Efficient Scientific Research
with (Multi)Markdown

Output 2014
Dipl.-Inf. Thomas Kühn

DRESDEN
concept
Exzellenz aus
Wissenschaft
und Kultur

Cartoon by Nick D Kim strange-matter.net. Used by permission.
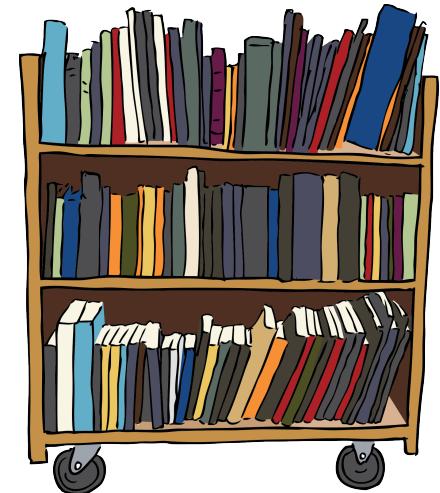
Reading   Writing   Organizing

Images from OpenClipart.org (Creative Commons by Steve Lambert)
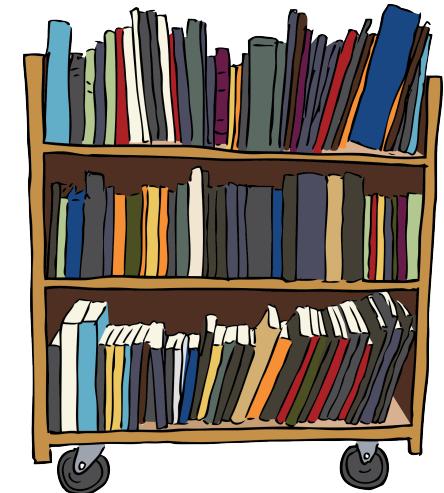
**Reading** **Writing** **Organizing**
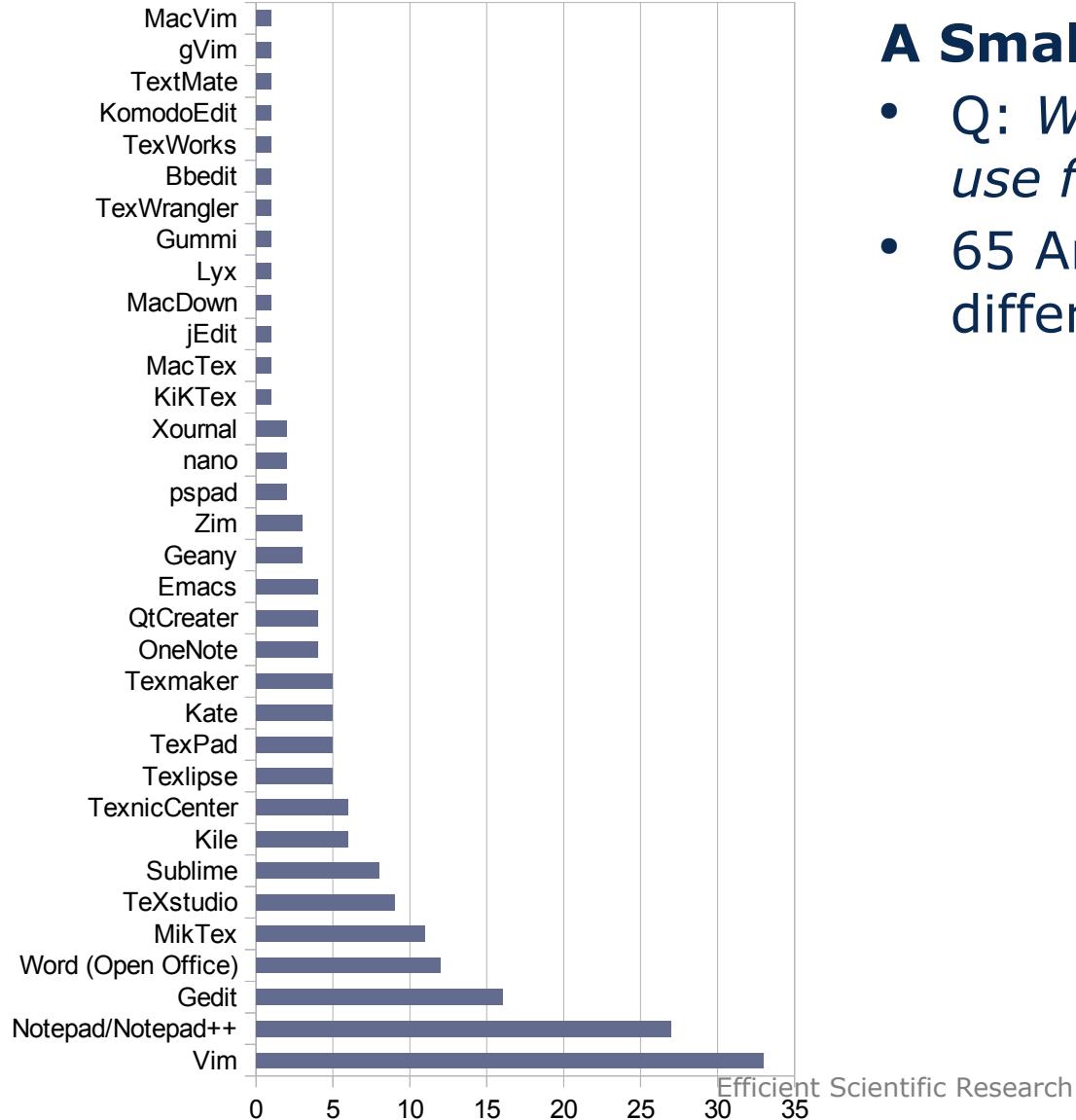
Images from OpenClipart.org (Creative Commons by Steve Lambert)

A Small Survey

- Q: *What tools do you all use for your daily writing?*

- 65 Answers named 34 different Tools

A Small Survey

- Q: *What tools do you all use for your daily writing?*
- 65 Answers named 34 different Tools

A Small Survey

- Q: *What tools do you all use for your daily writing?*

- 65 Answers named 34 different Tools

Bar chart (left) listing tools with approximate values:
- MacVim ~1
- gVim ~1
- TextMate ~1
- KomodoEdit ~1
- TexWorks ~1
- Bbedit ~1
- TexWrangler ~1
- Gummi ~1
- Lyx ~1
- MacDown ~1
- jEdit ~1
- MacTex ~1
- KiKTex ~1
- Xournal ~2
- nano ~2
- pspad ~2
- Zim ~3
- Geany ~3
- Emacs ~4
- QtCreater ~4
- OneNote ~4
- Texmaker ~5
- Kate ~5
- TexPad ~5
- Texlipse ~5
- TexnicCenter ~6
- Kile ~6
- Sublime ~8
- TeXstudio ~9
- MikTex ~11
- Word (Open Office) ~12
- Gedit ~16
- Notepad/Notepad++ ~27
- Vim ~33

Pie chart (right):
- Word Processor: 15, 8%
- Latex Environment: 56, 30%
- Text Editor: 114, 62%
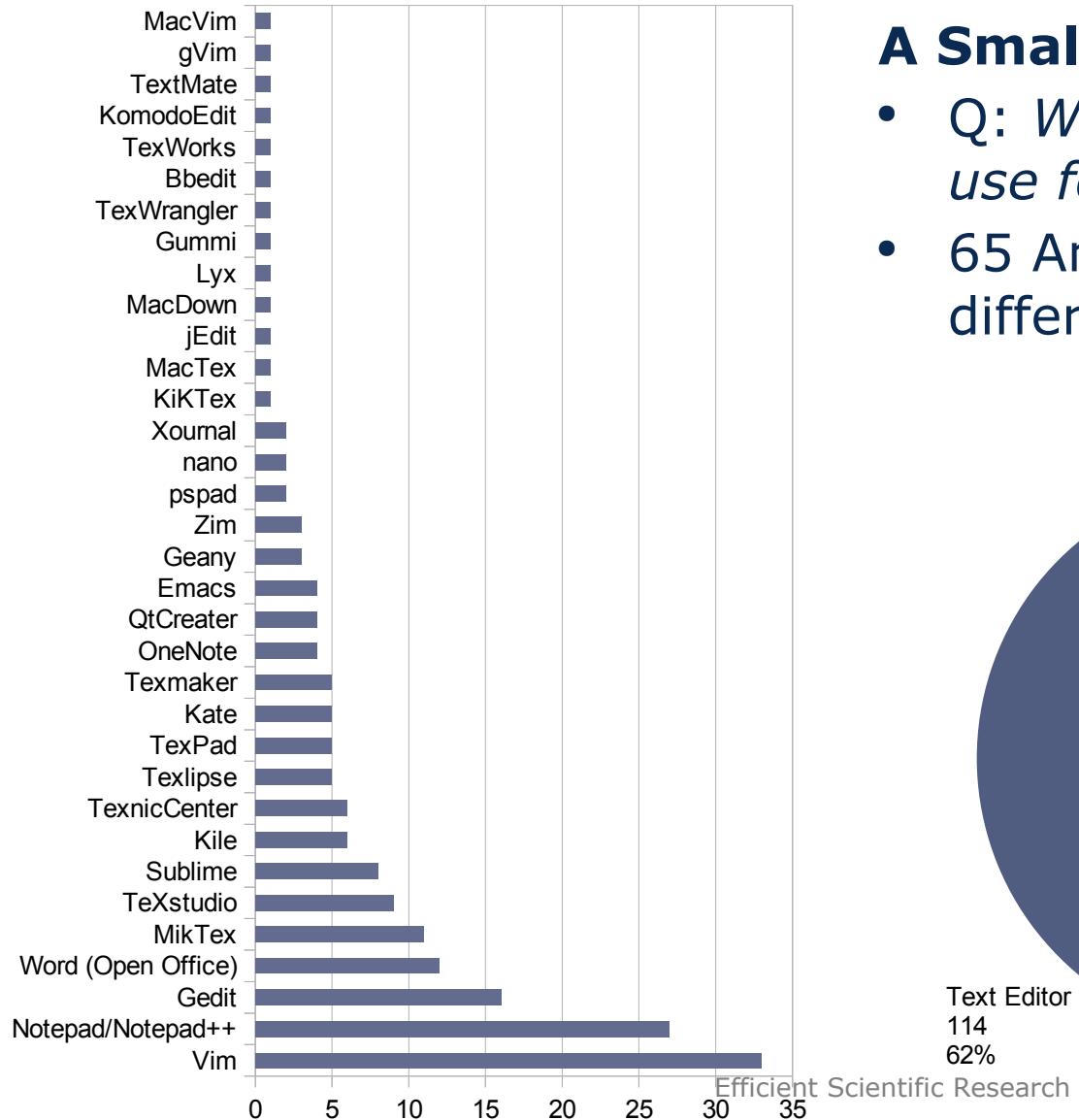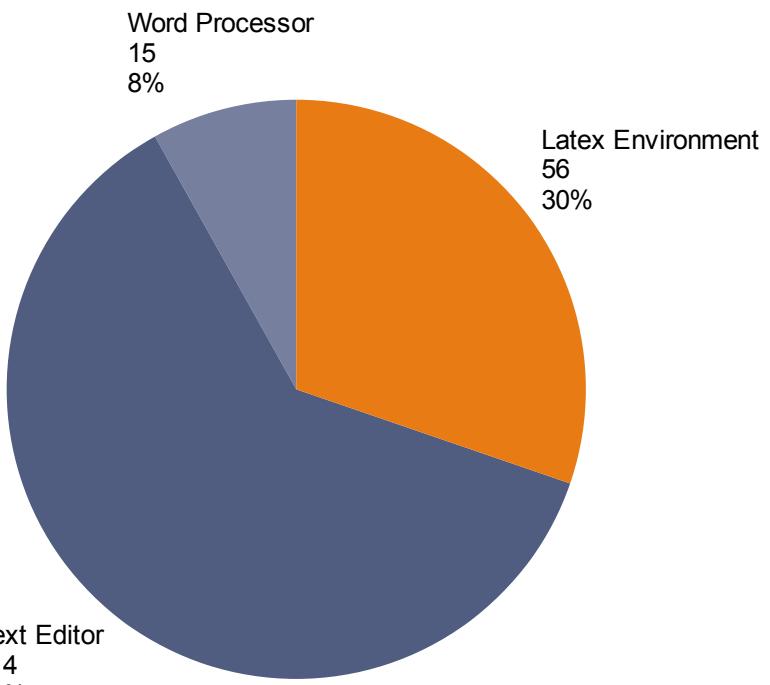
Survey Results

- Huge amount of tools in use
- Each participant used ~2,6 tools on average
- At least one text editor or a LaTeX environment

Consequences

- Unmanaged redundancy of knowledge between different document formats
- Leads to *Copy-And-Paste* behavior
- Writing performance is different among different tools
- LaTeX environments require the use of latex ideoms

What is a better strategie for efficient writing?

Requirements

- Focus on contant rather than ideoms
- Concise ways to structure text
- Direct support for citations, figures, tables
- Transformation to arbitrary formats
- **One content source many output formats**

Markdown[1]

- Focus a simple syntax to both *read* and *write*
- Developed by *John Gruber* to write for the web (HTML)

```
# This is another H1

## This is another H2

*This is an italic text*

 This is a text with an underscore 

**This is a bold text**

The following is a Quotation:

> Lorem ipsum dolor sit amet,

![This is an Image](rosi.png "This is the alt text")
```

```
## Lists

* Item of a bullet list
* Another Item

1. An Item of an ordered list
2. Yet another Item

## Highlights

This is a code block:
    $ pandoc --help
And this a link to [Pandoc]
(http://johnmacfarlane.net/pandoc/).
```

1) http://daringfireball.net/projects/markdown/syntax

(Multi)Markdown[2]

- Markdown extension developed by *Fletcher T. Penney* to support research activities
- Important features are: *Footnotes, Tables, Citations, …*

```
% title
% author(s) (separated by semicolons)
% date

# Citations and Footnotes

This is a text to cite [@author], whereas
the following Noun[^I am a footnote]
carries a footnote.

# Special Features

This ~~is striked out text.~~
H~2~O is a liquid.  2^10^ is 1024.
$\forall x \in X, \exists y \leq \epsilon$
```

```
# Tables

  Right      Left     Center     Default
 -------    ------   ----------   -------
      12     12         12             12
     123     123        123           123


 +--------------+--------------+
 | Fruit        | Price        |
 +==============+==============+
 | Bananas      | $1.34        |
 +--------------+--------------+
 | Oranges      | $2.10        |
 +--------------+--------------+
```

2) http://fletcherpenney.net/multimarkdown/

TECHNISCHE UNIVERSITÄT DRESDEN

(Multi)Markdown[2]

- Markdown extension developed by *Fletcher T. Penney* to support research activities
- Important features are: *Footnotes, Tables, Citations, ...*

```
[…]

\title{title}
\author{author(s) (separated by
semicolons)}
\date{date}

\begin{document}
\maketitle

\section{Citations and Footnotes}
This is a text to cite {[}@author{]},
whereas the following Noun{[}\^{}I
am a footnote{]} carries a footnote.

\section{Special Features}
This \sout{is striked out text.}
H\textsubscript{2}O is a liquid.
2\textsuperscript{10} is 1024.
$\forall x \in X, \exists y \leq \epsilon$
```
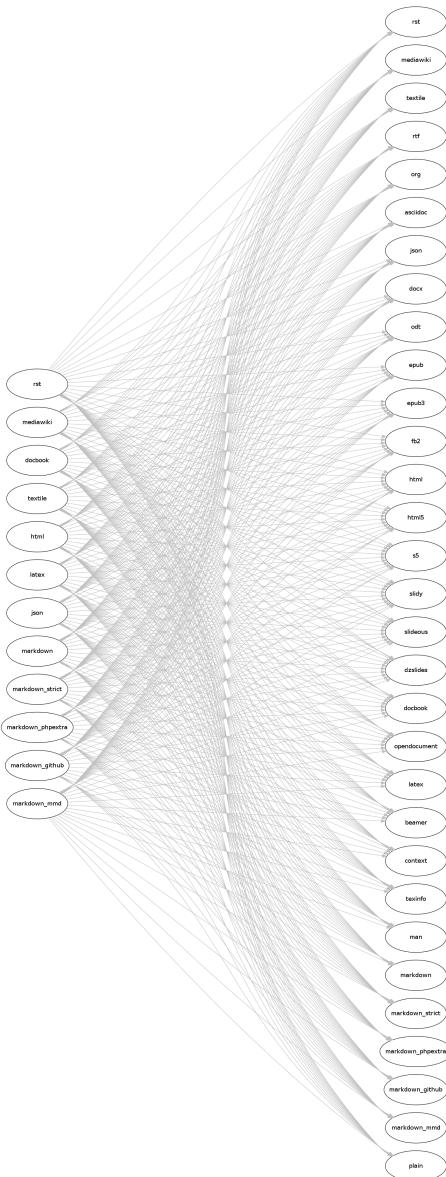
```
\section{Tables}

\begin{longtable}[c]{@{}rlcl@{}}
\toprule\addlinespace
Right & Left & Center & Default
\\\addlinespace
\midrule\endhead
12 & 12 & 12 & 12
\\\addlinespace
123 & 123 & 123 & 123
\\\addlinespace
\bottomrule
\end{longtable}

\begin{longtable}[c]{@{}ll@{}}
\toprule\addlinespace
\begin{minipage}[b
{0.21\columnwidth}\raggedright
Fruit
\end{minipage} & […]
```

Pandoc[3]

- Command line tool written in Haskell
- Available for **all** platforms
- Many extensions already included
- Direct transformation from
  - (multi)markdown, reStructuredText,
  - HTML, LaTeX, …
- To
  - XHTML, HTML 5,
  - LaTeX (including beamer slide shows),
  - ConTeXt, ODT, Word docx,
  - EPUB (v2 or v3), …
- Supports use of customized templates

```
$ pandoc -S test.md -o test.tex
```

3) http://johnmacfarlane.net/pandoc/

```
                    ttask.tex
[…]
\task{$tasktitle$}
\title{$title$}
$if(subtitle)$
\subtitle{$subtitle$}
$endif$
\author{$author$}
[…]
\begin{document}
[…]
$body$
[…]
```

Templates

- Pandoc's templates contain $variables$, conditionals, and for-loops
- $body$ is a special variable for the actual markdown content
- Large number of predefined templates

```
                    task.md
% Collaboration-Based Service-Oriented Architecture
% Dipl.-Inf. Thomas Kühn
% 30.01.2014

---
student: 'Max Mustermann'
startdate: 01.06.2014
enddate:   30.11.2014
[…]
...
# Motivation
[…]
```

Using Templates

- Add yaml_metadata_block extension
- Variable definition at the beginning of the document

```
$ pandoc -s -S "task.md"
    -f markdown+yaml_metadata_block
    -t latex --template="ttask.tex"
```

```
                    paper.md.rb
[…]
% Metamodel for Role-based Software Infrastructures
% Thomas Kühn (thomas.kuehn3@tu-dresden.de)
% 09.07.2013
[…]
<%= load("sections/sec2.md") %>
<%= load("sections/sec3.md") %>
<%= load("sections/sec4.md") %>
<%= load("sections/sec5.md") %>
<%= load("sections/sec6.md") %>
<%= load("sections/sec7.md") %>
<%= load("sections/ack.md") %>
```

```
                  section/sec5.md.rb
[…]
<% unless variant? == :beamer %>
[…]
<% else %>
## Overview
<% end %>

![Aspects of the RoSI metamodel]
(images/Merging.png)
[…]
```

Missing Features

- Include files into documents
- Conditional creation of content
- Embed source code in documents

PanRuby

- Ruby-based command line tool
- Employing ERB as template engine
- <% … %>
  Execute without return value
- <%= … %>
  Execute and return value
- github.com/Eden-06/panruby

```
$ ruby build.rb latex paper
  template.tex
```

Use Case 1 - Presentations

- Make your notes in (Multi)Markdown and convert them directly to slides

```
$ ruby build.rb beamer presentation
    ../trosi_presentation.tex
$ mlatex presentation
```
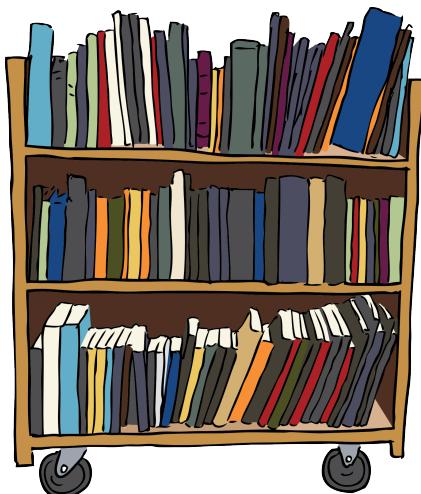
*Please note, that **mlatex** is an alias for **pdflatex** and **bibtex**, whereas the former is called multiple times.*

Conclusion

- Expressibility of (Mulit)Markdown sufficient for scientific writing
- Files writen in Markdown are easy to read and write
- Tool chain is platform independent
- Everyone can use their favorite Text Editor
  - Get fast with it
  - Write macros and helper functions
- Export your documentation to arbitrary formats
- Design your layouts and templates once

Focus on contant rather than LaTeX ideoms
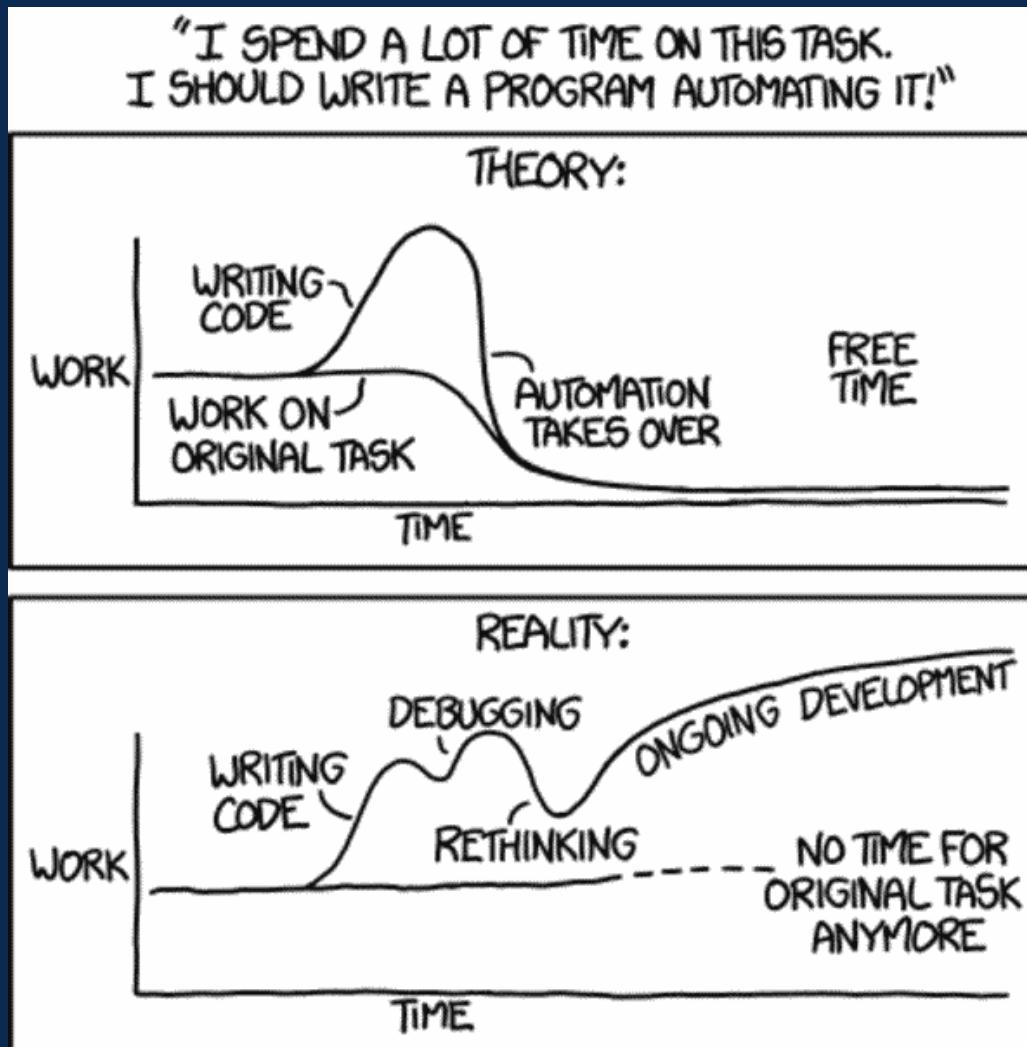
Organizing

Automated Organization

- Automated BibTex lookup for stored papers
- Automated BibTex lookup for specific publications from the web
- Automated filtering of large BibTex files
- Automated download of papers referenced by a BibTex file

Automation (By Randell Munroe from xkcd.com)