

LIBOCON19

Document Redaction with LibreOffice

By Muhammet Kara

**Consultant Software Engineer
at Collabora Productivity**

@muhamm3tkara





Muhammet Kara

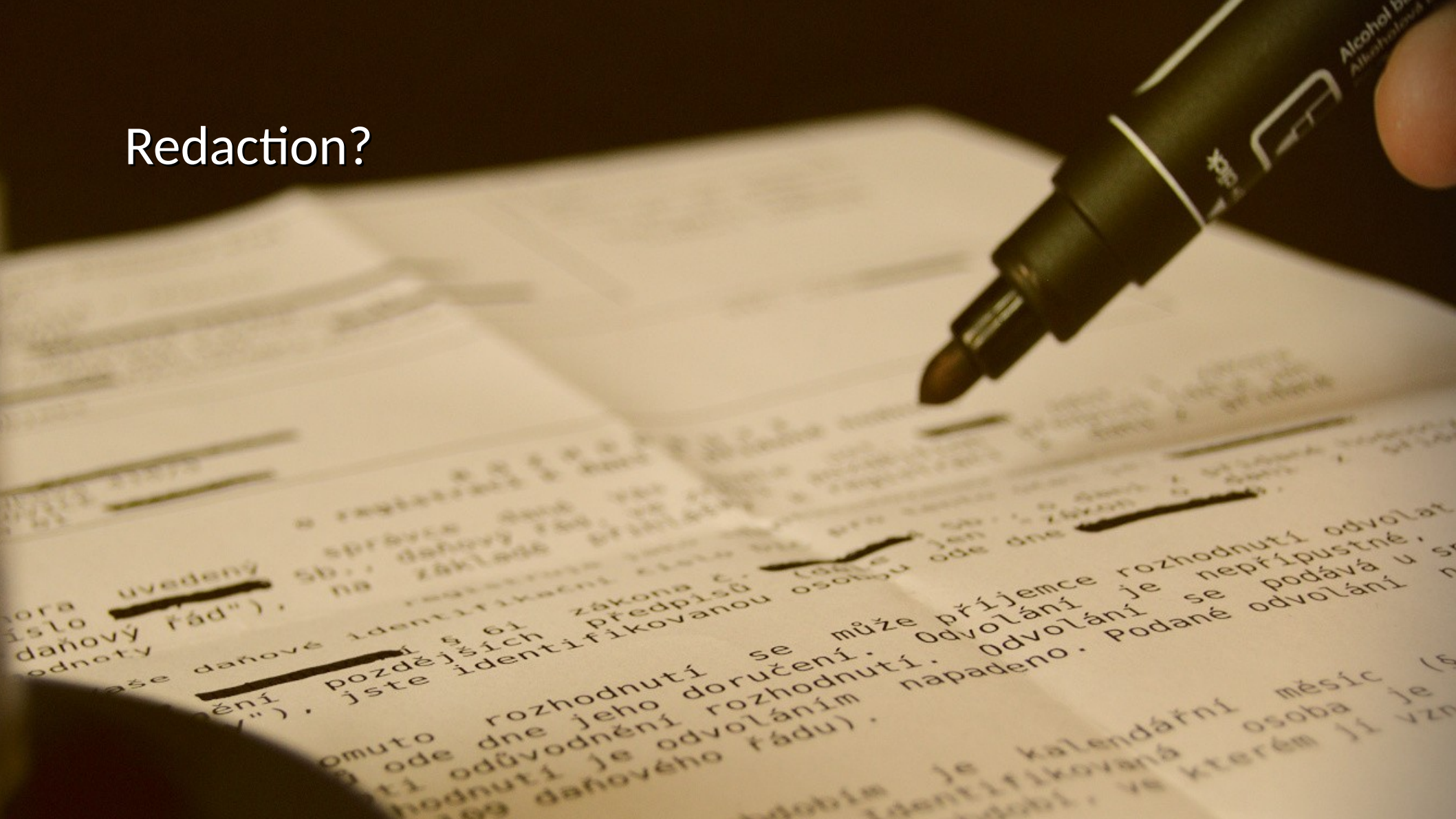
Work

- 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

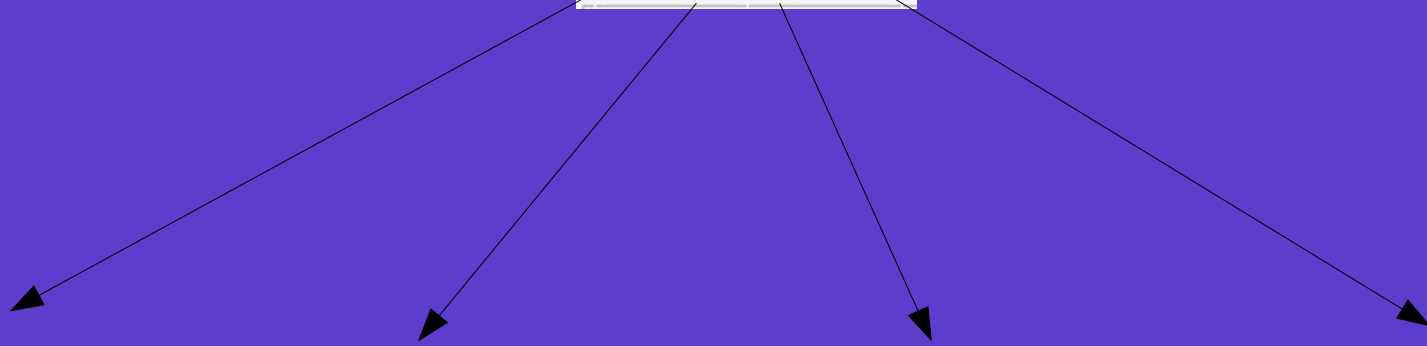
- **Open a document in Writer/Calc**
- **When you are ready to redact, click the redact button/menu**
- **Redact the document in Draw**
- **Export & share with the others**



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

The screenshot displays the Collabora OfficeDev Writer interface with the 'Tools' menu open. The 'Redact' option is highlighted in blue. The background document is a business proposal titled 'Business estimate' for John D. Stuart, dated 29th of May 2019. The document includes sections for 'Disclaimer', 'Description', and 'Road-map 1st week'. The status bar at the bottom indicates 'Page 1 of 1', '246 words, 1,564 characters', 'Default Style', 'English (USA)', and 'Outline Numbering : Level 1'.

BusinessOffer.odt - Collabora OfficeDev Writer

File Edit View Insert Format Styles Table Form Tools Window Help

Spelling... F7
Automatic Spell Checking Shift+F7
Thesaurus... Ctrl+F7
Language
Word Count
AutoCorrect
AutoText... Ctrl+F3
Redact
Chapter Numbering...
Line Numbering...
Footnotes and Endnotes...
Bibliography Database
Address Book Source...
Mail Merge Wizard...
Update
Sort...
Calculate Ctrl++
Macros
Fix Table Properties
XML Filter Settings...
Extension Manager... Ctrl+Alt+E
Customize...
Options... Alt+F12

~ Estimate ~
Business estimate
John D. Stuart
29th of May 2019
+420 723 220 522
mail@johndstuartomg.com
https://superdpbakery.com
Road-map
1st week

Disclaimer
Description

Page 1 of 1 | 246 words, 1,564 characters | Default Style | English (USA) | Outline Numbering : Level 1 | 110%

Use Redaction Toolbar

The screenshot displays the Collabra OfficeDev Draw interface. The main workspace shows a document titled "Estimate" with a disclaimer, description, and a road-map section. A redaction toolbar is visible in the top right corner, and a "Properties: Rectangle Redaction" panel is open on the right side of the screen. The toolbar includes icons for selection, redaction, and other drawing tools. The properties panel shows settings for the redaction, such as format, orientation, background, margin, and master page. The status bar at the bottom indicates the current slide (Slide 1 of 1), layout (Default), and zoom level (10.15 / -0.90).

Untitled 2 - Collabra OfficeDev Draw

File Edit View Insert Format Tools Modify Window Help

Pages

1

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 Stuart's web project: Project X and shall be deemed to be rejected unless the customer writes promptly informing of the need to adjust it within 7 days from dispatch.

Description

John D. stuart intends to promote the Project X and its language by having Donuts, an on-line showcase of all things bakery.

Project X will be, initially, curated by Sara Gomes but its essence relies on a community effort to spread the word about longwordwith and its possible applications.

The road-map will be divided in three main parts: Initial Analysis and research; Prototyping and informal meetings; and Testing. The goal is to provide the whole visual language with specifications document for bakery to come to light.

Throughout the creative process a couple of meetings will take place both to ensure that John's vision stays intact and to avoid any misunderstanding regarding the project and its application.

Road-map

1st week
Analysis and research

2nd to 3rd week
Prototyping with meetings

4th week
Testing and feedback

50% prepaid upfront and 50% upon completion

Estimate in A\$

Project briefing, analysis and requirements	free
Visual identity without brand book	123
Prototype: Analysis, User interface diagram and Low fidelity prototype with intermediary meetings	123456
Testing: User interface design	1234
Feedback: Handover files and specs	free
Total	12345 A\$

Properties: Rectangle Redaction

Format: Level

Orientation: Portrait

Background: None

Insert Image

Margin: Custom

Master Page: Default

Master Background

Master Objects

Slide 1 of 1 | Layout Controls Dimension Lines | Default | 10.15 / -0.90 | 0.00 x 0.00 | 54%

Select Redacted Export

The screenshot shows the Collabora OfficeDev Draw application interface. The main window displays a document titled "Estimate" for "Business estimate" by John D. Stuart, dated 29th of May 2019. The document content includes a disclaimer, a description of the project, a road-map, and an estimate table. A redaction shape is visible over the "50% preparation" text in the road-map section.

The Properties panel on the right is open, showing the "Redacted Export (Black)" option selected under the "Format" dropdown. Other options include "None" and "Solid". The "Shadow" and "Line" sections are also visible.

Estimate in AS	
Project briefing, analysis and requirements	free
Visual identity without brand book	123
Prototype: Analysis, User interface diagram and Low fidelity prototype with intermediary meetings	123456
Testing: User interface design	1234
Feedback: Handover files and specs	free
Total	12345 AS

Final PDF



1 of 1


Estimate-redacted.pdf

46.8%

Slide 1

~ Estimate ~
Business estimate
John D. Stuart
29th of May 2019

+420 723 220 522
mail@johndstuartmg.com
<https://suspendbakery.com>



Road-map

- 1st week
Analysis
and research
- 2nd to 3rd week
Prototype
with meetings
- 4th week
Testing
and feedback

prepaid
upfront and
upon completion

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 writes promptly informing of the need to adjust it within 7 days from dispatch.

Description

[REDACTED] intends to promote the [REDACTED] and its language by having Donuts, an on-line showcase of all things bakery.

[REDACTED] will be, initially, curated by [REDACTED] but its essence relies on a community effort to spread the word about longwordwith and its possible applications.

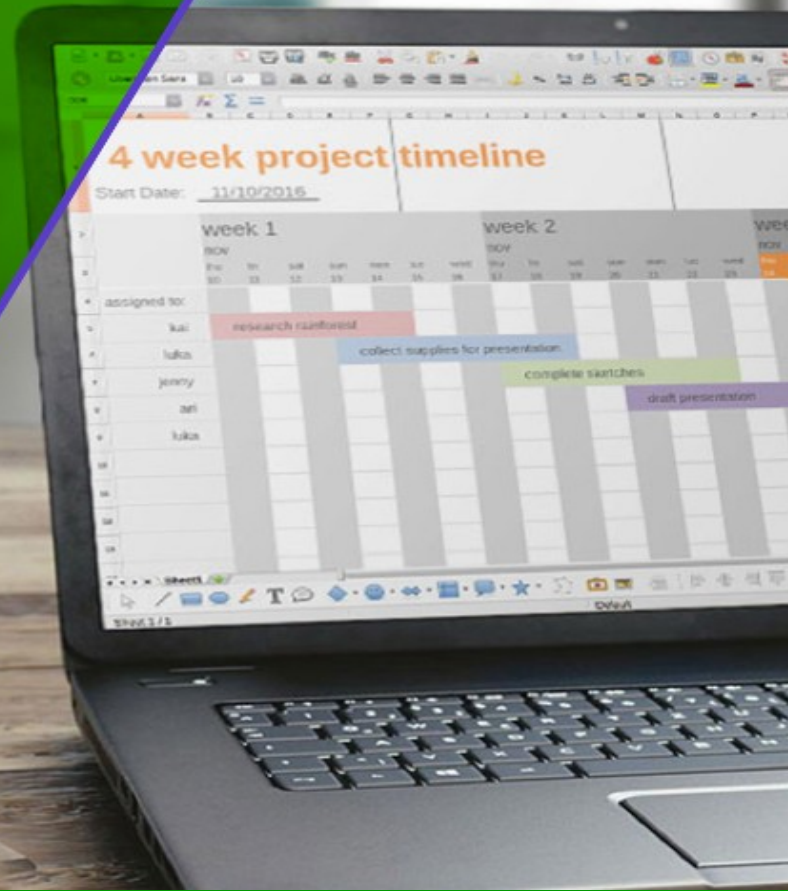
The road-map will be divided in three main parts: Initial Analysis and research, [REDACTED], and Testing. The goal is to provide the whole visual language with specifications document for bakery to come to light.

Throughout the creative process a couple of meetings will take place both to ensure that John's vision stays intact and to avoid any misunderstanding regarding the project and its application.

Estimate in AS

Project briefing, analysis and requirements	[REDACTED]
Visual identity without brand book	[REDACTED]
Prototype: Analysis, User interface diagram and Low fidelity prototype with intermediary meetings	[REDACTED]
Testing: User interface design	[REDACTED]
Feedback: Handover files and specs	[REDACTED]
Total	[REDACTED]

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
- ...



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

The screenshot displays the Collabora Office 6.0-34 Writer interface. The 'Tools' menu is open, and 'Auto-Redact' is highlighted. The document content includes a business estimate for John D. Stuart, dated 29th of May 2019. The document is titled 'BusinessOffer.odt' and contains 246 words and 1,564 characters. The status bar at the bottom shows 'Page 1 of 1', '246 words, 1,564 characters', 'Default Style', 'English (USA)', and a zoom level of 110%.

BusinessOffer.odt - Collabora Office 6.0-34 Writer

File Edit View Insert Format Styles Table Form **Tools** Window Help

Spelling... F7
Automatic Spell Checking Shift+F7
Thesaurus... Ctrl+F7
Language
Word Count
AutoCorrect
AutoText... Ctrl+F3
Redact
Auto-Redact
Chapter Numbering...
Line Numbering...
Footnotes and Endnotes...
Bibliography Database
Address Book Source...
Mail Merge Wizard...
Update
Sort...
Calculate Ctrl++
Macros
Fix Table Properties
XML Filter Settings...
Extension Manager... Ctrl+Alt+E
Customize...
Options... Alt+F12

~ Estimate ~
Business estimate
John D. Stuart
29th of May 2019
+420 723 220 522
mailto:johndstuartomg.com
https://superdjbakery.com
Road-map
1st week
Analysis

The present design service shall be deemed to be provided for all purposes and shall be promptly dispatched.
Disclaimer

John D. Stuart having done so, hereby offers to provide its language by bakery.
Description

Project X relies on a and its position out longwordwith
Initial Analysis and Testing. The goal is to provide the whole visual language with specifications document for bakery to come to light.

Throughout the creative process a couple of meetings will take place both to ensure that John's vision stays intact and to avoid any misunderstanding regarding the project and its application.

Page 1 of 1 | 246 words, 1,564 characters | Default Style | English (USA) | 110%

Add Targets

BusinessOffer.odt - Collabora Office 6.0-34 Writer

File Edit View Insert Format Styles Table Form Tools Window Help

Text Body Jotia 11 B I U X² X₂

~ Estimate ~
Business estimate
John D. Stuart
29th of May 2019

+420 723 220 522
mail@johndstuartomg.com
<https://superdpbakery.com>

Road-map
1st week
Analysis

The presented document is a design services on Stuart's deemed to be rejected unless informing of the need to act.

John D. stuart intends to provide having Donuts, an on-line service.

Project X will be, initially, completely relies on a community effort and its possible application.

The road-map will be divided into and research; Prototyping a goal is to provide the whole document for bakery to collect.

Throughout the creative process both to ensure that John's vision stays intact and to avoid any misunderstanding regarding the project and its application.

Automatic Redaction

Target Name	Type	Content	Case Sensitive	Whole Words
Name	Text	Stuart	No	Yes

Add Target

Name: ProjectNamen
Type: Text
Content: Project X
 Case Sensitive Whole Words Only

Buttons: Add, Edit, Delete, Help, Cancel, OK

Page 1 of 1 | 246 words, 1,564 characters | Default Style | English (USA) | 110%

3 Types of Targets



The screenshot shows the 'Automatic Redaction' dialog box. It contains a table of 'Redaction Targets' and an 'Add Target' sub-dialog.

Target Name	Type	Content	Case Sensitive	Whole Words
Name	Text	Stuart	No	Yes
Target 1	Prede	Credit card numbers	No	No

The 'Add Target' sub-dialog is open, showing the following fields:

- Name: Target 2
- Type: Text (dropdown menu is open showing options: Text, **Regex**, Predefined)
- Content: (empty text field)
- Case Sensitive
- Whole Words Only

Buttons in the 'Add Target' dialog: Help, Cancel, OK.

Buttons in the 'Automatic Redaction' dialog: Load Targets, Save Targets, Add, Edit, Delete, Help, OK, Cancel.

Predefined Targets



The screenshot shows the 'Automatic Redaction' dialog box. It features a table of 'Redaction Targets' with one entry: 'Name' (Type: Text, Content: Stuart, Case Sensitive: No, Whole Words: Yes). There are 'Load Targets' and 'Save Targets' buttons. An 'Add Target' sub-dialog is open, showing 'Name: Target 1', 'Type: Predefined', and 'Content: Credit card numbers'. A dropdown menu is open for 'Content', listing 'Credit card numbers', 'Email addresses', 'IP addresses', 'Dates (Numerical)', 'National Insurance Number (UK)', and 'Social Security Number (US)'. The 'Dates (Numerical)' option is highlighted. The sub-dialog also has a 'Help' button and 'OK'/'Cancel' buttons.

Target Name	Type	Content	Case Sensitive	Whole Words
Name	Text	Stuart	No	Yes

Add Target

Name: Target 1

Type: Predefined

Content: Credit card numbers

- Credit card numbers
- Email addresses
- IP addresses
- Dates (Numerical)
- National Insurance Number (UK)
- Social Security Number (US)

Buttons: Add, Edit, Delete, Help, OK, Cancel

Save Targets

BusinessOffer.odt - Collabora Office 6.0-34 Writer

File Edit View Insert Format Styles Table Form Tools Window Help

Text Body Jotia 11 B I U X² X₂

~ Estimate ~
Business estimate
John D. Stuart
29th of May 2019

+420 723 220 522
mail@johndstuartomg.com
<https://superdpbakery.com>

Road-map
1st week
Analysis

The presented document is a design services on Stuart's deemed to be rejected unless informing of the need to a

John D. stuart intends to having Donuts, an on-li

Project X will be, initiall relies on a community e and its possible applica

The road-map will be d and research; Prototypi goal is to provide the w document for bakery to

Throughout the creati both to ensure that Joh misunderstanding rega

Automatic Redaction

Redaction Target Name	Type	Content	Case Sensitive	Whole Words
Name	Text	Stuart	No	Yes
ProjectName	Text	Project X	Yes	No
CompName	Text	John D. Stuart	Yes	No

Load Targets

Save Targets

Save Targets

Name: redaction_ProjectX.json

Save in folder: [Folder Icon]

Create Folder

Places	Name	Size	Modified
Search			26-10-18
Recently Used			15-03-18
			09-03-18
			19-12-17
Desktop			
File System	redaction_ProjectX.json	682 bytes	11:30
Documenten			
Muziek			

Page 1 of 1 | 246 words, 1,564 characters | Default Style

Finalize and Export as Redacted PDF



The screenshot shows the Collabora Office 6.0-34 Draw interface. The main window displays a document titled "Untitled 1 - Collabora Office 6.0-34 Draw". The document content includes a business estimate for "Project X" dated 29th of May 2019, a road-map section, and a table of costs. A redaction tool is active, showing a dropdown menu with the option "Redacted Export (Black)" highlighted. The Properties panel on the right shows page settings: Format: Letter, Orientation: Portrait, Background: None, Margin: Custom, Master Page: Default, Master Background checked, and Master Objects checked. The status bar at the bottom indicates "Page 1 of 1", "246 words, 1,564 characters", "Default Style", "English (USA)", "25,26 / -1,39", "0,00 x 0,00", and "56%".

Redaction

- Redacted Export (Black)

Properties

Page

Format: Letter

Orientation: Portrait

Background: None

Insert Image

Margin: Custom

Master Page: Default

Master Background

Master Objects

Slide 1 of 1 | Page 1 of 1 | 246 words, 1,564 characters | Default Style | English (USA) | 25,26 / -1,39 | 0,00 x 0,00 | 56%



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...



Collabora Productivity

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