String pString,
boolean pTrmPrsng)

Phrase **processTerm** (String pTerm)



MMTx

How does it Work?





MMTx

Where can I get it?

- <http://mmtx.nlm.nih.gov>
- The download is password restricted
- Preconditions:
 - Signed the UMLS license agreement
 - UMLS Knowledge source server account name and password



MMTx

Minimum Requirements

- OS's supported:
 - Solaris/Windows NT/2000/XP, Linux, Mac OS/X
- Java 1.4 or better
- 400 mb disk space for software
- 3GB disk space for each Year's Data per model
- 600+ MB or more RAM/Swap space



MMTx

Terms and Conditions of Use

- Must be a UMLS Signatory
- Must customize the data to the vocabularies that you have rights to use
- The MMTx software soon to be under an open source agreement:
 - Attribution, redistribution in-total, no NLM endorsement, indemnity



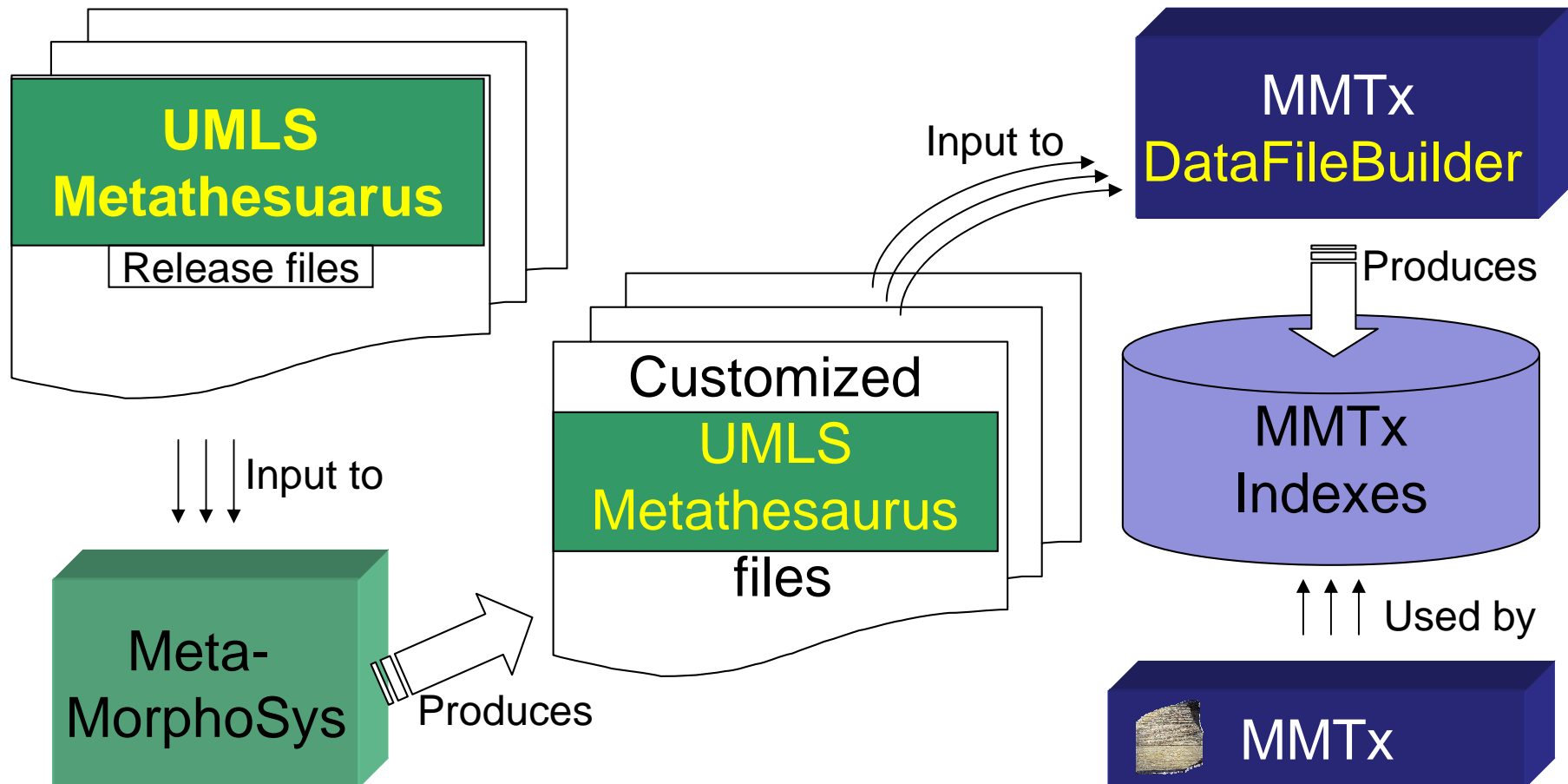
MMTx

Can I customize the data?

- You will need to customize it because ...
 - Need to extract only those vocabularies you have rights to
 - Extract only those vocabularies that make sense for your application

MMTx

How can I customize the data

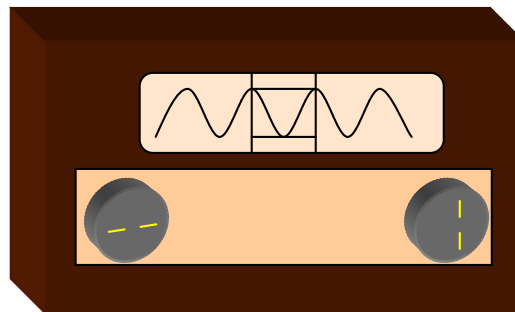




MMTx

How do I tune it?

- Limit matches to conservative variation
 - Noun/adjective derivations, unique acronyms and expansions
- Filter by semantic type(s)
- Filter by vocabulary source(s)
- Match on (longer) composite phrases





MMTx

Team Members
and
Points of Contact

mmtx@nlm.nih.gov

<http://mmtx.nlm.nih.gov>