OFFICE SYSTEMS KEY PROBLEMS and some TOOLS

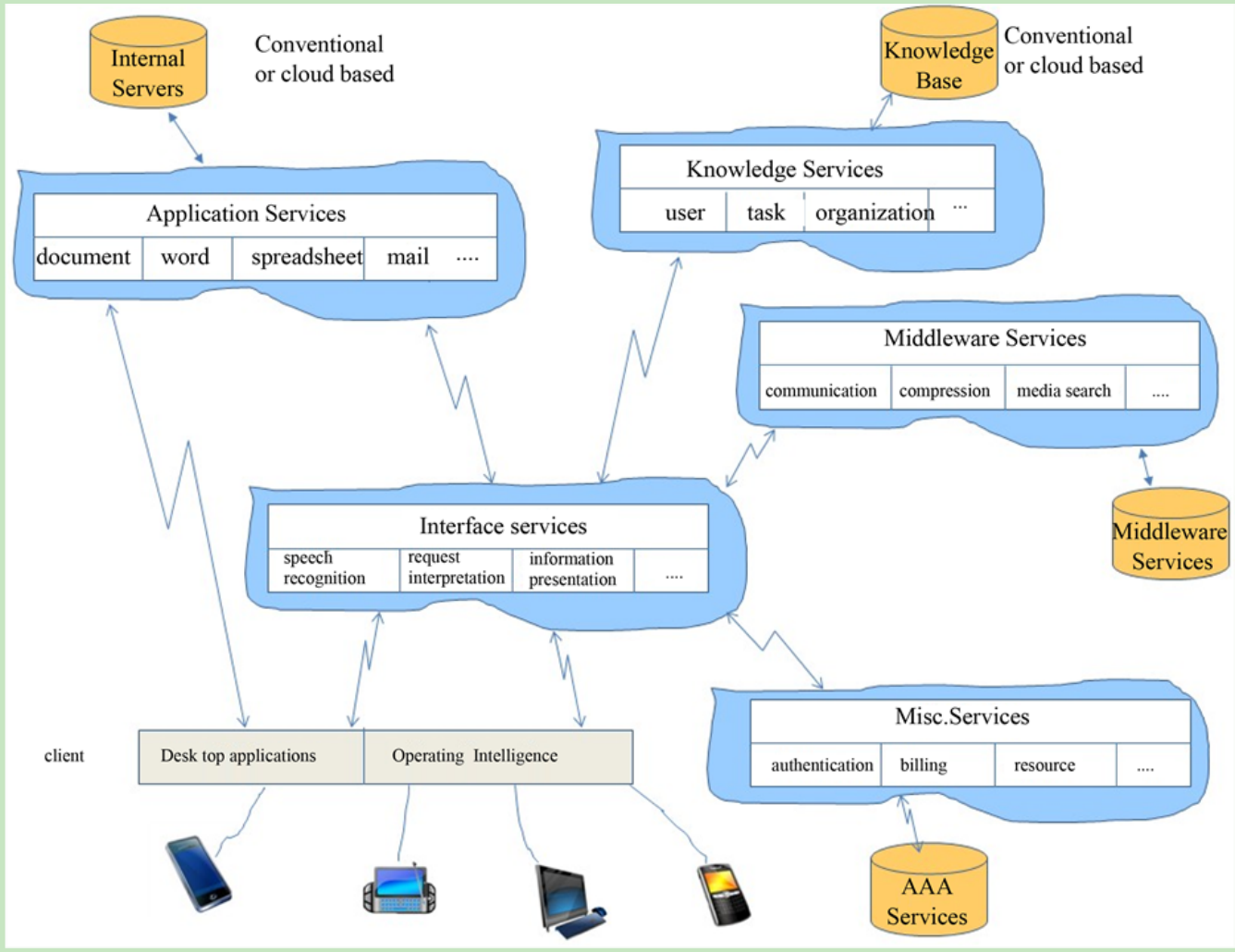
Tomasz Przechlewski

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
ArmsChandra 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 :-))

Formats (aka notations; open vs proprietary)

`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

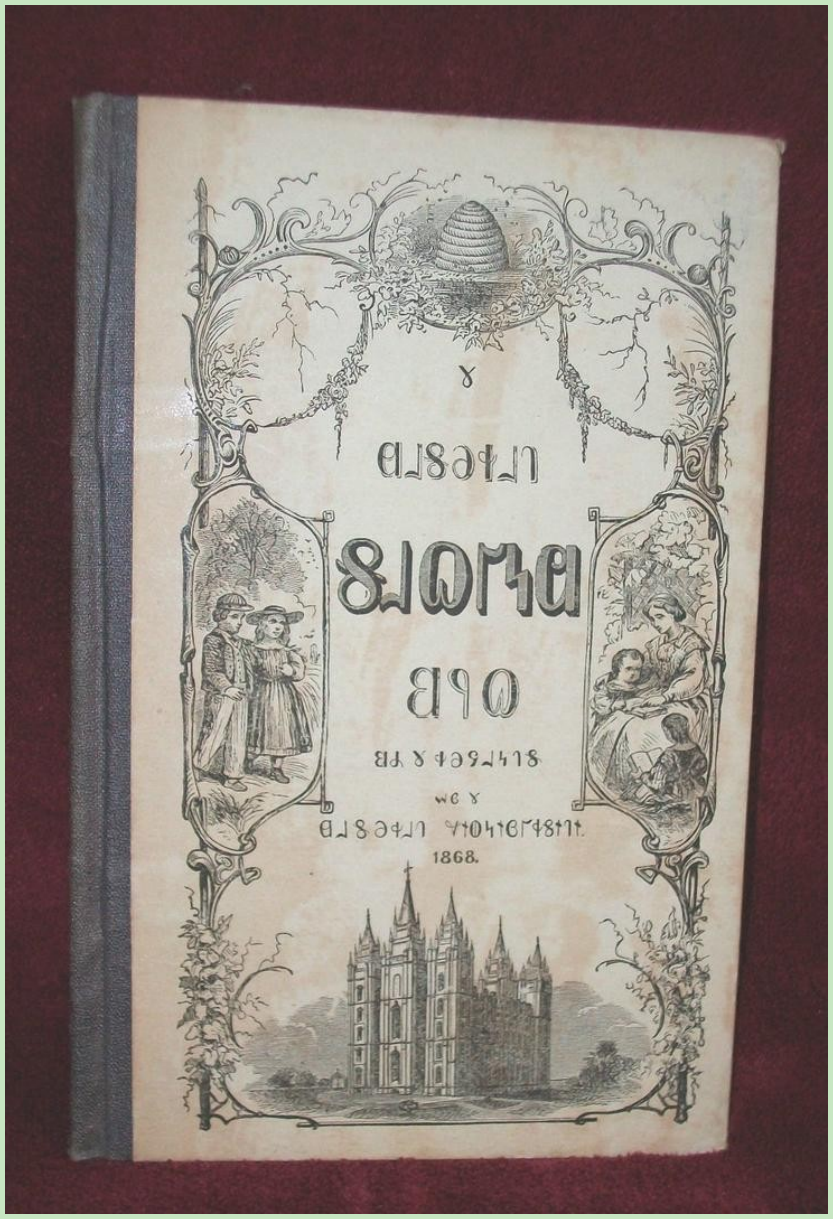
Tools [you know]...

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.he
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

Half-way: LaTeX vs LyX;

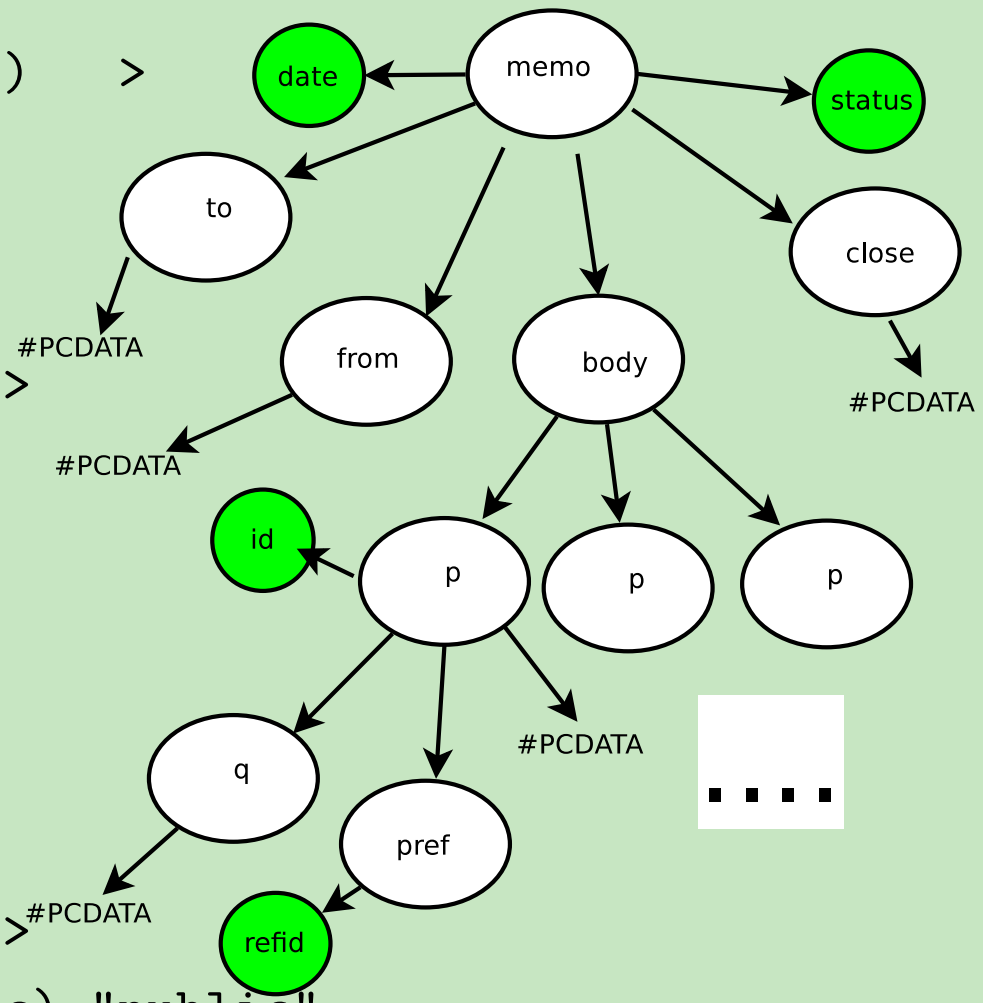
XML validators (aka parsers): xmllint, WWW browsers

XSLT transformers: WWW browsers

Publishing: Adobe Acrobat, WWW browser

Schemas

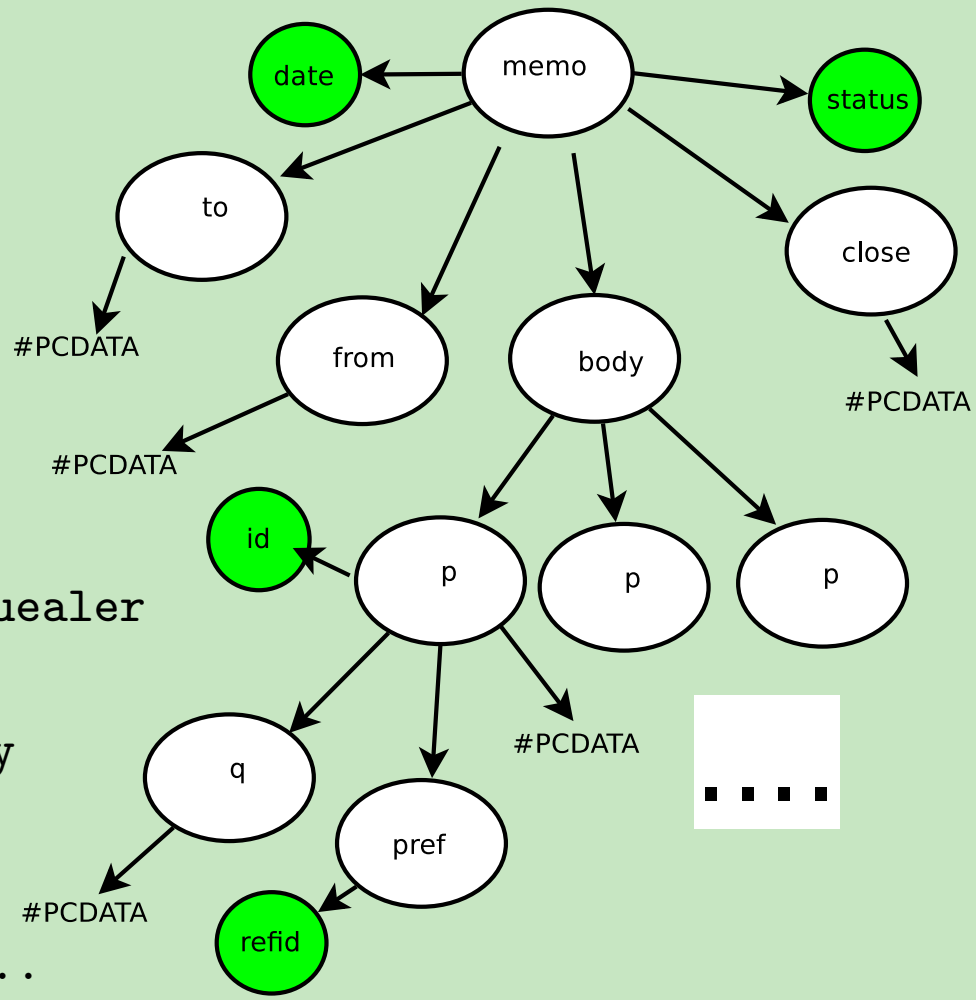
```
<!-- DTD for simple office memoranda -->
<!-- Elements -->
<!ELEMENT memo   (to, from, body, close?)    >
<!ELEMENT to     (#PCDATA) >
<!ELEMENT from   (#PCDATA) >
<!ELEMENT body   (p*) >
<!ELEMENT p      (#PCDATA | q | pref )* >
<!ELEMENT q      (#PCDATA) >
<!ELEMENT pref   EMPTY >
<!ELEMENT close (#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 declard in a document with all attributes of ID type. Thus ID/IDREF procides functionality similar to primary/foreign key in databases.

Schemas

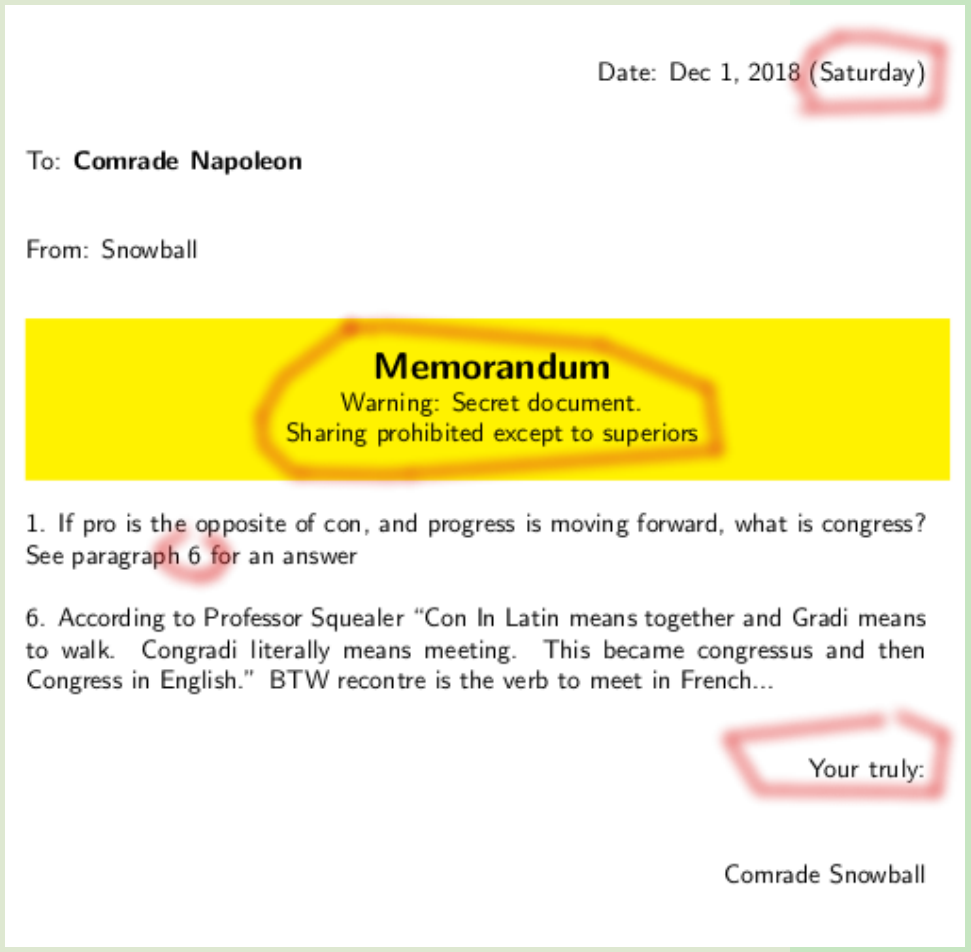
```
<?xml version='1.0' encoding='utf-8' ?>
<!DOCTYPE Memo SYSTEM "./memo.dtd">
<memo>
<to>Comrade Napoleon</to>
<from>Snowball</from>
<body>
<p>
If pro is the opposite of con, and progress
is moving forward, what is congress?
See paragraph <pref refid='congr.def'/> for
an answer </p>
<!-- ... omitted 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 appearence of content–colors, fonts, margins, skips/spaces etc... Fonts: family, variant (italic, caps and smallcaps), weight (bold), size. Font technologies OTF (Unicode font)



Date: Dec 1, 2018 (Saturday)

To: **Comrade Napoleon**

From: Snowball

Memorandum
Warning: Secret document.
Sharing prohibited except to superiors

1. If pro is the opposite of con, and progress is moving forward, what is congress? See paragraph 6 for an answer

6. According to Professor Squealer "Con In Latin means together and Gradi means to walk. Congradi literally means meeting. This became congressus and then Congress in English." BTW recontre is the verb to meet in French...

Your truly:

Comrade Snowball

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

`https://fonts.google.com/`
`https://fonts.google.com/specimen/Lato`
`https://fontforge.github.io/en-US/`
`https://blog.fontlab.com/author/adamtwr39258/`

Sideways story: Polish fonts

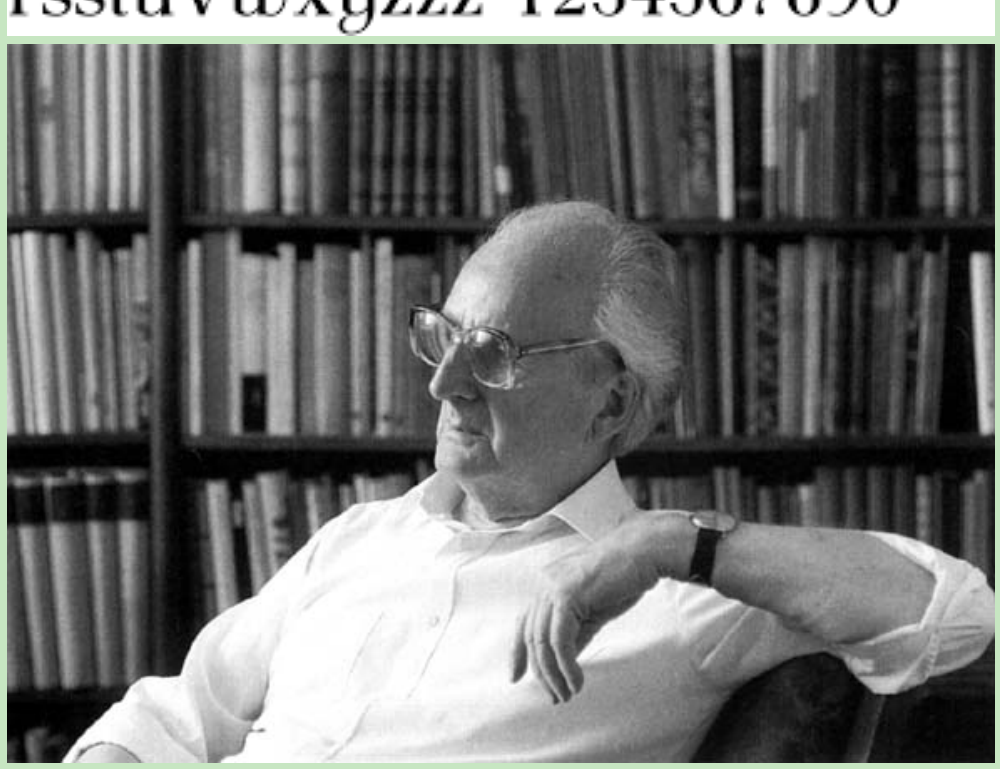
Pre-computer era: Adam Półtawski / Zygfryd Gardzielewski

Antykwa Półtawskiego; Antykwa Toruńska

`https://ctan.org/tex-archive/fonts/iwona`

`c:/windows/fonts`

AĄBCĆDEĘFGHIJKLŁMNŃ
OÓPQRSŚTUVWXYZŹŻ
aąbcćdeęfghijklłmnńoópq
rsśtuvwxyzźż 1234567890



ABCDEFGHIJKLMN
abcdefghijklmnopqrst
1234567890 (!@#$%

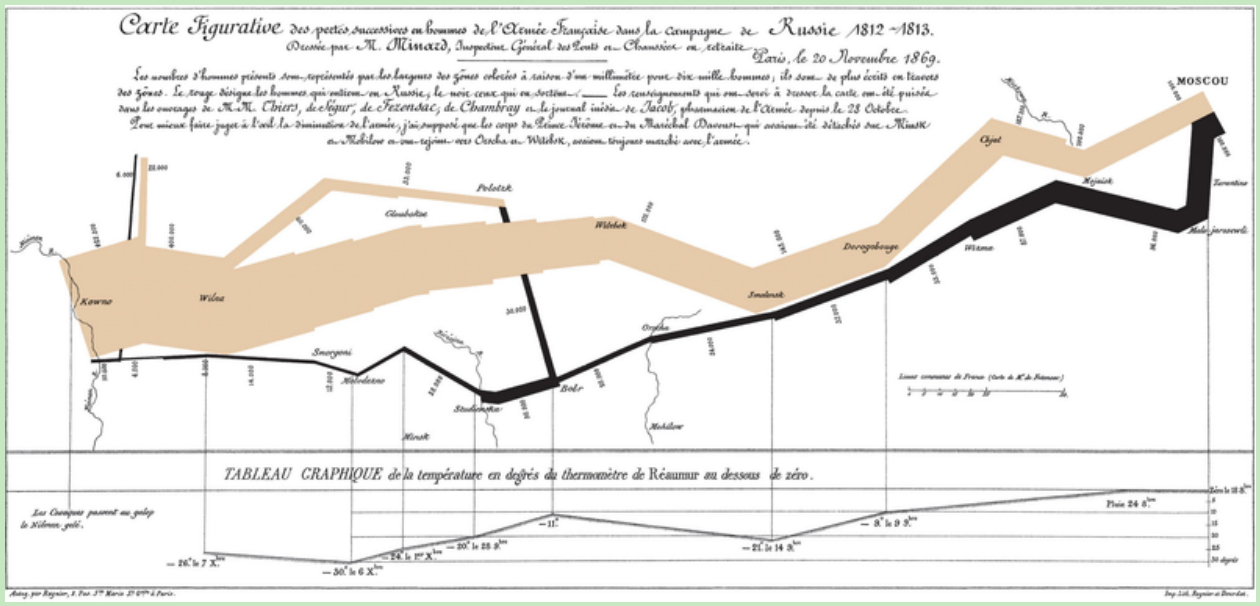
ABCDEFGHIJKLMN
abcdefghijklmnopqrst
1234567890 (!@#$%

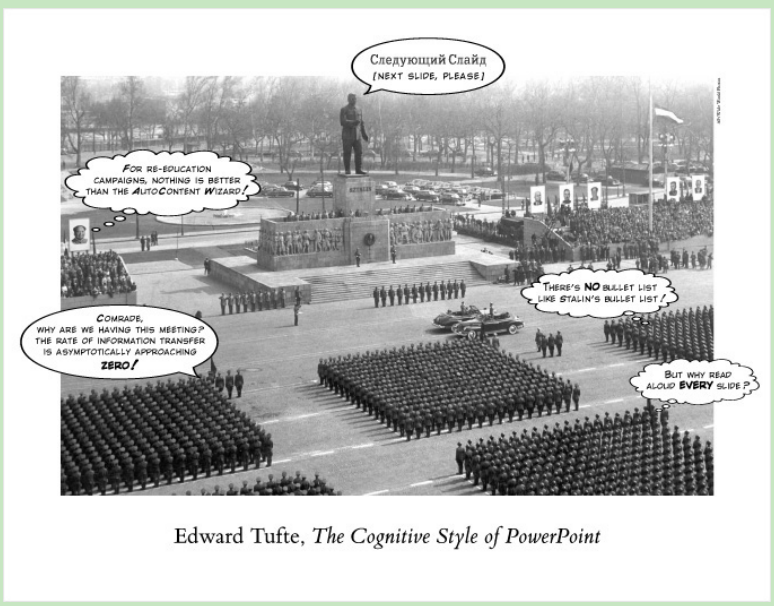
Sideways story: PP Presentations and such

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 (premanent)

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