Document Redaction with LibreOffice

By Muhammet Kara
Consultant Software Engineer at Collabora Productivity
@muhamm3tkara
Muhammet Kara

Work

- Working on the core & UI
- Since Dec 2018
- Was @Pardus project before

Community

- LibreOffice certified developer
- Long-time GNOME contributor
- Member of The Document Foundation & the GNOME Foundation
Redaction?
What is Redaction, aka Sanitization?

“Redaction/Sanitization is the process of removing sensitive information from a document or other message, so that the document may be distributed to a broader audience.”
Why?

Some of the reasons

- Unclassify the classified documents
- Privacy protection & data anonymization
- Prevent information leak in collaborative environments
- ...
Who?

Some of the possible users

- Government & military
- Public institutions
- Anyone who wants to share info without disclosing personal/private information
- Actually, this feature development was ordered by South Eastern Health & Social Care Trust (SEHSCT)
Process

- Open a document in Writer/Calc
- When you are ready to redact, click the redact button/menu
- Redact the document in Draw
- Export & share with the others
Technical Process

- Convert the document into a metafile
- Spawn a new Draw instance
- Insert the metafile into Draw
- Open the Redaction toolbar
- Flatten/pixelize document while exporting
Redaction Toolbar

- Rectangle Redaction
- Free-form Redaction
- Export as Redacted PDF
- Export as Preview PDF
Select Redact
Use Redaction Toolbar
Select Redacted Export
~ Estimate ~

Business estimate
John D. Stuart
29th of May 2019

---

~ Disclaimer ~

The presented document intends to be an estimate offer for all design services on [redacted] and shall be deemed to be rejected unless the customer signs promptly informing of the need to adjust it within 7 days from dispatch.

---

~ Description ~

Road-map
1st week
Analysis and research

2nd to 3rd week
Prototype with meetings

4th week
Testing and feedback

Additional costs:
Upholt and upon completion

---

~ Estimate in £ ~

Visual identity without brand tools
Prototype: Analysis, User interface diagram and Low-fidelity prototype with intermediary meetings
Testing: User interface design
Feedback: Handover files and specs

---

46.8%
So is that all?
Automatic Redaction
Automatic Redaction?

Even more efficient redaction with:

- Allowing redaction target definition with ease
- Automatically redacting matching text
- Letting user to fine-tune and finish with manual redaction
- ...

collaboraonline.com
Process

• Open a document in Writer/Calc/Impress

• Select Tools > Auto-Redact from the menu

• Add/define targets in the Automatic Redaction dialog

• Click OK, and let LibreOffice do its magic
Features

- Allows adding targets via a friendly dialog
- Guides user with informative messages
- Three different types of targets:
  - Text
  - Regex
  - Predefined
- Allows saving and loading targets to/from JSON files
Features (continued)

- Dialog remembers last used targets
- Can be chosen to redact case-sensitive or case-insensitive
- Targets can be added/deleted/edited on the dialog
Select Auto-Redact
3 Types of Targets
### Predefined Targets

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Content</th>
<th>Case Sensitive</th>
<th>Whole Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stuart</td>
<td>Text</td>
<td>Stuart</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Save Targets
Finalize and Export as Redacted PDF
Availability

Manual Redaction

- LibreOffice 6.3

Automatic Redaction

- LibreOffice master branch ( → 6.4)

Collabora Office 6.0 and 6.2 have both Manual and Automatic Redaction, available to Collabora customers.
We will be working on rough edges as they get discovered...
Thank you for listening!

... and thank you, South Eastern Health & Social Care Trust (SEHSCT) for funding this feature development!

By Muhammet Kara
@muhamm3tkara
muhammet.kara@collabora.com