

# BIBTEX

## Eine Vivisektion

*Gerd Neugebauer*

<http://www.gerd-neugebauer.de>  
[gene@gerd-neugebauer.de](mailto:gene@gerd-neugebauer.de)

Im Lerchelsbühl 5  
64521 Groß-Gerau

### Vivisektion

Eingriff am lebenden Tier zu  
wissenschaftlichen  
Versuchszwecken.

# Inhalt

## Die Literaturdatenbank aux-Datei bib-Dateien bst-Datei BibTeX 1.0 Therapie: Reimplementierung

Die Literaturdatenbank  
Kooperation  
aux-Datei  
bib-Dateien  
bib-Dateien: Beispiel  
bib-Dateien: Befunde  
bst-Datei  
Komponenten  
Strukturierung des  
Interpreters  
Das Hauptprogramm  
Das Hauptprogramm: Beispiel  
Eingebaute Funktionen  
Der Sprach-Kern  
Der Sprach-Kern: Befunde  
Die Datenbank-Funktionen  
Die Typesetting-Funktionen  
Die Text-Funktionen  
Dokumentenlage  
BibTeX 1.0  
Therapie: Reimplementierung  
Und wohin geht die Reise?

# Die Literaturdatenbank

## ➤ Aufgaben

- ◆ Verwaltung von Literaturstellen
- ◆ Anbindung an L<sup>A</sup>T<sub>E</sub>X
- ◆ Selektion der relevanten Einträge
- ◆ Formatierung der Einträge
- ◆ BIB<sub>T</sub>EX weiß nichts über Literatur

## ➤ Realisierung

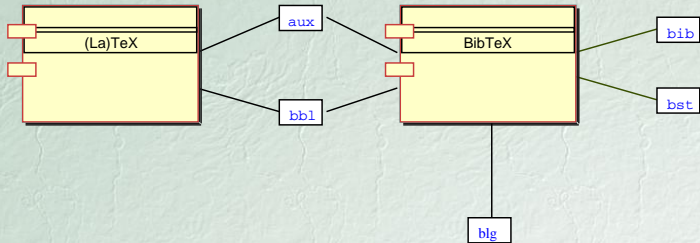
- ◆ WEB

## ➤ Geschichte

- ◆ BIB<sub>T</sub>EX 0.99c  
Oren Patashnik 1985
- ◆ BIB<sub>T</sub>EX 1.0  
Lange angekündigt

Die Literaturdatenbank ✕  
Kooperation  
aux-Datei  
bib-Dateien  
bib-Dateien: Beispiel  
bib-Dateien: Befunde  
bst-Datei  
Komponenten  
Strukturierung des  
Interpreters  
Das Hauptprogramm  
Das Hauptprogramm: Beispiel  
Eingebaute Funktionen  
Der Sprach-Kern  
Der Sprach-Kern: Befunde  
Die Datenbank-Funktionen  
Die Typesetting-Funktionen  
Die Text-Funktionen  
Dokumentenlage  
Bib<sub>T</sub>EX 1.0  
Therapie: Reimplementierung  
Und wohin geht die Reise?

## Kooperation



- Datei-basiert
- Abwechselnde Aktivität  
damit mehrere L<sup>A</sup>T<sub>E</sub>X-Läufe

## aux-Datei

- Wenige Anweisungen sind relevant:

```
\citation{*}  
\bibstyle{alpha}  
\bibdata{xampl}
```

## Befunde

- Die Anweisungen sind fragil:
  - ◆ Keine Variationen im Format erlaubt
- Unproblematisch, da von L<sup>A</sup>T<sub>E</sub>X-Makros erzeugt

# bib-Dateien

- Wenige eingebaute Typen:
  - ◆ @preamble
  - ◆ @string
  - ◆ @comment
- Vererbung von Attributen (crossref)
- Makros (Strings)
- Konkatenierung
- Definierbare Typen und Felder:  
festgelegt in der bst-Datei
- Alles zwischen den Einträgen ist Kommentar

## bib-Dateien: Beispiel

```
@preamble{ "\newcommand{\noopsort}[1]{ } "
           # "\newcommand{\printfirst}[2]{#1} "
           # "\newcommand{\singleletter}[1]{#1} "
           # "\newcommand{\switchargs}[2]{#2#1} "

@ARTICLE{article-minimal,
  author = {L[eslie] A. Aamport},
  title = {The Gnats and Gnus Document Preparation
           System},
  journal = {\mbox{G-Animal's} Journal},
  year = 1986,
}

@ARTICLE{article-full,
  author = {L[eslie] A. Aamport},
  title = {The Gnats and Gnus Document Preparation System},
  journal = {\mbox{G-Animal's} Journal},
  year = 1986,
  volume = 41,
  number = 7,
  pages = "73+",
  month = jul,
  note = "This is a full ARTICLE entry",
}
```

Die Literaturdatenbank ✓  
Kooperation ✓  
aux-Datei ✓  
bib-Dateien ✓  
bib-Dateien: Beispiel x  
bib-Dateien: Befunde  
bst-Datei  
Komponenten  
Strukturierung des  
Interpreters  
Das Hauptprogramm  
Das Hauptprogramm: Beispiel  
Eingebaute Funktionen  
Der Sprach-Kern  
Der Sprach-Kern: Befunde  
Die Datenbank-Funktionen  
Die Typesetting-Funktionen  
Die Text-Funktionen  
Dokumentenlage  
Bib<sub>T</sub>X 1.0  
Therapie: Reimplementierung  
Und wohin geht die Reise?

## bib-Dateien: Befunde

- **@comment** ist kontraintuitiv
- **Kein @** in Kommentaren (E-Mail-Adresse)
- **crossref** nur eine Stufe tief
- **Reihenfolge bei crossref (one-pass)**

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
<b>bib-Dateien: Befunde</b>	<b>x</b>
bst-Datei	
Komponenten	
Strukturierung des	
Interpreters	
Das Hauptprogramm	
Das Hauptprogramm: Beispiel	
Eingebaute Funktionen	
Der Sprach-Kern	
Der Sprach-Kern: Befunde	
Die Datenbank-Funktionen	
Die Typesetting-Funktionen	
Die Text-Funktionen	
Dokumentenlage	
BibTeX 1.0	
Therapie: Reimplementierung	
Und wohin geht die Reise?	



# bst-Datei

- Enthält das Programm
- Konzeptionell einfache Stack-Sprache:
  - ◆ Der Stack als zentrales Mittel zum Datenaustausch
  - ◆ Der Stack kann Daten und Programmteile aufnehmen
  - ◆ Das oberste Element des Stack wird vom Stack genommen und ausgeführt.
- Wenige eingebaute Funktionen

## Befunde

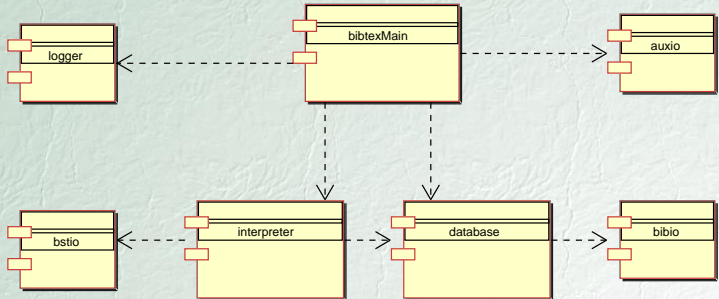
- Keine Modularisierung  
(außer über die aux-Datei)
- Kein Escape-Mechanismus für "

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
<b>bst-Datei</b>	<b>x</b>
Komponenten	
Strukturierung des Interpreters	
Das Hauptprogramm	
Das Hauptprogramm: Beispiel	
Eingebaute Funktionen	
Der Sprach-Kern	
Der Sprach-Kern: Befunde	
Die Datenbank-Funktionen	
Die Typesetting-Funktionen	
Die Text-Funktionen	
Dokumentenlage	
BibTeX 1.0	
Therapie: Reimplementierung	
Und wohin geht die Reise?	

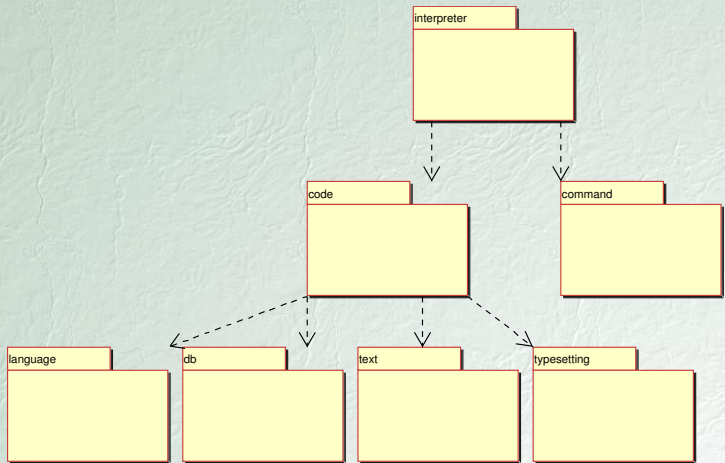
# Komponenten

➤ BIB<sub>T</sub>E<sub>X</sub> lässt sich in „logische“ Komponenten zerlegen

Die Literaturlatenbank ✓  
Kooperation ✓  
aux-Datei ✓  
bib-Dateien ✓  
bib-Dateien: Beispiel ✓  
bib-Dateien: Befunde ✓  
bst-Datei ✓  
**Komponenten** ✗  
Strukturierung des  
Interpreters  
Das Hauptprogramm  
Das Hauptprogramm: Beispiel  
Eingebaute Funktionen  
Der Sprach-Kern  
Der Sprach-Kern: Befunde  
Die Datenbank-Funktionen  
Die Typesetting-Funktionen  
Die Text-Funktionen  
Dokumententlage  
Bib<sub>T</sub>E<sub>X</sub> 1.0  
Therapie: Reimplementierung  
Und wohin geht die Reise?

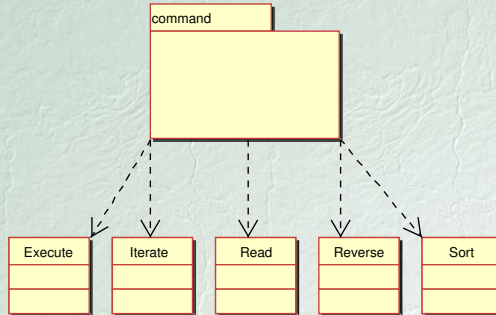


# Strukturierung des Interpreters



- Die Literaturdatenbank ✓
- Kooperation ✓
- aux-Datei ✓
- bib-Dateien ✓
- bib-Dateien: Beispiel ✓
- bib-Dateien: Befunde ✓
- bst-Datei ✓
- Komponenten ✓
- Strukturierung des Interpreters ✗
- Das Hauptprogramm
- Das Hauptprogramm: Beispiel
- Eingebaute Funktionen
- Der Sprach-Kern
- Der Sprach-Kern: Befunde
- Die Datenbank-Funktionen
- Die Typesetting-Funktionen
- Die Text-Funktionen
- Dokumentenlage
- Bib<sub>T</sub>X 1.0
- Therapie: Reimplementierung
- Und wohin geht die Reise?

# Das Hauptprogramm



## ➤ Zusätzlich Deklarationen:

- ◆ ENTRY
- ◆ INTEGERS
- ◆ STRINGS
- ◆ MACRO
- ◆ FUNCTION

Die Literaturlatenbank ✓  
Kooperation ✓  
aux-Datei ✓  
bib-Dateien ✓  
bib-Dateien: Beispiel ✓  
bib-Dateien: Befunde ✓  
bst-Datei ✓  
Komponenten ✓  
Strukturierung des ✓  
Interpreters ✓  
**Das Hauptprogramm** x  
Das Hauptprogramm: Beispiel  
Eingebaute Funktionen  
Der Sprach-Kern  
Der Sprach-Kern: Befunde  
Die Datenbank-Funktionen  
Die Typesetting-Funktionen  
Die Text-Funktionen  
Dokumentenlage  
BibTeX 1.0  
Therapie: Reimplementierung  
Und wohin geht die Reise?

# Das Hauptprogramm: Beispiel

**READ**

**EXECUTE** {initialize.et.al.char.used}

**ITERATE** {presort}

**SORT**

**EXECUTE** {initialize.longest.label}

**ITERATE** {forward.pass}

**REVERSE** {reverse.pass}

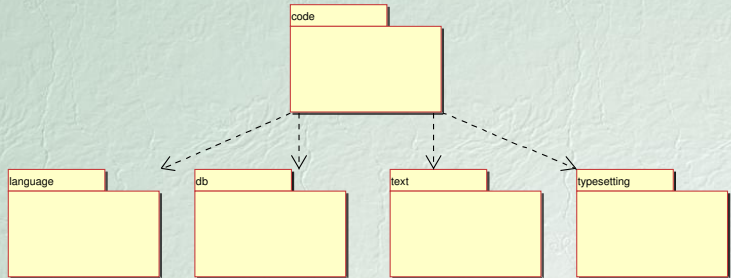
**EXECUTE** {begin.bib}

**EXECUTE** {init.state.consts}

**ITERATE** {call.type\$}

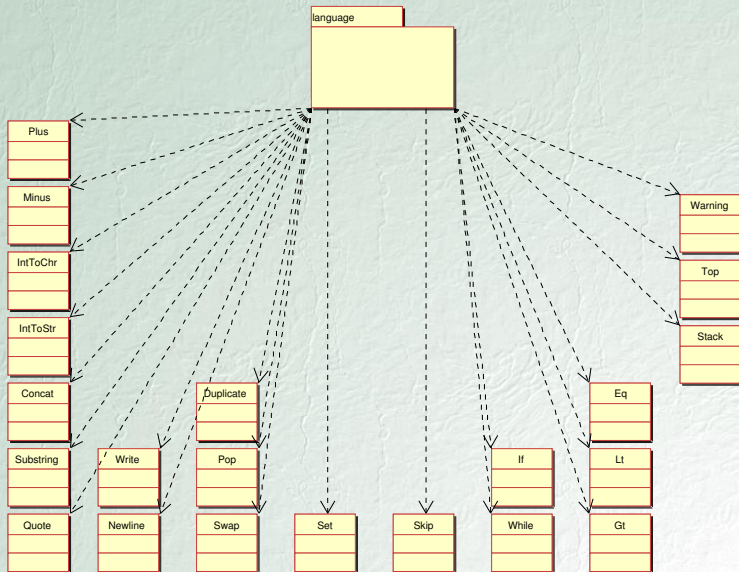
Die Literaturlatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des	
Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	
<b>x</b>	
Eingebaute Funktionen	
Der Sprach-Kern	
Der Sprach-Kern: Befunde	
Die Datenbank-Funktionen	
Die Typesetting-Funktionen	
Die Text-Funktionen	
Dokumentation	
BibTeX 1.0	
Therapie: Reimplementierung	
Und wohin geht die Reise?	

## Eingebaute Funktionen



- **language: Der Sprach-Kern**
- **db: Datenbank-Funktionen**
- **text: Text-Funktionen**
- **typesetting: Typesetting-Funktionen**

# Der Sprach-Kern



- Die Literaturlatenbank Kooperation ✓
- aux-Datei ✓
- bib-Dateien ✓
- bib-Dateien: Beispiel ✓
- bib-Dateien: Befunde ✓
- bst-Datei ✓
- Komponenten ✓
- Strukturierung des Interpreters ✓
- Das Hauptprogramm: Beispiel ✓
- Das Hauptprogramm: Eingebaute Funktionen ✓
- Der Sprach-Kern x**
- Der Sprach-Kern: Befunde ✓
- Die Datenbank-Funktionen ✓
- Die Typesetting-Funktionen ✓
- Die Text-Funktionen ✓
- Dokumentenlage ✓
- BibTeX 1.0 ✓
- Therapie: Reimplementierung ✓
- Und wohin geht die Reise? ✓

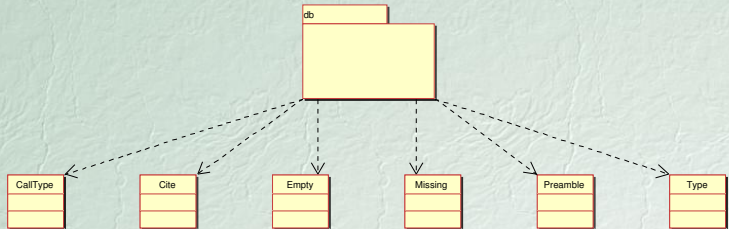
## Der Sprach-Kern: Befunde

- Schlechtes Debugging
- Eingabe und Ausgabe nur ASCII (oder höchstens 8-Bit-clean)
- Feste Größe für interne Datenstrukturen
- 80 Zeichen breite Ausgabezeilen
  - ◆ Automatischer Umbruch
  - ◆ % wird eingefügt
- Nur `newline$` leert den Ausgabepuffer.  
Fehlt das Am Ende, geht ein Teil der Ausgabe verloren.

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des	
Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	✓
Eingebaute Funktionen	✓
Der Sprach-Kern	✓
Der Sprach-Kern: Befunde	✗
Die Datenbank-Funktionen	
Die Typesetting-Funktionen	
Die Text-Funktionen	
Dokumentenlage	
BibTeX 1.0	
Therapie: Reimplementierung	
Und wohin geht die Reise?	



## Die Datenbank-Funktionen

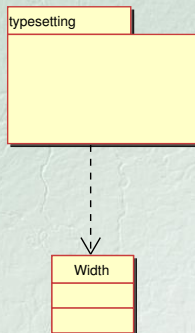


## Befunde

- Kein Konstrukt zum Iterieren über alle Felder
- Keine Manipulation der Datenbank:  
Einfügen, Löschen

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des	
Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	✓
Eingebaute Funktionen	✓
Der Sprach-Kern	✓
Der Sprach-Kern: Befunde	✓
Die Datenbank-Funktionen	✗
Die Typesetting-Funktionen	
Die Text-Funktionen	
Dokumentenlage	
BibTeX 1.0	
Therapie: Reimplementierung	
Und wohin geht die Reise?	

# Die Typesetting-Funktionen

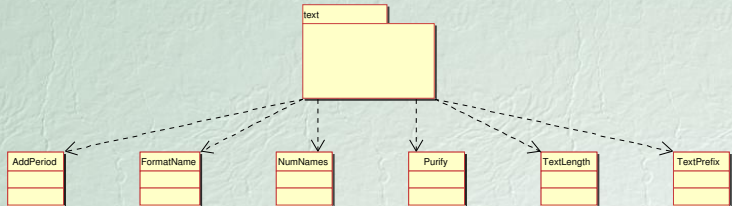


## Befunde

- cmr10 fest eingebaut
- Keine Berücksichtigung von  $\TeX$ -Konstrukten

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des	
Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	✓
Eingebaute Funktionen	✓
Der Sprach-Kern	✓
Der Sprach-Kern: Befunde	✓
Die Datenbank-Funktionen	✓
Die Typesetting-Funktionen	✗
Die Text-Funktionen	
Dokumentenlage	
Bib <sub>TeX</sub> 1.0	
Therapie: Reimplementierung	
Und wohin geht die Reise?	

## Die Text-Funktionen



### Befunde

- Klammer-Zählung
- Hochgradig heuristisch
- Nicht dokumentiert

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des	
Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	✓
Eingebaute Funktionen	
Der Sprach-Kern	✓
Der Sprach-Kern: Befunde	✓
Die Datenbank-Funktionen	✓
Die Typesetting-Funktionen	✓
Die Text-Funktionen	✗
Dokumentenlage	
Bib <sub>TeX</sub> 1.0	
Therapie: Reimplementierung	
Und wohin geht die Reise?	

## Dokumentenlage

`purify$`

Pops the top (string) literal, removes nonalphanumeric characters except for white-space characters and hyphens and ties (these all get converted to a space), removes certain alphabetic characters contained in the control sequences associated with a “special character”, and pushes the resulting string.

aus: Oren Patashnik: Designing Bib<sub>T</sub>EX Styles”

## Befunde

- Keine ausreichende Spezifikation
- Quellen sind die letzte Quelle der Wahrheit

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des	
Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	✓
Eingebaute Funktionen	✓
Der Sprach-Kern	✓
Der Sprach-Kern: Befunde	✓
Die Datenbank-Funktionen	✓
Die Typesetting-Funktionen	✓
Die Text-Funktionen	✓
Dokumentenlage	✗
Bib <sub>T</sub> EX 1.0	
Therapie: Reimplementierung	
Und wohin geht die Reise?	

# BibTeX 1.0

- Angekündigt in TUGboat 15(3), September 1994
- Keine Neuerungen für die Sprache
- Neues für die Standard-Styles
- Neuerungen für die Datenbank-Dateien

- ◆ Namespaces für Datenbanken

- ◆ @include

```
@include{file}
```

- ◆ @alias

```
@alias{newKey=oldKey}
```

- ◆ @modify

```
@modify{key, field=value}
```

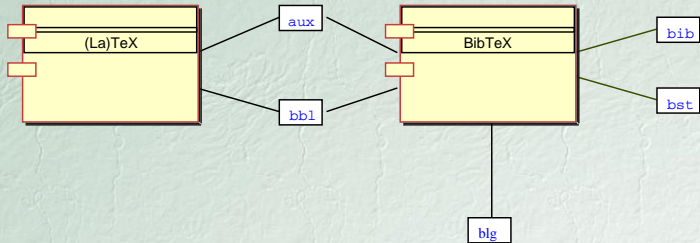
Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des	
Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	✓
Eingebaute Funktionen	✓
Der Sprach-Kern	✓
Der Sprach-Kern: Befunde	✓
Die Datenbank-Funktionen	✓
Die Typesetting-Funktionen	✓
Die Text-Funktionen	✓
Dokumentenlage	✓
BibTeX 1.0	✗
Therapie: Reimplementierung	
Und wohin geht die Reise?	

# Therapie: Reimplementierung

- Erster Prototyp in Perl abgebrochen
- Implementierungssprache: Java
- alpha, plain, unsrt, abbrev auf sample.bib läuft
- Einige Absonderlichkeiten sind noch nicht tief genug analysiert und nachgebildet.

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	✓
Eingebaute Funktionen	✓
Der Sprach-Kern	✓
Der Sprach-Kern: Befunde	✓
Die Datenbank-Funktionen	✓
Die Typesetting-Funktionen	✓
Die Text-Funktionen	✓
Dokumentenlage	✓
Bib <sub>T</sub> EX 1.0	✓
Therapie: Reimplementierung	
✗	
Und wohin geht die Reise?	

## Und wohin geht die Reise?



- Enge Integration der Komponenten erwünscht
- Welche Komponenten werden benötigt?
- Wie sieht die Komponentenarchitektur für ein  $\text{\TeX}$ -basiertes Textsatzsystem aus?

Die Literaturdatenbank	✓
Kooperation	✓
aux-Datei	✓
bib-Dateien	✓
bib-Dateien: Beispiel	✓
bib-Dateien: Befunde	✓
bst-Datei	✓
Komponenten	✓
Strukturierung des	
Interpreters	✓
Das Hauptprogramm	✓
Das Hauptprogramm: Beispiel	✓
Eingebaute Funktionen	✓
Der Sprach-Kern	✓
Der Sprach-Kern: Befunde	✓
Die Datenbank-Funktionen	✓
Die Typesetting-Funktionen	✓
Die Text-Funktionen	✓
Dokumentenlage	✓
Bib <sub>TeX</sub> 1.0	✓
Therapie: Reimplementierung	✓
Und wohin geht die Reise?	✗