

# Converting Different Document Formats

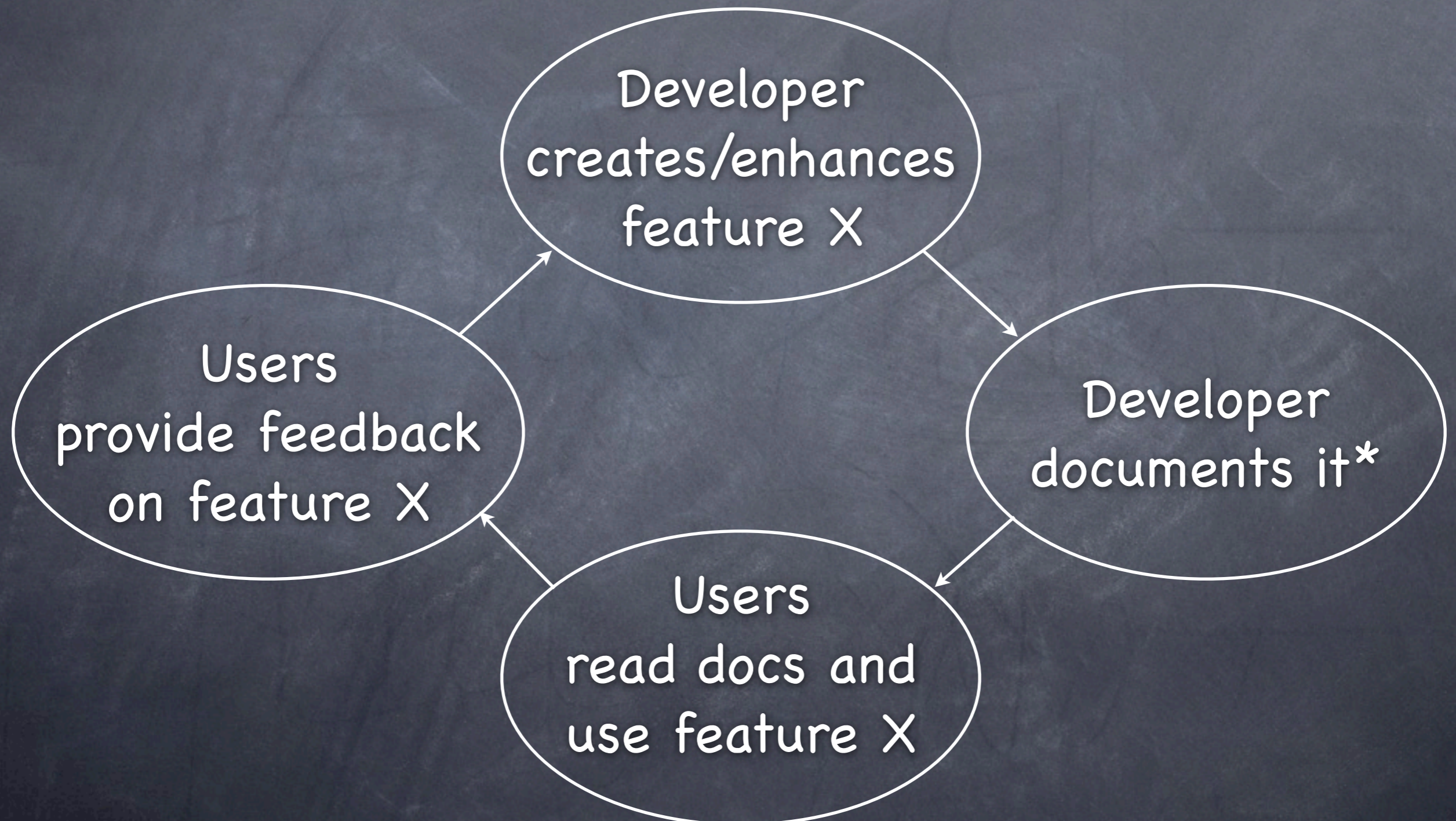
Benedict Reuschling  
bcr@FreeBSD.org

Developer Summit Track, EuroBSDcon 2011  
Maarssen, The Netherlands

# Why this talk?

- Make life easier for developers and documenters alike
- Create more documentation for our community
- Turn notes into rough draft documents for documentation committers to refine

# Development cycle from a documentation perspective



\* Involve documentation community early to get help

# Document formats that FreeBSD uses right now

- man/mandoc
- DocBook SGML, converted into
  - plain-text
  - XHTML
  - PDF
- MoinMoin wiki syntax ([wiki.FreeBSD.org](http://wiki.FreeBSD.org))

# Document formats that are becoming relevant

- DocBook XML
- HTML (5)
- OpenOffice/LibreOffice ODT
- EPUB (see also: my talk at BSDCan 2011)
- non-XML markup languages (AsciiDoc, reStructured Text)

# pandoc – A universal document converter

- Author: John MacFarlane, Philosophy Professor, University of California, Berkeley
- Homepage: <http://johnmacfarlane.net/pandoc/>
- Source code: <http://code.google.com/p/pandoc/>
- Available from:
  - pkgsrc: <http://pkgsrc.se/wip/pandoc>
  - Ports Collection: `textproc/hs-pandoc`
- Warning: long compile time, use the package!

# Basic Syntax

from

to

output

```
$ pandoc -f fooformat -t barformat infile -o outfile
```

## Example:

```
$ pandoc -f html -t rtf wikipage.html -o wikipage.rtf
```

Let pandoc detect input/output based on file extension:

```
$ pandoc wikipage.html -o wikipage.rtf
```

# Converting a wiki page (XHTML) to DocBook XML

- Why?

- Good howto's available in the FreeBSD wiki

- DocBook XML is the successor of DocBook SGML

- How:

```
$ pandoc -s -S -t docbook wikipage -o wikipage.db
```



# Create an EPUB

- Why?

- Read documents on your tablet/smart phone

- Source: XHTML document(s), including images and CSS

- Conversion:

```
pandoc foo.html -o mybook.epub
```

- Test version of FreeBSD handbook:

```
http://people.freebsd.org/~bcr/epub/handbook.epub
```

# Summary

- Pandoc is powerful in transforming document formats
- Produces good results, even for large/complex docs
- Manual cleanup may be required afterwards
- Make some test conversions, see if it fits your needs
- Try out the online converter:  
<http://johnmacfarlane.net/pandoc/try>

Q&A

Thanks for listening!