



OFFICE SYSTEMS KEY PROBLEMS and some TOOLS

Tomasz Przechlewski
2018

Key office applications

email everybody knows :-)
http://file.scirp.org/Html/9-9301914_48115.htm

Journal of Software Engineering
and Applications Vol.7 No.8(2014),
Article ID:48115,13 pages

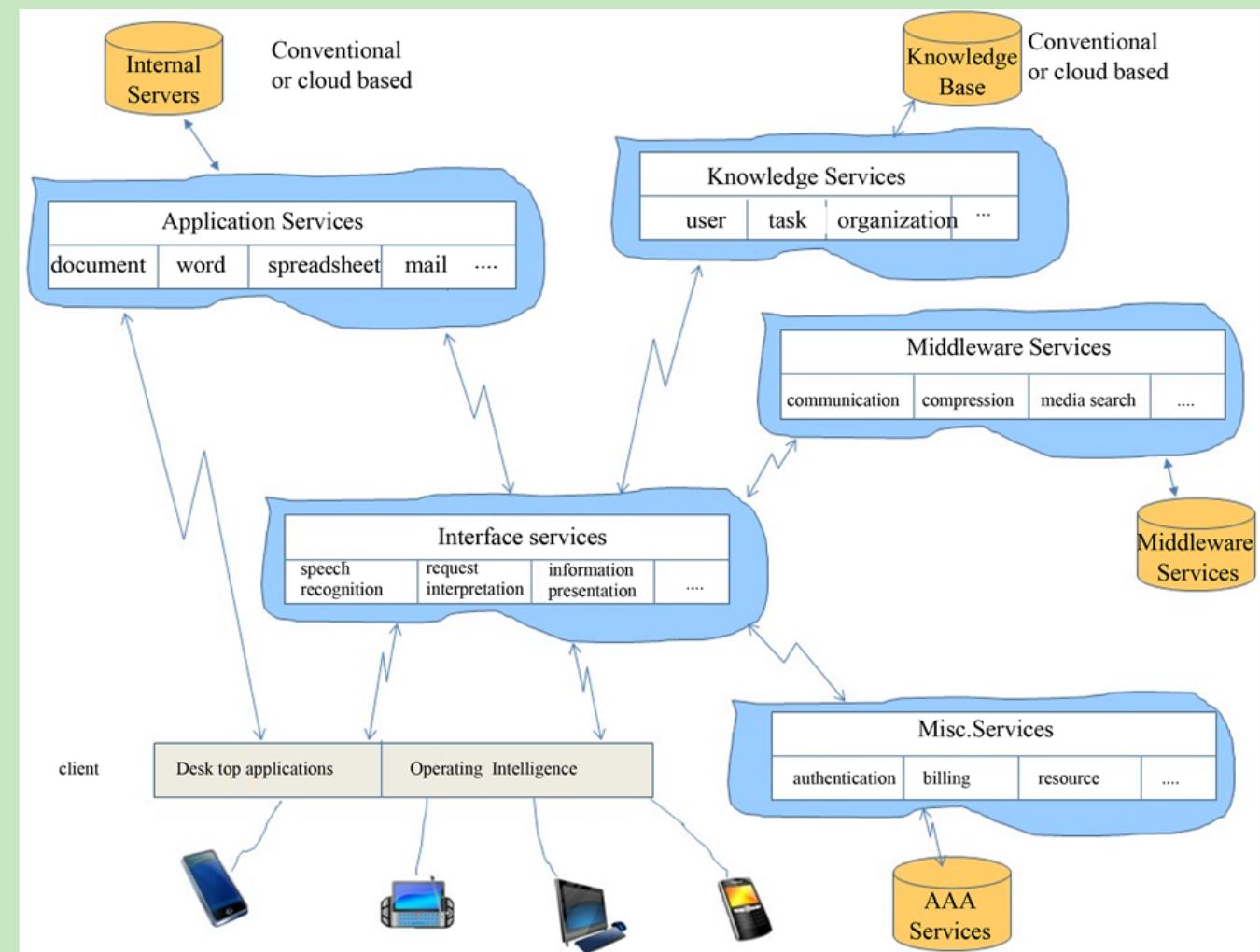
DOI:10.4236/jsea.2014.78065

Office Information Systems: A
Retrospective and a Call to
Arms
Chandra S. Amaravadi

document = MS Word ??

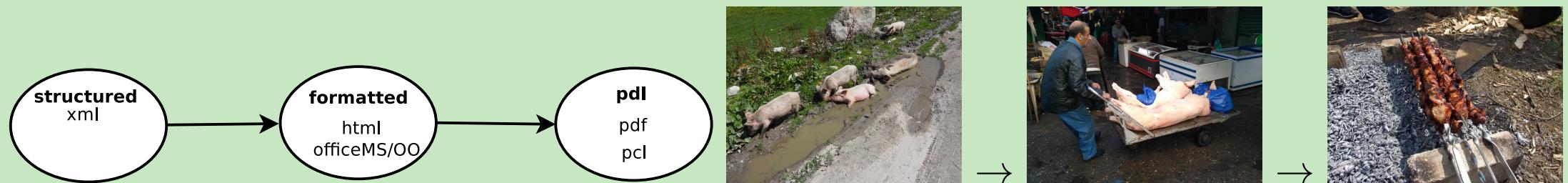
spreadsheet = MS Excel ??

presentation = MS PP



Editors

Concepts: WYSIWYG vs WYSIWYM, Direct Manipulation, Templates,
From Structural to Formatted to PDL (difficult to move opposite way as in case of pig→pork
transformation:-))



Formats (aka notations; open vs proprietary || text-based vs binary)

<https://wiki.openoffice.org/wiki/Documentation/UserGuide/Formatting>

In Apache OpenOffice, you can apply formatting to documents in two different ways. You can do so directly using menus and toolbars, or you can use styles and templates.

Markup (visual vs structural),

Key problems: consistency (visual); maintenance (document reuse/modification etc...);
preservation (write & forget (invoice) vs document life cycle); portability (time / space)

Key technologies: XML/Schema (valid document);

Formatting = transformation + stylization

Publishing = PDL, PDF (??, Document Format), HTML...

Permanent storage: document databases (noSQL), JSON

Tools [you know]...

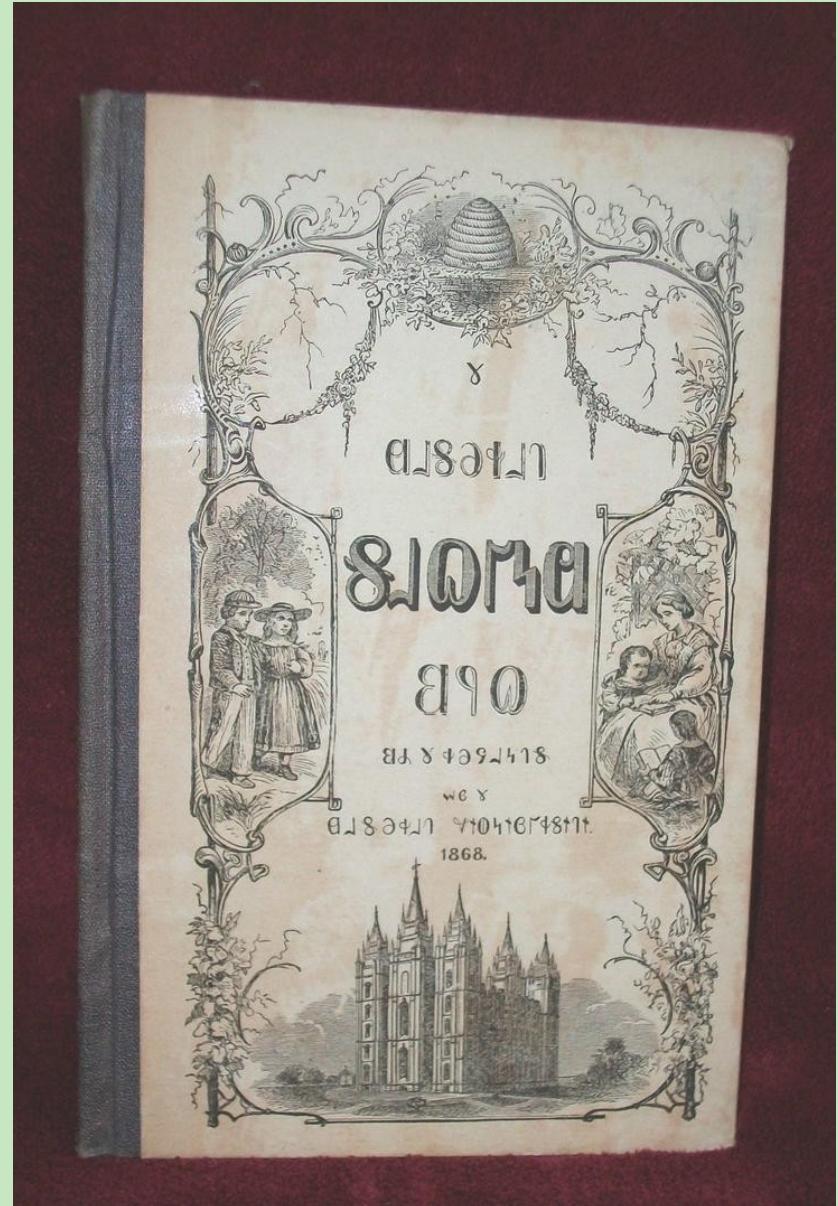
Sideways story: Deseret Alphabet

Typesetting the Deseret Alphabet with LaTeX and metafont

https://www.researchgate.net/publication/220993881_Typesetting_the_Deseret_Alphabet_with_LaTeX_and_metafont

The Deseret Alphabet was an orthographical reform for English, designed by Pitman and promoted by the Church of Jesus Christ of Latter-day Saints (the Mormons) between about 1854 and, 1875. Ultimately rejected, the Deseret Alphabet was used in four printed books, several newspapers and unprinted book manuscripts, journals, meeting minutes, letters etc...

Turkish connection = Atatrk's Reforms



Tools

WYSIWYG editors: Word, OpenOffice, LibreOffice

Schema-driven editors (WYSiQYM): Emacs (nXML mode)

Half-way: TeX/LaTeX vs LyX;

XML validators (aka parsers): xmllint, WWW browsers

XSLT transformers: WWW browsers

Publishing: Adobe Acrobat, WWW browser

```
%% TeX (LaTeX)
\begin{document}
\section{Kingdom Fungi}
The organisms
in \emph{kingdom fungi} include:

\begin{itemize}
\item mushrooms,
\item yeasts,
\item molds,
\item others.
\end{itemize}
More than 70,000 species
of fungi have been identified.

\section{Kingdom Fungi Classification}
\subsection{Division Ascomycota: Sac Fungi}
\end{document}
```

```
<!-- XML -->
<body>
<h1>Kingdom Fungi</h1>

<p>The organisms
in <em>kingdom fungi</em> include:</p>
<ul>
<li> mushrooms,
<li> yeasts,</li>
<li> molds,
<li> others.
</ul>
<p>More than 70,000 species
of fungi have been identified.</p>

<h1>Kingdom Fungi Classification</h1>
<h2*>Division Ascomycota: Sac Fungi</h2*>
</body>
```

1 Kingdom Fungi

The organisms in *kingdom fungi* include:

- mushrooms,
- yeasts,
- molds,
- others.

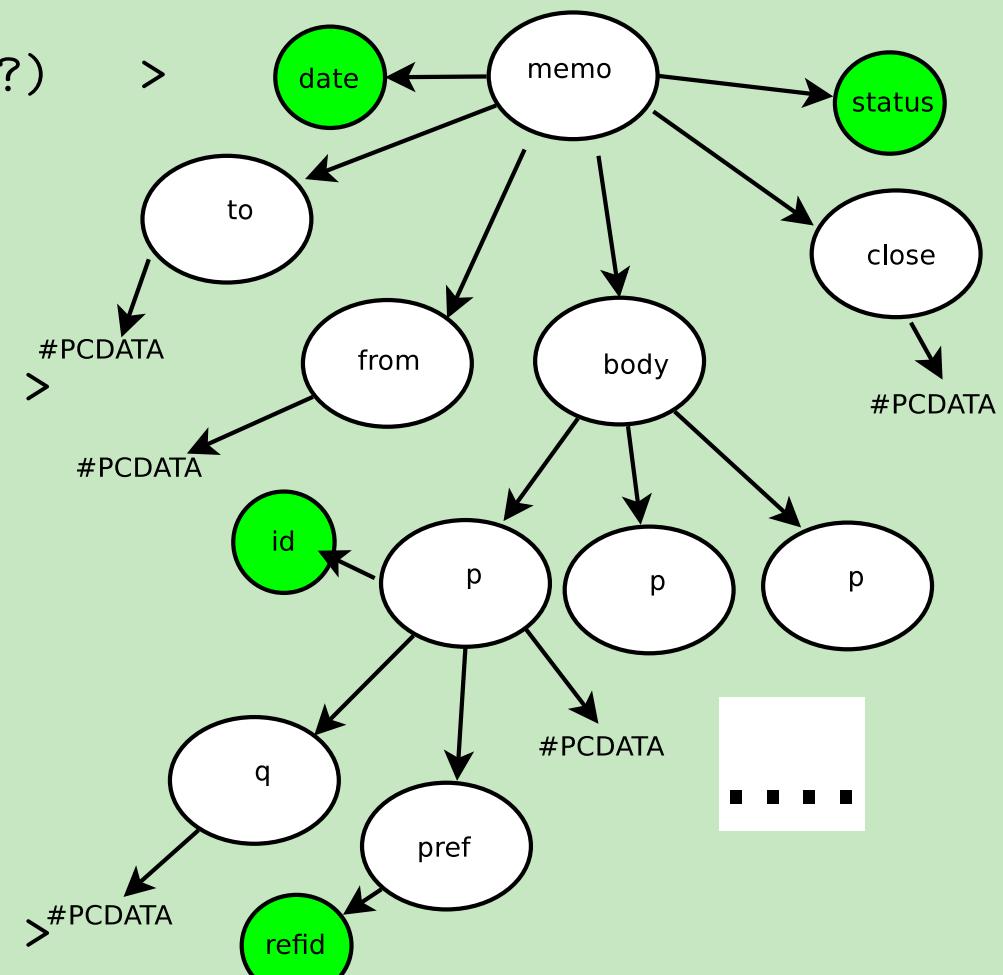
More than 70,000 species of fungi have been identified.

2 Kingdom Fungi Classification

2.1 Division Ascomycota: Sac Fungi

Schemas

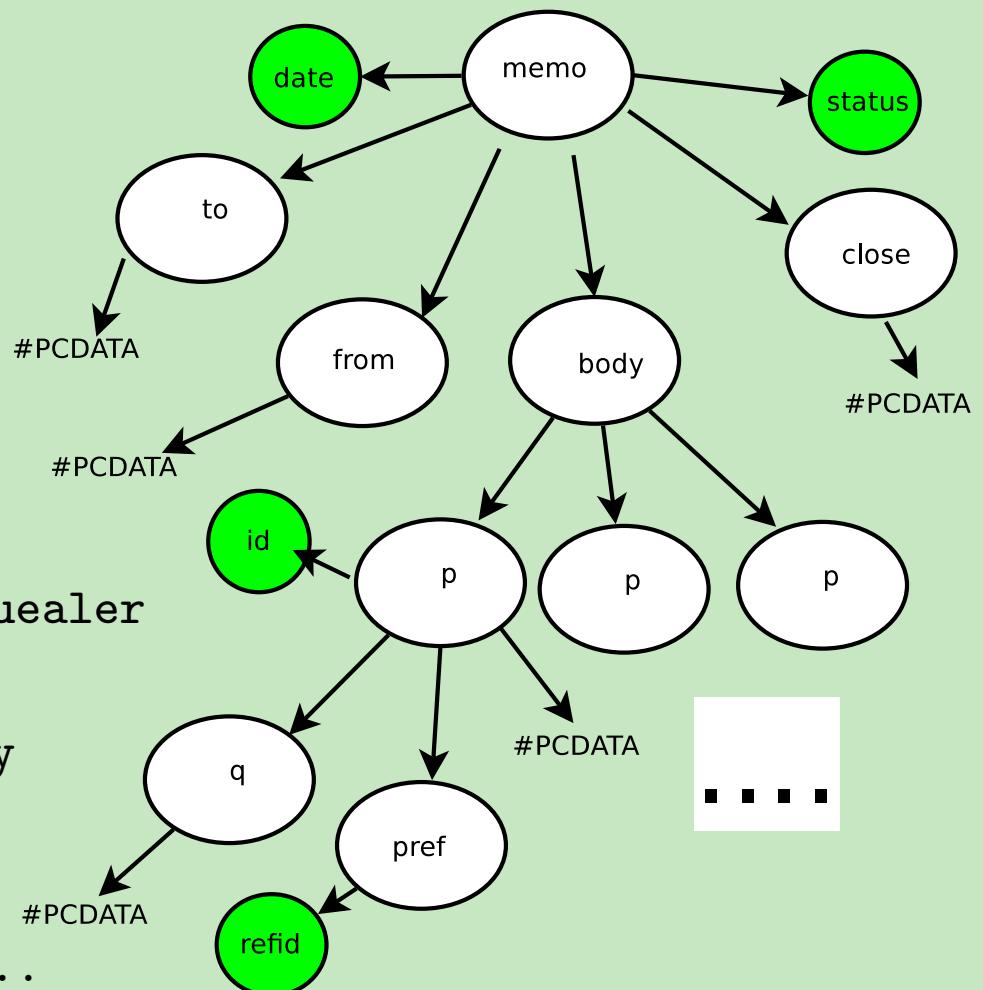
```
<!-- DTD for simple office memoranda -->
<!-- Elements -->
<!ELEMENT memo (to, from, body, close?) >
<!ELEMENT to (#PCDATA) >
<!ELEMENT from (#PCDATA) >
<!ELEMENT body (p*) >
<!ELEMENT p (#PCDATA | q | pref )* >
<!ELEMENT q (#PCDATA) >
<!ELEMENT pref EMPTY >
<!ELEMENT close (#PCDATA) >
<!-- Attributes -->
<!-- ELEMENTS NAME VALUE DEFAULT -->
<!ATTLIST p id ID #IMPLIED >
<!ATTLIST pref refid IDREF #REQUIRED >
<!ATTLIST memo status (confiden | public) "public"
          date CDATA #IMPLIED >
```



ID is a set of unique labels; IDREF must be a label declared in a document with all attributes of ID type. Thus ID/IDREF provides functionality similar to primary/foreign key in databases.

Schemas

```
<?xml version='1.0' encoding='utf-8' ?>
<!DOCTYPE Memo SYSTEM "./memo.dtd">
<memo>
<to>Comrade Napoleon</to>
<from>Snowball</from>
<body>
<p>
If pro is the opposite of con, and progress
is moving forward, what is congress?
See paragraph <pref refid='congr.def'> for
an answer </p>
<!-- ... omitted p elements -->
<p id='congr.def'>According to Professor Squealer
<q>Con In Latin means together
and Gradi means to walk. Congradi literally
means meeting. This became congressus
and then Congress in English.</q>
BTW recontre is the verb to meet in French...
</p>
</body>
<close>Comrade Snowball</close>
</memo>
```



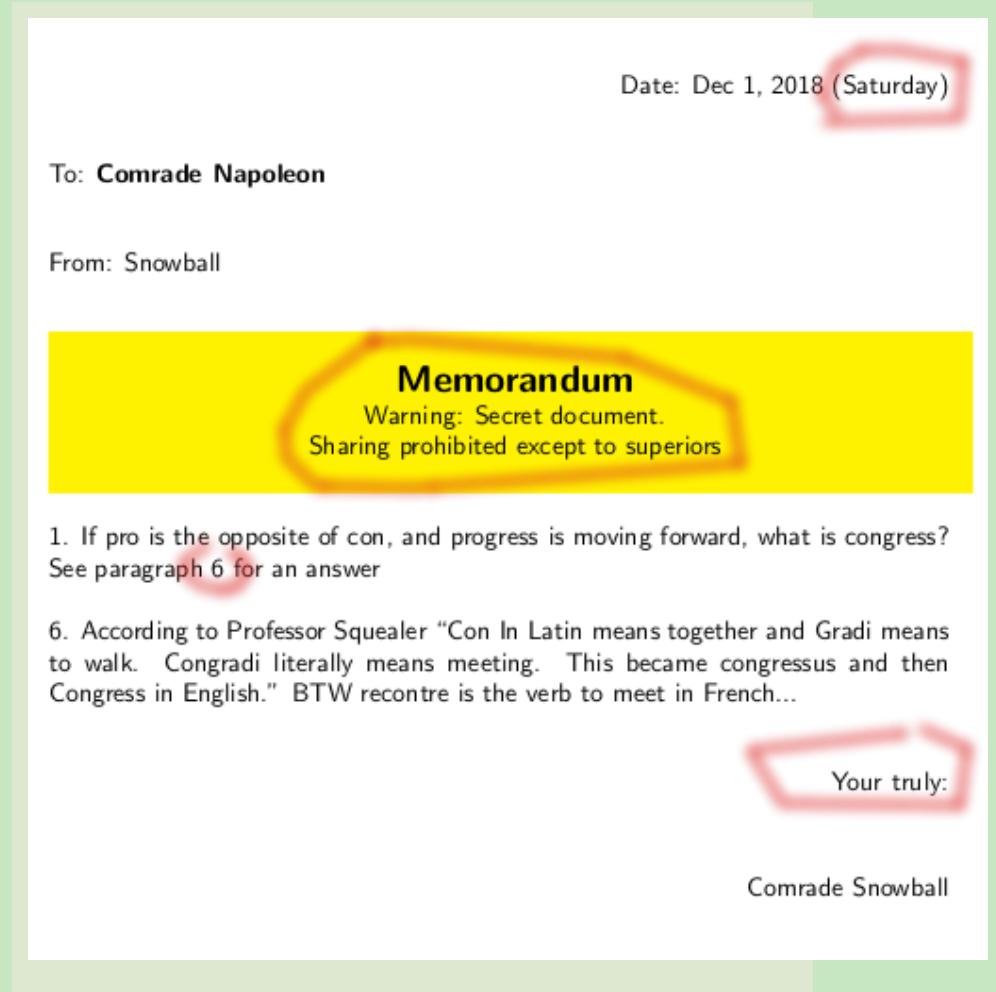
Transformation (aka formatting)

Transformation: (OO manual again) Beyond this, styles can provide a lot more than just consistency and ease of administration. They permit the creation of an automatic table of contents, numbered chapters, the setting of colors, insertion of symbols, and more

Stylization: ie declaring/defining visual appearance of content—colors, fonts, margins, skips/spaces etc...

Fonts: family, variant (italic, caps and smallcaps), weight (bold), size. Font technologies OTF (Unicode font)

<https://fonts.googleapis.com/>
<https://fonts.googleapis.com/specimen/Lato>
<https://fontforge.github.io/en-US/>
<https://blog.fontlab.com/author/adamtwr39258/>



Memo = transformed (text added marked with red circles) then stylized (boldface, fontsize etc...)

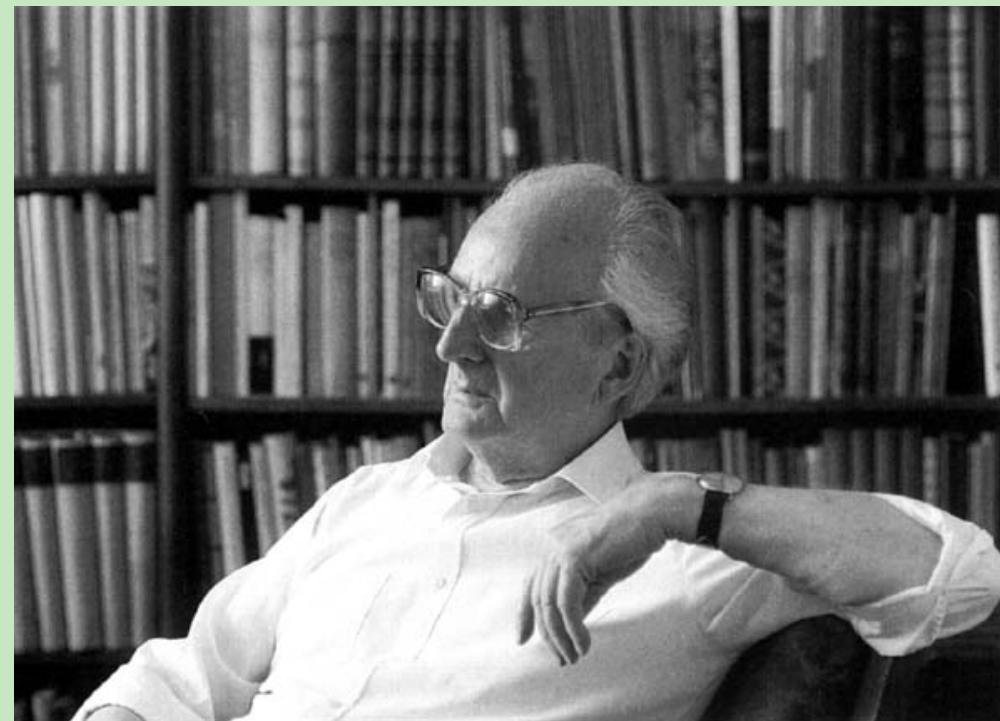
Sideways story: Polish fonts

Pre-computer era: Adam Półtawski / Zygfryd Gardzielewski

Antykwa Półtawskiego; Antykwa Toruńska

<https://ctan.org/tex-archive/fonts/iwona>
c:/windows/fonts

AĄBCĆDEĘFGHIJKLŁMNŃ
OÓPQRSŚTUVWXYZŹŻ
aąbcćdeęfgijklłmnńoópq
rsśtuvwxyzźż 1234567890



ABCDEF^{GHIJKLMN}
abcdefghijklmnopqrstuvwxyz
1234567890 (!@#\$%)

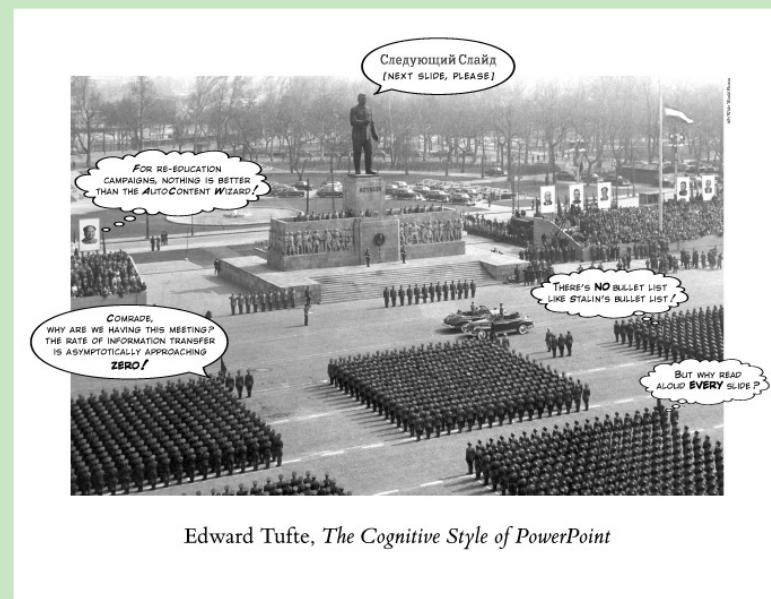
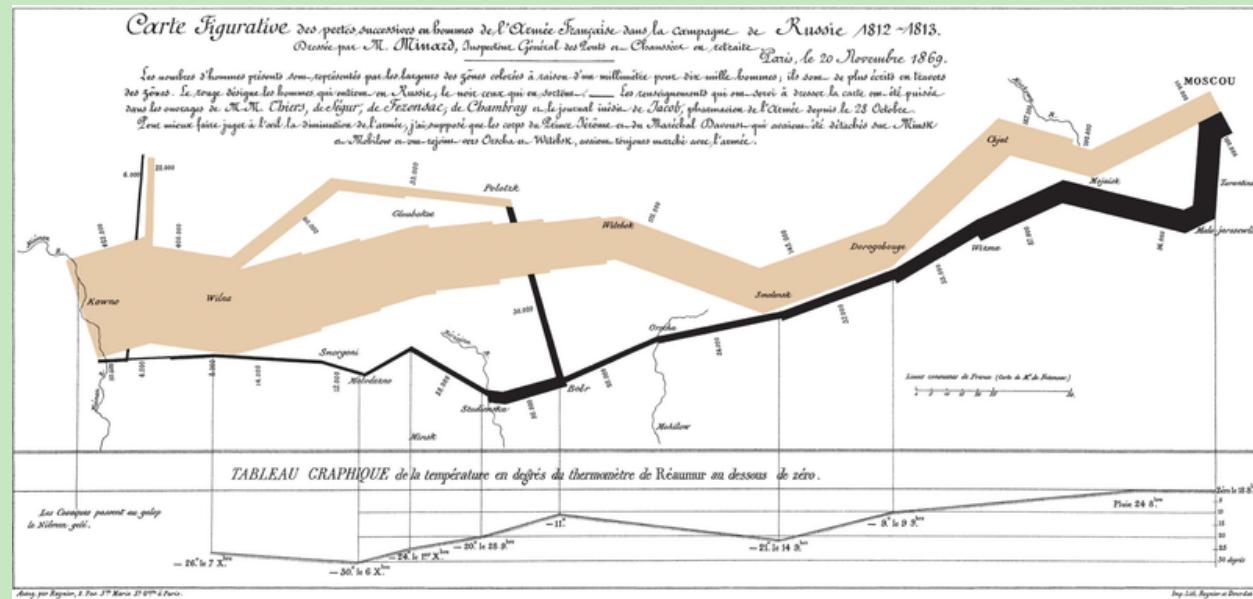
ABCDEF^{GHIJKLMN}
abcdefghijklmnopqrstuvwxyz
1234567890 (!@#\$%)

Sideways story: PP Presentations and such

Huston we have a problem [with publishing/presenting information NOT text]

Information visualization, information design principles

Edward Tufte <https://www.edwardtufte.com/tufte/>; (see also Tufte's The Cognitive Style of PowerPoint https://www.inf.ed.ac.uk/teaching/courses/pi/2016_2017/phil/tufte-powerpoint.pdf)



Edward Tufte, *The Cognitive Style of PowerPoint*

Trends: B-W, Color, Dynamic/animated

Tools: Excel (awful, google: excel chart junk pdf), R, googlecharts, shiny, D3

https://developers.google.com/chart/interactive/docs/quick_start

<https://plot.ly/>

storage (permanent)

Permanent storage = document databases (noSQL) vs relational databases

Schema change

<https://aws.amazon.com/nosql/document/>

A document database is a type of nonrelational database that is designed to store semistructured data as documents. [...] A document database is a great choice for content management applications such as blogs and video platforms.

<https://firebase.google.com/docs/firestore/quickstart>

Search Full text search vs querying structured documents; Metadata; RDF; Semantic Web

Tim Berners-Lee, James Hendler and Ora Lassila *The Semantic Web* Scientific American 2001.

<https://scholar.google.com/citations?user=udmYLRAAAAAJ>

Project

Prepare slides using MarkDown, github and pandoc

<https://github.com/>

<https://pandoc.org/demos.html>

Minimal example:

title: Test

...

Test!

This is a test of *pandoc*.

- list one

- list two

conversion

pandoc.exe test1.md -f markdown -t html -s -o test1.html

pandoc.exe" test1.md -f markdown -t odt -s -o test1.odt

The End/Thank you