

# Bib<sup>L</sup>A<sub>T</sub>E<sup>X</sup>and Biber, some history and plans for the future

Philip Kime

19.10.2019 / DANTE Herbsttagung

# What is Bib $\text{\LaTeX}$ ?

- ▶ Bib $\text{\LaTeX}$  is a comprehensive bibliography and references typesetting system for  $\text{\LaTeX}$

# What is Bib $\LaTeX$ ?

- ▶ Bib $\LaTeX$  is a comprehensive bibliography and references typesetting system for  $\LaTeX$
- ▶ Styles are completely separate from the backend (Bib $\TeX$ , Biber)

# What is Bib $\text{\LaTeX}$ ?

- ▶ Bib $\text{\LaTeX}$  is a comprehensive bibliography and references typesetting system for  $\text{\LaTeX}$
- ▶ Styles are completely separate from the backend (Bib $\text{\TeX}$ , Biber)
- ▶ It uses the `.bb1` file like a database

# Origins of BibL<sup>A</sup>T<sub>E</sub>X

Introduction

BibL<sup>A</sup>T<sub>E</sub>X

Biber

Statistics and  
Features

Some Statistics

Present and  
Future

Present

Future

- ▶ Started in 2006 by Philipp Lehman ...

# Origins of Bib $\text{\LaTeX}$

- ▶ Started in 2006 by Philipp Lehman ...
- ▶ ... who disappeared in 2011

- ▶ Started in 2006 by Philipp Lehman ...
- ▶ ... who disappeared in 2011
- ▶ Sophisticated styles for humanities and data/style separation

- ▶ Started in 2006 by Philipp Lehman ...
- ▶ ... who disappeared in 2011
- ▶ Sophisticated styles for humanities and data/style separation
- ▶ Started to run into limitations imposed by Bib $\TeX$  as backend in 2008

# What is Biber?

- ▶ Biber is the default backend for Bib $\text{\LaTeX}$ , written for Bib $\text{\LaTeX}$  and designed to overcome Bib $\text{\TeX}$  limitations

# What is Biber?

- ▶ Biber is the default backend for Bib $\text{\LaTeX}$ , written for Bib $\text{\LaTeX}$  and designed to overcome Bib $\text{\TeX}$  limitations
- ▶ Its main job is sorting (full multi-field, UCA with CLDR tailoring)

# What is Biber?

- ▶ Biber is the default backend for Bib $\LaTeX$ , written for Bib $\LaTeX$  and designed to overcome Bib $\TeX$  limitations
- ▶ Its main job is sorting (full multi-field, UCA with CLDR tailoring)
- ▶ It also does a **lot** of other things ...

# Origins of Biber

- ▶ Started around 2008 by François Charette as dedicated Bib $\LaTeX$  backend

## Introduction

Bib $\LaTeX$

**Biber**

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

- ▶ Started around 2008 by François Charette as dedicated Bib $\LaTeX$  backend
- ▶ PL put me in touch with FC when I found limitations in Bib $\TeX$  preventing implementation of APA style

# Origins of Biber

## Introduction

Bib $\LaTeX$

**Biber**

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

- ▶ Started around 2008 by François Charette as dedicated Bib $\LaTeX$  backend
- ▶ PL put me in touch with FC when I found limitations in Bib $\TeX$  preventing implementation of APA style
- ▶ My first commit April 14th 2009

- ▶ Started around 2008 by François Charette as dedicated Bib $\LaTeX$  backend
- ▶ PL put me in touch with FC when I found limitations in Bib $\TeX$  preventing implementation of APA style
- ▶ My first commit April 14th 2009
- ▶ Became sole developer early 2011

# Since 2009

▶ ~4000 total commits

## Introduction

Bib $\LaTeX$

Biber

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

# Since 2009

- ▶ ~4000 total commits
- ▶ ~250k total **manual** downloads

## Introduction

Bib $\LaTeX$   
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Since 2009

- ▶ ~4000 total commits
- ▶ ~250k total **manual** downloads
- ▶ ~50k lines of code

## Introduction

Bib<sub>A</sub>T<sub>E</sub>X  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Since 2009

- ▶ ~4000 total commits
- ▶ ~250k total **manual** downloads
- ▶ ~50k lines of code
- ▶ ~1k issues/bugs, 95% resolved

## Introduction

Bib $\LaTeX$

Biber

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

# Since 2009

- ▶ ~4000 total commits
- ▶ ~250k total **manual** downloads
- ▶ ~50k lines of code
- ▶ ~1k issues/bugs, 95% resolved
- ▶ Germany is largest user (~33%) by Sourceforge downloads (twice US)

## Introduction

Bib<sub>A</sub>T<sub>E</sub>X

Biber

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

# Features

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\text{L}\text{A}\text{T}\text{E}\text{X}\leftrightarrow\text{UTF-8}$  conversion

# Features

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\text{\LaTeX}\leftrightarrow\text{UTF-8}$  conversion
- ▶ Best-in-class sorting

# Features

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\text{\LaTeX}\leftrightarrow\text{UTF-8}$  conversion
- ▶ Best-in-class sorting
- ▶ Completely customisable datamodel

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}\leftrightarrow\text{UTF-8}$  conversion
- ▶ Best-in-class sorting
- ▶ Completely customisable datamodel
- ▶ Dynamic data mapping

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\text{\LaTeX}\leftrightarrow\text{UTF-8}$  conversion
- ▶ Best-in-class sorting
- ▶ Completely customisable datamodel
- ▶ Dynamic data mapping
- ▶ Automatic name and name list disambiguation

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X} \leftrightarrow \text{UTF-8}$  conversion
- ▶ Best-in-class sorting
- ▶ Completely customisable datamodel
- ▶ Dynamic data mapping
- ▶ Automatic name and name list disambiguation
- ▶ Biber Tool mode

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X} \leftrightarrow \text{UTF-8}$  conversion
- ▶ Best-in-class sorting
- ▶ Completely customisable datamodel
- ▶ Dynamic data mapping
- ▶ Automatic name and name list disambiguation
- ▶ Biber Tool mode
- ▶ Generalised name support

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\LaTeX \leftrightarrow$  UTF-8 conversion
- ▶ Best-in-class sorting
- ▶ Completely customisable datamodel
- ▶ Dynamic data mapping
- ▶ Automatic name and name list disambiguation
- ▶ Biber Tool mode
- ▶ Generalised name support
- ▶ Sophisticated label auto-disambiguation and generation

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\LaTeX \leftrightarrow$  UTF-8 conversion
- ▶ Best-in-class sorting
- ▶ Completely customisable datamodel
- ▶ Dynamic data mapping
- ▶ Automatic name and name list disambiguation
- ▶ Biber Tool mode
- ▶ Generalised name support
- ▶ Sophisticated label auto-disambiguation and generation
- ▶ Customisable data inheritance

- ▶ Full Unicode (UTF-8) support including auto-reencoding and  $\LaTeX \leftrightarrow$  UTF-8 conversion
- ▶ Best-in-class sorting
- ▶ Completely customisable datamodel
- ▶ Dynamic data mapping
- ▶ Automatic name and name list disambiguation
- ▶ Biber Tool mode
- ▶ Generalised name support
- ▶ Sophisticated label auto-disambiguation and generation
- ▶ Customisable data inheritance
- ▶ Full support for all types of “related” entries

# Technical History

29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format

## Introduction

Bib $\LaTeX$

Biber

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

# Technical History

29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format

18.05.2009 Moved from Subversion to Git

## Introduction

Bib $\LaTeX$

Biber

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`

## Introduction

Bib<sub>La</sub>T<sub>E</sub>X  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support

## Introduction

Bib<sub>La</sub>T<sub>E</sub>X  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib

## Introduction

Bib<sub>A</sub>T<sub>E</sub>X  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries

## Introduction

Bib<sub>A</sub>T<sub>E</sub>X  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT

## Introduction

Bib<sub>L</sub>E<sub>X</sub>  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT
- 2011 Became sole maintainer

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT
- 2011 Became sole maintainer
- 04.2011 `uniquelist` option implemented, APA style now fully possible...

## Introduction

Bib<sub>L</sub>E<sub>T</sub>X  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT
- 2011 Became sole maintainer
- 04.2011 `uniquelist` option implemented, APA style now fully possible...
- 06.2011 customisable `labelalpha` and auto-disambiguation methods

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT
- 2011 Became sole maintainer
- 04.2011 `uniquelist` option implemented, APA style now fully possible...
- 06.2011 customisable `labelalpha` and auto-disambiguation methods
- 06.2011 Datasource mapping functionality

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT
- 2011 Became sole maintainer
- 04.2011 `uniquelist` option implemented, APA style now fully possible...
- 06.2011 customisable `labelalpha` and auto-disambiguation methods
- 06.2011 Datasource mapping functionality
- 04.2012 Generalised sorting lists to make multiple, differently sorted bibliographies and other bib lists possible

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT
- 2011 Became sole maintainer
- 04.2011 `uniquelist` option implemented, APA style now fully possible...
- 06.2011 customisable `labelalpha` and auto-disambiguation methods
- 06.2011 Datasource mapping functionality
- 04.2012 Generalised sorting lists to make multiple, differently sorted bibliographies and other bib lists possible
- 2012 Generalised, explicit datamodel implemented and passed in `.bcf`

## Introduction

Bib<sub>L</sub>A<sub>T</sub>E<sub>X</sub>  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT
- 2011 Became sole maintainer
- 04.2011 `uniquelist` option implemented, APA style now fully possible...
- 06.2011 customisable `labelalpha` and auto-disambiguation methods
- 06.2011 Datasource mapping functionality
- 04.2012 Generalised sorting lists to make multiple, differently sorted bibliographies and other bib lists possible
- 2012 Generalised, explicit datamodel implemented and passed in `.bcf`
- 12.2012 Biber Tool mode

## Introduction

Bib<sub>La</sub>T<sub>E</sub>X  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
Future

# Technical History

- 29.04.2009 Started on `.bcf` format (XML) which is ~2000 lines as opposed to the **one** line original control file format
- 18.05.2009 Moved from Subversion to Git
- 05.2010 Started hacking `btparse`
- 09.2010 “binary” distribution and real CLDR support
- 10.2010 TLContrib
- 11.2010 Dynamic entry sets, related entries
- 12.2010 Real multi-field sorting, `.bcf` to HTML via XSLT
- 2011 Became sole maintainer
- 04.2011 `uniquelist` option implemented, APA style now fully possible...
- 06.2011 customisable `labelalpha` and auto-disambiguation methods
- 06.2011 Datasource mapping functionality
- 04.2012 Generalised sorting lists to make multiple, differently sorted bibliographies and other bib lists possible
- 2012 Generalised, explicit datamodel implemented and passed in `.bcf`
- 12.2012 Biber Tool mode
- 2016–2017 Data annotations, extended name format, full generalisation of all aspects of name handling (sorting key, name parts etc.), ISO8601 date handling

- ▶ Bib $\LaTeX$ - Philip Kime and Moritz Wemheuer

- ▶ Bib $\LaTeX$ - Philip Kime and Moritz Wemheuer
- ▶ Biber - Philip Kime (plus release builders)

# Planned Developments

## Introduction

Bib<sup>L</sup>A<sub>T</sub>E<sub>X</sub>

Biber

## Statistics and Features

Some Statistics

## Present and Future

Present

**Future**

- ▶ Multiscript support

# Planned Developments

- ▶ Multiscript support
- ▶ Quick-start documentation

## Introduction

Bib<sub>A</sub>T<sub>E</sub>X  
Biber

## Statistics and Features

Some Statistics

## Present and Future

Present  
**Future**

# Planned Developments

- ▶ Multiscript support
- ▶ Quick-start documentation
- ▶ Granular regression test setup for Bib $\text{\LaTeX}$

## Introduction

Bib $\text{\LaTeX}$

Biber

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

# Planned Developments

- ▶ Multiscript support
- ▶ Quick-start documentation
- ▶ Granular regression test setup for Bib $\LaTeX$
- ▶ “Stand-alone” `.bbl`

## Introduction

Bib $\LaTeX$

Biber

## Statistics and Features

Some Statistics

## Present and Future

Present

Future

# Planned Developments

- ▶ Multiscript support
- ▶ Quick-start documentation
- ▶ Granular regression test setup for Bib $\text{\LaTeX}$
- ▶ “Stand-alone” `.bbl`
- ▶ Finding someone to help support Biber