



PANGAEA.



Staats- und
Universitätsbibliothek
Bremen

gesis
Leibniz-Institut
für Sozialwissenschaften

KonsortSWD
Konsortium für die
Sozial-, Verhaltens-, Bildungs- und
Wirtschaftswissenschaften



Forschungsinstitut
Gesellschaftlicher
Zusammenhalt

DR|EX



SocioHub

EVIFA
FACHPORTAL ETHNOLOGIE

POLLUX Informationsdienst
Politikwissenschaft

akkreditiert durch
RatSWD.

Rat für Sozial- und
Wirtschaftsdaten

gefördert von
DFG Deutsche
Forschungsgemeinschaft



Bundesministerium
für Bildung
und Forschung



Das Qualiservice Anonymisierungstool für Textdaten: QualiAnon

Theresa Manderscheid

Thüringer FDM-Tage

21. Juni 2023



Universität
Bremen

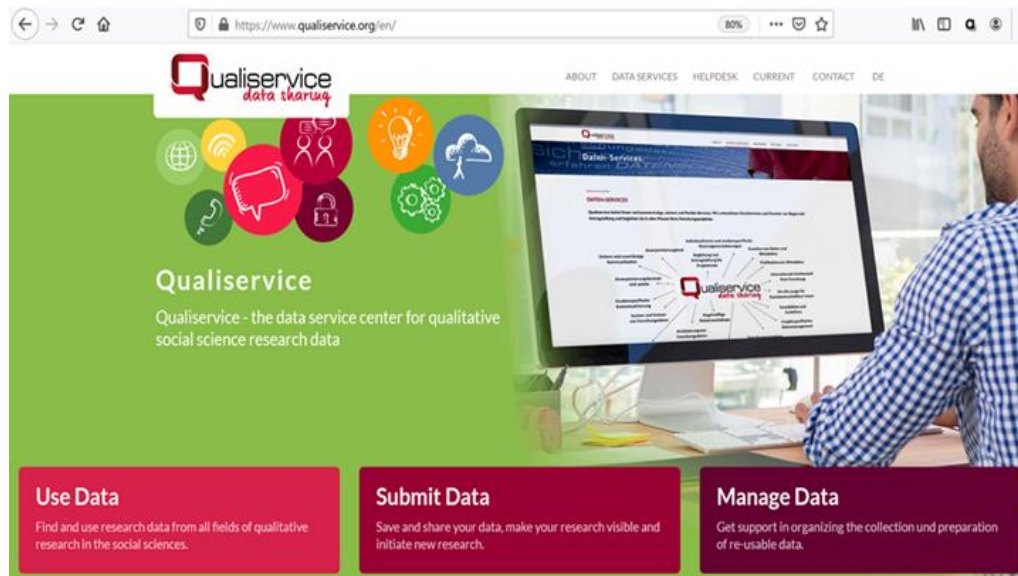
info@qualiservice.org

socium
Forschungszentrum
Ungleichheit und Sozialpolitik

Inhalte

1. Background (Qualiservice-Konzept der „Flexiblen Anonymisierung“)
2. Funktionalitäten von QualiAnon („halbautomatisches“ Tool)
3. Zugang, Manual etc.

Das Forschungsdatenzentrum Qualiservice



<https://www.qualiservice.org/>

- FDZ für qualitative Forschungsdaten der Sozialforschung (themenungebunden)
- Z.B.: Interviewtranskripte und -aufnahmen, Beobachtungsprotokolle, Feldnotizen, Fotos, audiovisuelle Daten
- Konsortium: PANGAEA, SuUB, GESIS Leibniz Institut für Sozialwissenschaften, FIDs (aktuell Sozial- und Kulturanthropologie, Pollux, SocioHub)
- Verortet an der Universität Bremen, SOCIUM
- Leitung: Prof. Dr. Betina Hollstein
- Akkreditiert durch den RatSWD, Mitglied von KonsortSWD

Besonderheiten qualitativer Forschung & qualitativer Daten

- Beinhalten oft Methodenkombinationen für Erhebung und Analyse
- Offener und iterativer Forschungsprozess, Forschungsdesign wird im Prozess angepasst
- Vertrauensverhältnis zwischen Forschenden und Beforschten
- Große Vielfalt und Dichte des Materials
- beinhaltet i. d. R. Umgang mit personenbezogenen Informationen
- hohe Kontextdichte und Verweise innerhalb der Daten (Beispiel: biographische Erzählung über Krankheitsbewältigung)
- Detaillierte Kontextinformationen als Rahmeninformationen für Datenanalyse

>> hohe datenschutzrechtliche Anforderungen und forschungsethische Standards

Herausforderungen für die Nachnutzung qualitativer Daten

- Zwei potenziell konfligierende Interessen: Persönlichkeitsrechte schützen und Forschung ermöglichen
- Anonymisierung und Pseudonymisierung als zwei der zentralen Instrumente
- Braucht Konzept, welches sozialwissenschaftlich relevante Informationen erhält ohne ‚falsche Fährten‘ zu legen
- einfache Löschungen oder Ersetzungen könnten zu falschen Ergebnissen führen

Anonymisierung/Pseudonymisierung

Das Qualiservice-Konzept der „Flexiblen Anonymisierung“

- Sensible personenbezogene Daten werden mit unterschiedlich stark abstrahierenden sozialwissenschaftlich relevanten Informationen ersetzt
- Ersetzungen können auf unterschiedlichen Abstraktionsebenen wieder geöffnet werden
- Für bestimmte Fragestellungen können so spezifischere Informationen in ‚weniger‘ anonymisierten Datensätzen bereit gestellt werden (onsite)
- Teilstandardisierung durch Listennutzung möglich



<https://nbn-resolving.org/urn:nbn:de:0168-ssoar-47605-2>

Arbeitspapier zur Konzeptentwicklung der Anonymisierungs-/Pseudonymisierung in Qualiservice
Kretzer, Susanne

Erstveröffentlichung / Primary Publication
Arbeitspapier / working paper

Empfohlene Zitierung / Suggested Citation:
Kretzer, S. (2013). Arbeitspapier zur Konzeptentwicklung der Anonymisierungs-/Pseudonymisierung in Qualiservice..
<https://nbn-resolving.org/urn:nbn:de:0168-ssoar-47605-2>



Kati Mozygamba et al.

Anonymisierung und Pseudonymisierung qualitativer textbasierter Forschungsdaten – eine Handreichung

University of Bremen – Qualiservice Working Papers

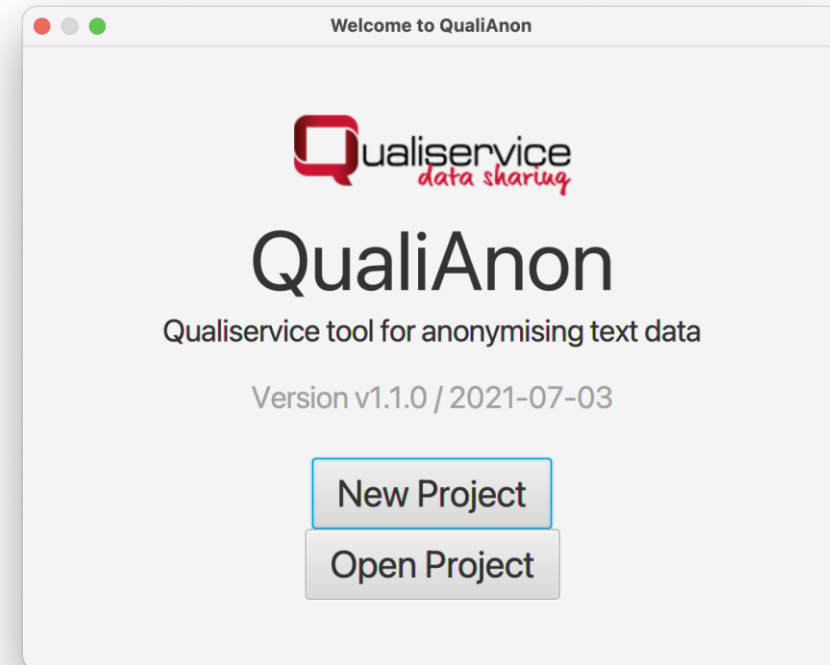
Im Erscheinen

Beispiele

Kategorie	Subkategorie	Beispiel	Replacement			
			Original	Level 1	Level 2	Level 3
Orte	Land, Region, Bundesland, Stadt, Straße etc.	Berlin	Ort	Ort 1, Stadt in Deutschland	Ort 1, große Großstadt in Deutschland	Ort 1, große Großstadt in Deutschland mit mehr als 1 Mio. Einwohner
berufliche Tätigkeiten	Ausbildung, Berufe, berufliche Tätigkeiten	Altenpfleger	Beruf des Interviewten	Beruf des Interviewten, Dienstleistungsberufe und Verkäufer, ISCO-08: Code 05	Beruf des Interviewten, Betreuungsberufe, ISCO-08 Code 53	Beruf des Interviewten, Betreuungsberufe im Gesundheitswesen ISCO-08 Code 532
Individuelle Charakteristik	Krankheiten, Preise, Unfälle...	Schwerer Herzinfarkt	Krankheit	Krankheit 1	Schwere Erkrankung 1	Schwere Erkrankung des Kreislaufsystems

QualiAnon – Tool für Textdaten

- Umsetzung des Konzeptes der „flexiblen Anonymisierung“
- Nutzung im FDZ (Kuration) und für Forschende
- Nachhaltigkeit
- User in control

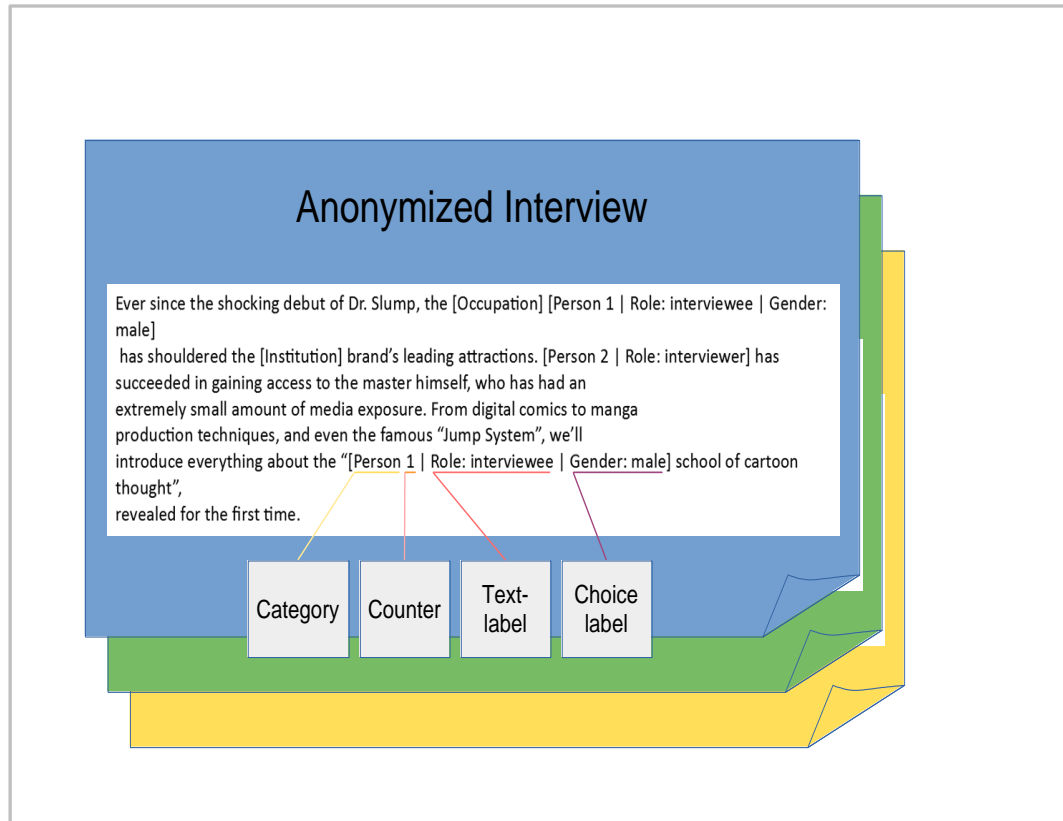


Reference: Tom Nicolai, Kati Mozygemba, Susanne Kretzer, Betina Hollstein (2021): QualiAnon – Qualiservice tool for anonymizing text data (version 1.1.0). Qualiservice. University of Bremen. Software available at: <https://github.com/pangaea-data-publisher/qualianon>

Ausgewählte Funktionen

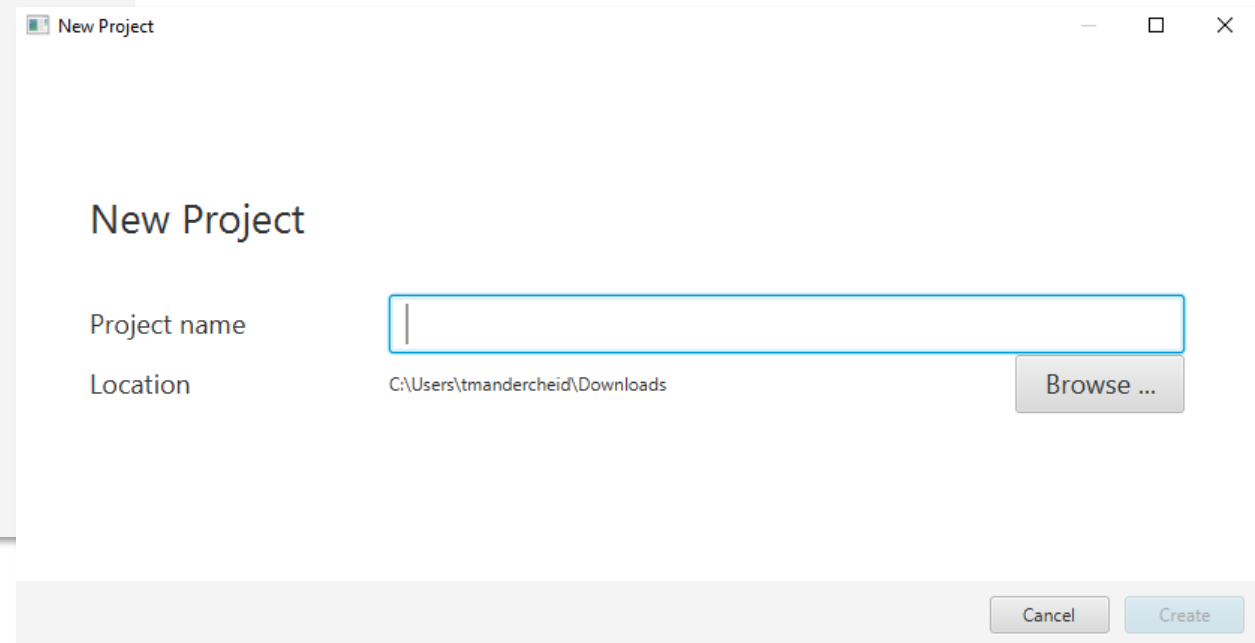
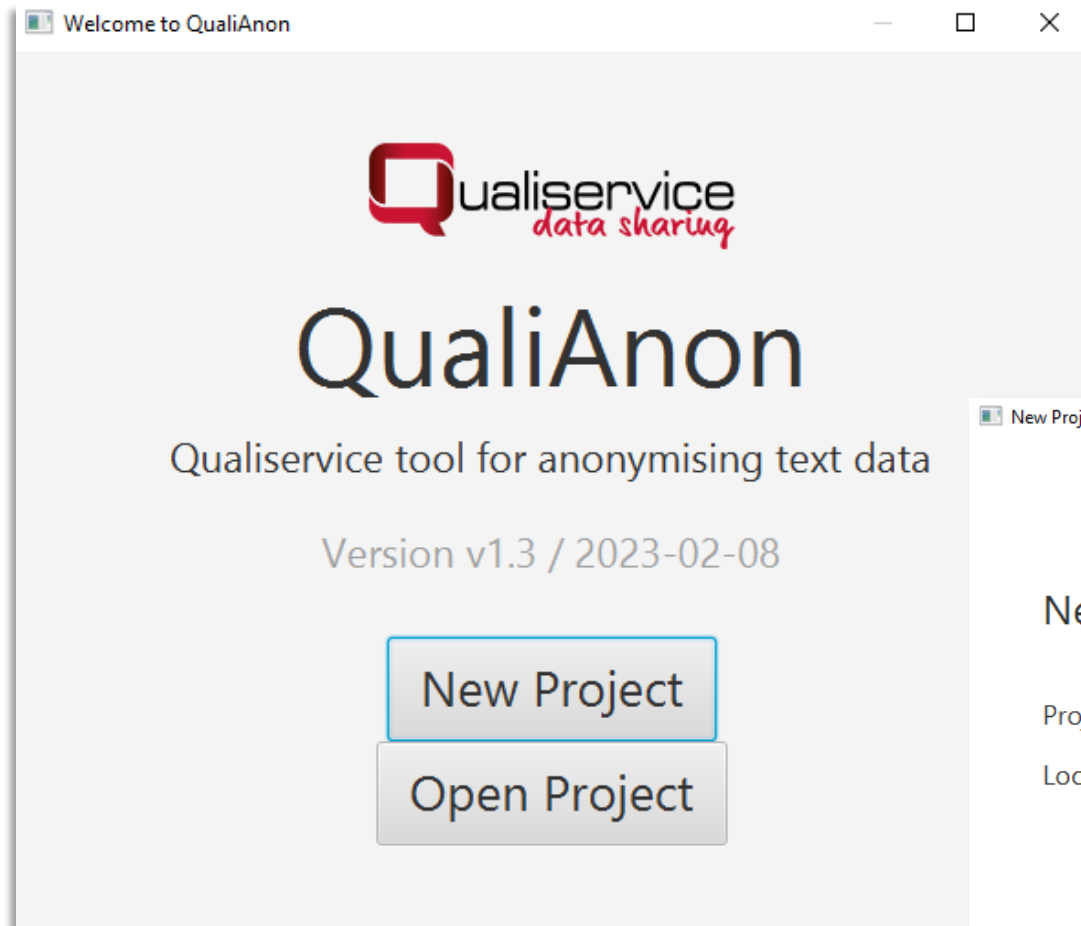
- halbautomatische Anonymisierung/Pseudonymisierung von Textdaten (user in control)
- Ersetzungen auf Studien- und Fallebene
- Markieren & Ersetzen von sensiblen Informationen auf unterschiedlichen Abstraktionslevels
- Dokumentieren der Ersetzungen, schnelle Anpassungen ermöglichen
- Unterschiedlich stark abstrahierte Exporte
- Integration von (standardisierten) Listen
- Erstellen und Teilen von themenspezifischen Ersetzungsschemata
- Getrennte Speicherung

Beispiel für abstrahierten Text



- Für jede Entität, die QualiAnon ersetzt, kann aus einem bearbeitbaren System von Ersetzungskategorien, eine Kategorie gewählt (z.B. Person, Beruf, Ort, etc.) werden.
- In eckigen Klammern findet sich die Information, die zur Ersetzung verwendet wird:
hier: „[Person 1 | Role: Interviewee | Gender: male]“.
- Nummerierung von Ersetzungskategorien
- Flexible Anonymisierung ermöglicht unterschiedlich stark abstrahierte Exporte
- Export und Weiterentwicklung der Kategorienschemata

Blick ins QualiAnon-Tool



Demo Project - QualiAnon

File Edit Window Help

Documents Replacements Categories Exports

Category / Originals	Labels	#
Occupation		
cartoonist	Level 1: Professionals Level 2: Legal, social and cu... Level 3: Creative and perfor... Level 4: Visual artists	1
Person		
Toriyama, Akira Toriyama, Akira Toriyama	Gender: Male Pseudo: Peter	33

interview-a interview-b X

1 Akira Toriyama Interview

2

3 Ever since the shocking debut of Dr. Slump, the cartoonist Akira

4 Toriyama has shouldered the Jump brand's leading attractions. Wired has

5 succeeded in gaining access to the master himself, who has had an

6 extremely small amount of media exposure. From digital comics to manga

7 production techniques, and even the famous "Jump System", we'll

8 introduce everything about the "Toriyama school of cartoon thought",

9 revealed for the first time.

10

11 Wired: It seems that the move to digital has begun, with even things

12 like comics being put on CD-ROM. Do you think that digital comics will

13 continue to increase in number from now on?

14

15 Toriyama: I think they'll probably keep increasing gradually. If we're

16 just talking about CD-ROMs, then ultimately it's almost entirely the

Log Search: Toriyama Markers: Akira Toriyama X

Person | Gender: Male, Pseudo: Peter

interview-b

- 1: Akira Toriyama Interview
- 3: Ever since the shocking debut of Dr. Slump, the cartoonist Akira
- 8: introduce everything about the "Toriyama school of cartoon thought",
- 15: Toriyama: I think they'll probably keep increasing gradually. If we're
- 27: Toriyama: Well, that in itself is obviously not a problem, as far as

Selected: 3 New replacement Reassign Remove

12 like comics being put on CD-ROM. Do you think that digital comics

13 will

14 continue to increase in number from now on?

15 Toriyama: I think they'll probably keep increasing gradually. If we're

16 just talking about CD-ROMs, then ultimately it's almost entirely the

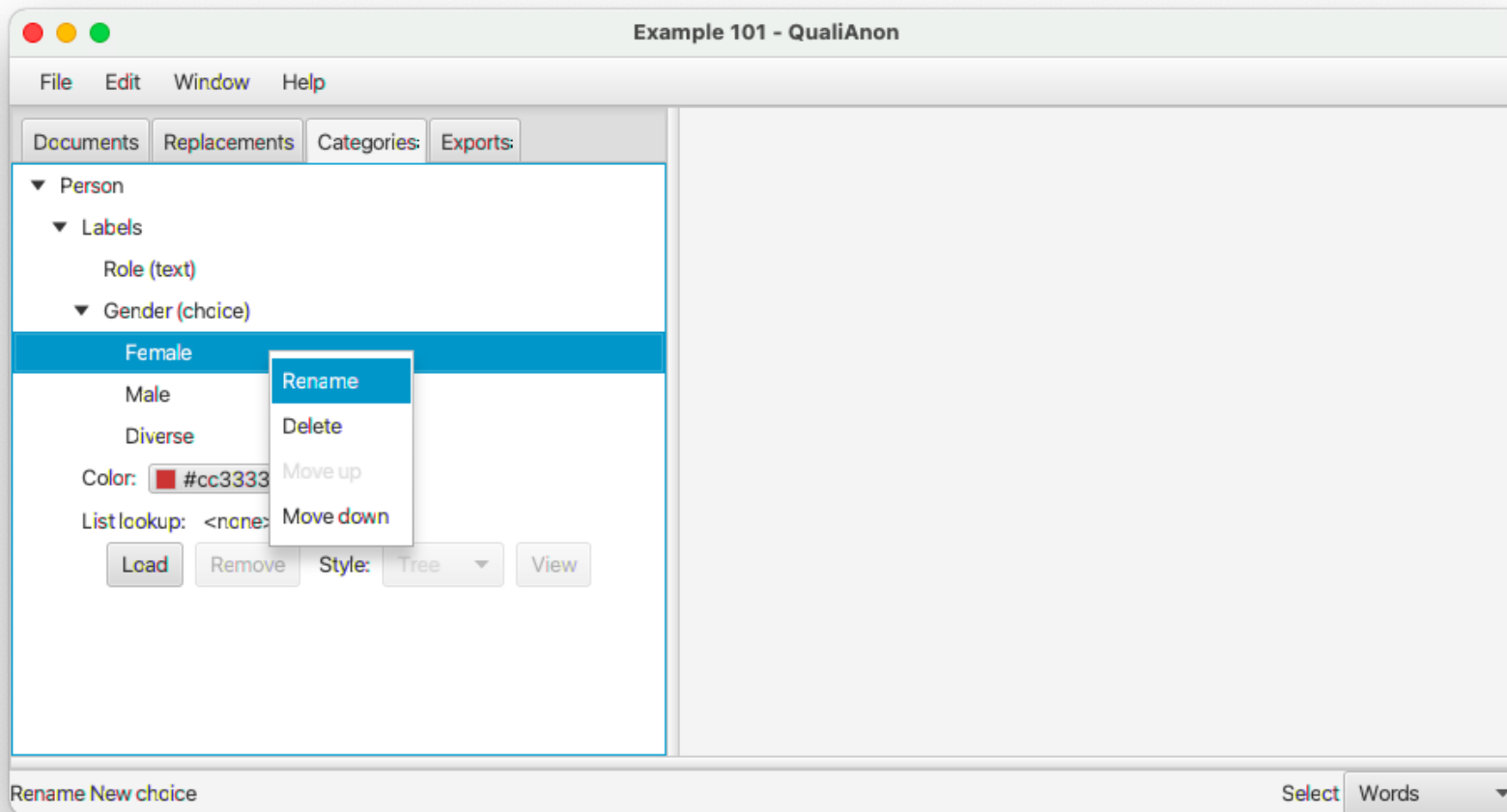
17 same work as reading comics [in print], since you're entirely free to

18 do things like go back [in the book]. CD-ROMs might even be better

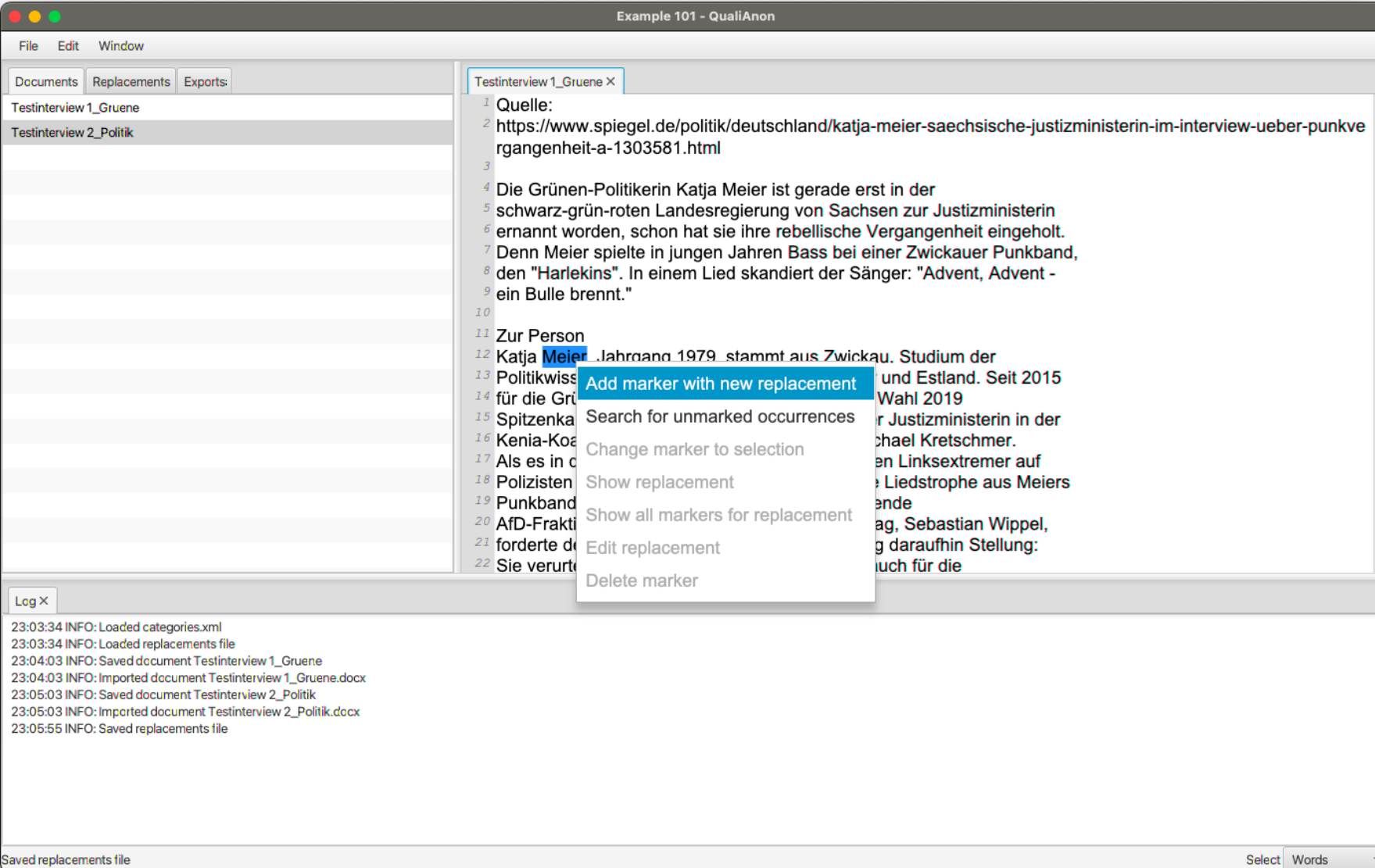
19 in

that they don't get discarded and thrown away like paper comics.

Saved document interview-b Select Words



Kategorien erstellen
und bearbeiten



The screenshot shows the QualiAnon application window titled "Example 101 - QualiAnon". The interface includes a menu bar (File, Edit, Window), a document list on the left, a main text editor, and a log window at the bottom. The text editor displays a document named "Testinterview 1_Gruene X" with the following content:

```

1 Quelle:
2 https://www.spiegel.de/politik/deutschland/katja-meier-saechsische-justizministerin-im-interview-ueber-punkve
3 rgangenheit-a-1303581.html
4
5 Die Grünen-Politikerin Katja Meier ist gerade erst in der
6 schwarz-grün-roten Landesregierung von Sachsen zur Justizministerin
7 ernannt worden, schon hat sie ihre rebellische Vergangenheit eingeholt.
8 Denn Meier spielte in jungen Jahren Bass bei einer Zwickauer Punkband,
9 den "Harlekins". In einem Lied skandiert der Sänger: "Advent, Advent -
10 ein Bulle brennt."
11
12 Zur Person
13 Katja Meier, Jahrgang 1979, stammt aus Zwickau. Studium der
14 Politikwissenschaften in Zwickau und Estland. Seit 2015
15 für die Grünen in der Bundestagswahl 2019
16 Spitzenkandidatin für die Justizministerin in der
17 Kenia-Koalition von Ministerpräsident Michael Kretschmer.
18 Als es in der Bundestagswahl 2019 um den Linksextremer auf
19 Polizisten und die Liedstrophe aus Meiers
20 Punkband "Harlekins" geht, fordert er die
21 AfD-Fraktion auf, Sebastian Wippel,
22 forderte die Partei daraufhin Stellung:
    
```

A context menu is open over the word "Meier" on line 13, with the following options:

- Add marker with new replacement
- Search for unmarked occurrences
- Change marker to selection
- Show replacement
- Show all markers for replacement
- Edit replacement
- Delete marker

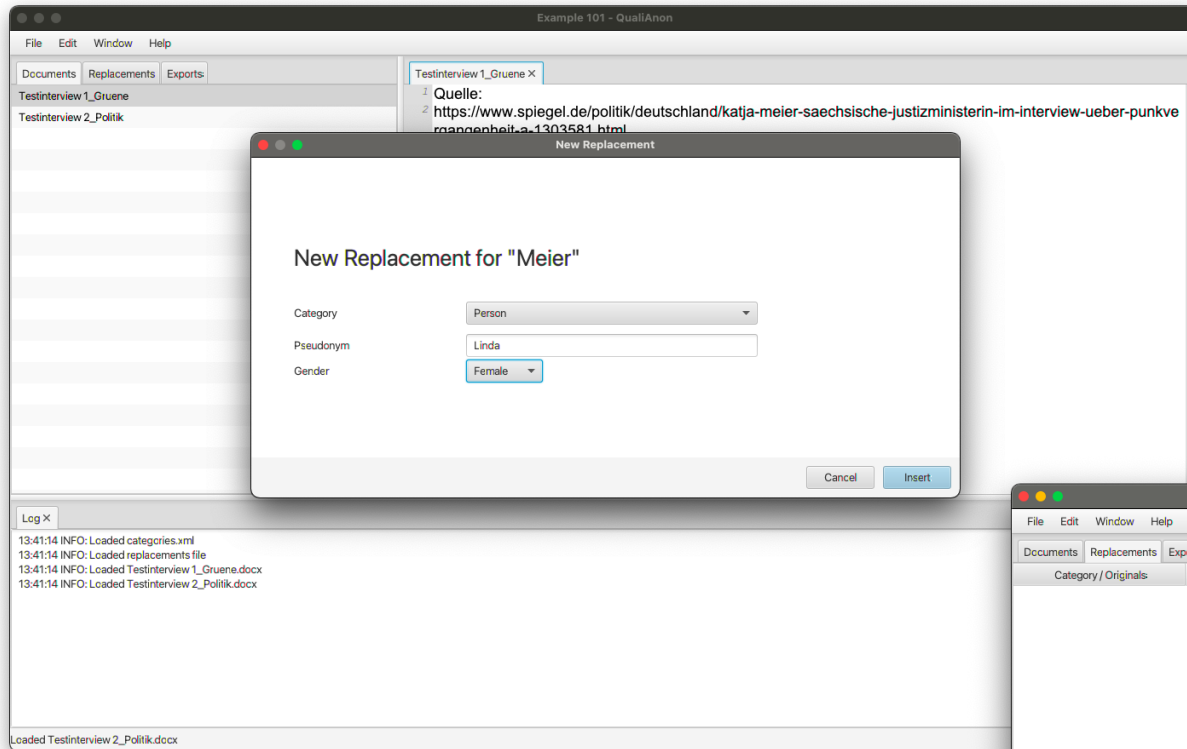
The log window at the bottom shows the following entries:

```

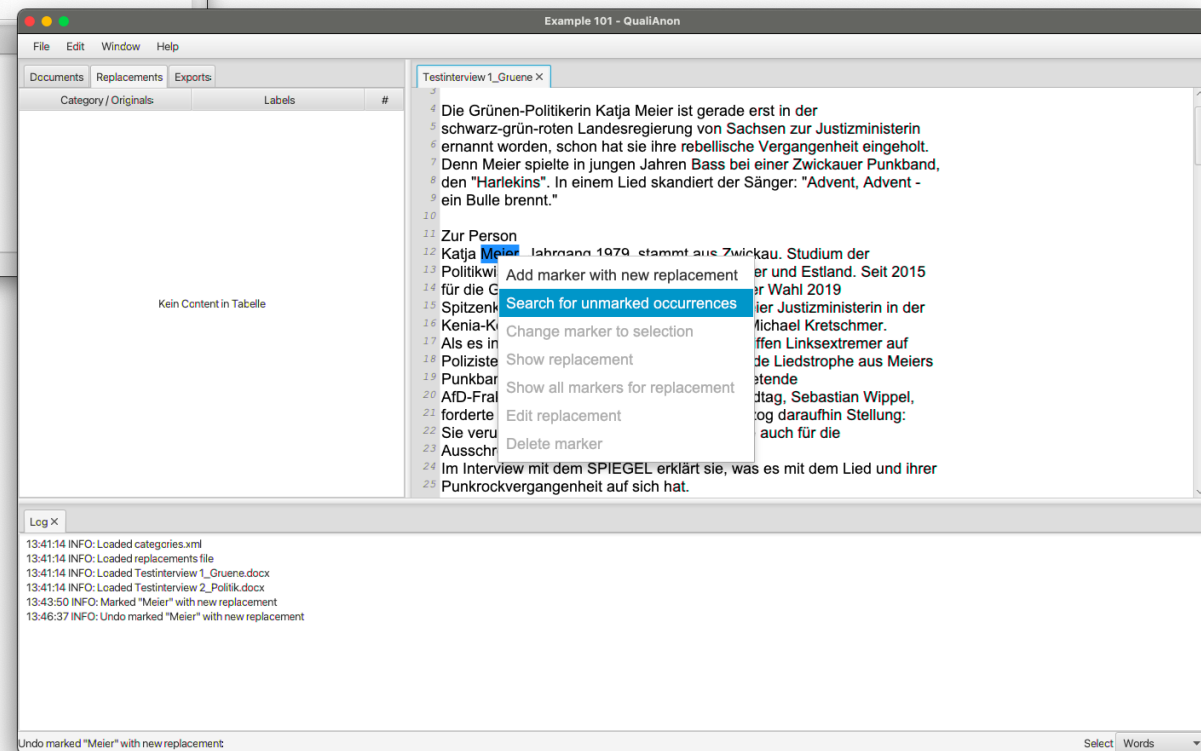
23:03:34 INFO: Loaded categories.xml
23:03:34 INFO: Loaded replacements file
23:04:03 INFO: Saved document Testinterview 1_Gruene
23:04:03 INFO: Imported document Testinterview 1_Gruene.docx
23:05:03 INFO: Saved document Testinterview 2_Politik
23:05:03 INFO: Imported document Testinterview 2_Politik.docx
23:05:55 INFO: Saved replacements file
    
```

At the bottom right of the application, there are buttons for "Select" and "Words".

Markieren einer Textstelle und neue Ersetzung erstellen

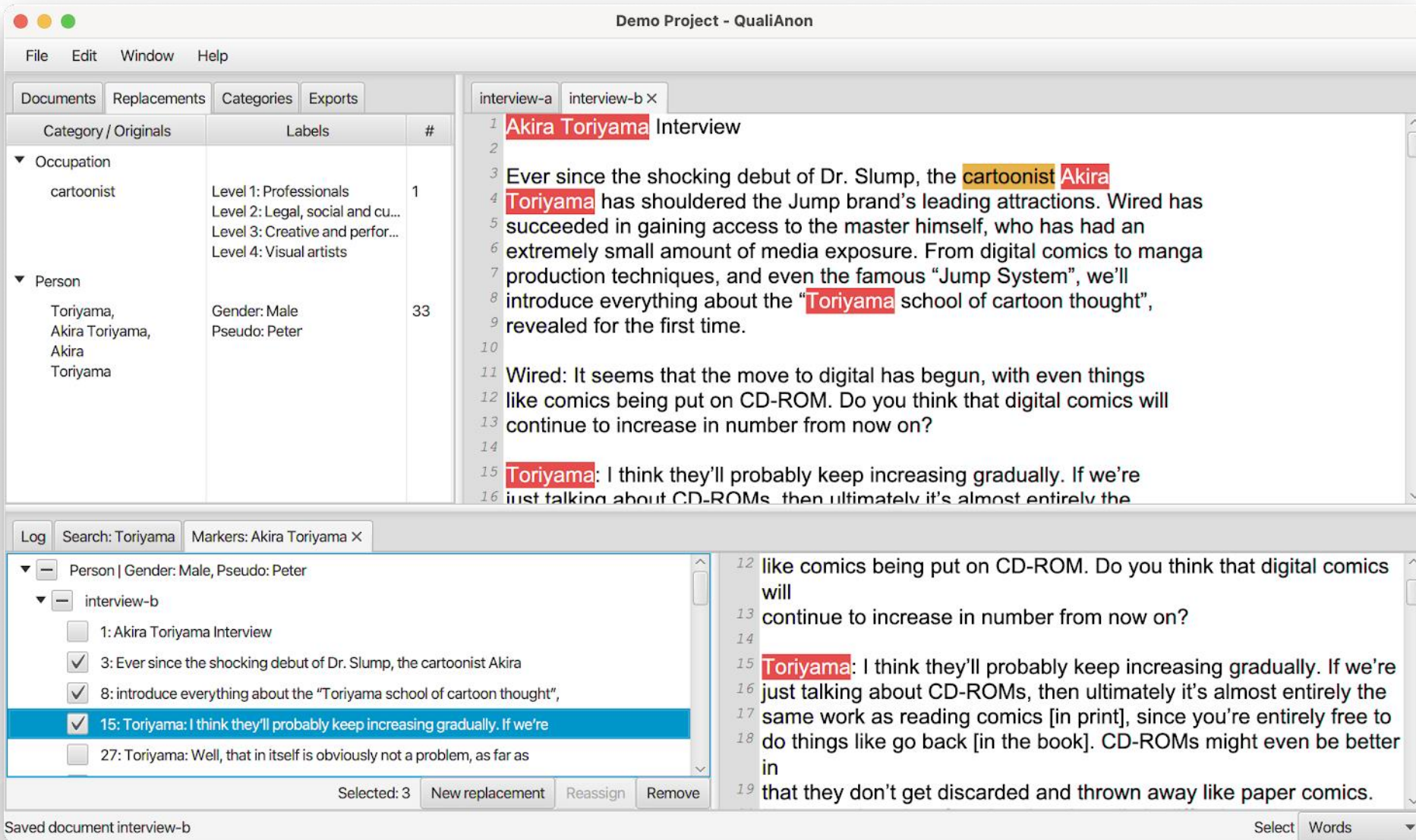


Nach weiteren
Vorkommnissen suchen



Ersetzung wählen

Ansicht der Ersetzungen



The screenshot shows the QualiAnon software interface. The main window displays a document titled "interview-b" with several paragraphs of text. The text contains several instances of the name "Akira Toriyama" and "Toriyama", which are highlighted in red. The interface includes a menu bar (File, Edit, Window, Help), a toolbar (Documents, Replacements, Categories, Exports), and a sidebar with a tree view showing the document structure. The tree view shows a "Person" category with "Toriyama, Akira Toriyama, Akira Toriyama" listed, and a "Person | Gender: Male, Pseudo: Peter" category with "interview-b" listed. The "interview-b" category is expanded, showing a list of segments with checkboxes. The segment "15: Toriyama: I think they'll probably keep increasing gradually. If we're" is selected. The main text area shows the text of the document, with the selected segment highlighted in blue. The status bar at the bottom indicates "Saved document interview-b" and "Selected: 3".

Category / Originals	Labels	#
Occupation		
cartoonist	Level 1: Professionals Level 2: Legal, social and cu... Level 3: Creative and perfor... Level 4: Visual artists	1
Person		
Toriyama, Akira Toriyama, Akira Toriyama	Gender: Male Pseudo: Peter	33

interview-a interview-b X

1 Akira Toriyama Interview

2

3 Ever since the shocking debut of Dr. Slump, the cartoonist Akira

4 Toriyama has shouldered the Jump brand's leading attractions. Wired has

5 succeeded in gaining access to the master himself, who has had an

6 extremely small amount of media exposure. From digital comics to manga

7 production techniques, and even the famous "Jump System", we'll

8 introduce everything about the "Toriyama school of cartoon thought",

9 revealed for the first time.

10

11 Wired: It seems that the move to digital has begun, with even things

12 like comics being put on CD-ROM. Do you think that digital comics will

13 continue to increase in number from now on?

14

15 Toriyama: I think they'll probably keep increasing gradually. If we're

16 just talking about CD-ROMs, then ultimately it's almost entirely the

17 like comics being put on CD-ROM. Do you think that digital comics

18 will

19 continue to increase in number from now on?

20

21 Toriyama: I think they'll probably keep increasing gradually. If we're

22 just talking about CD-ROMs, then ultimately it's almost entirely the

23 same work as reading comics [in print], since you're entirely free to

24 do things like go back [in the book]. CD-ROMs might even be better

25 in

26 that they don't get discarded and thrown away like paper comics.

27

Log Search: Toriyama Markers: Akira Toriyama X

Person | Gender: Male, Pseudo: Peter

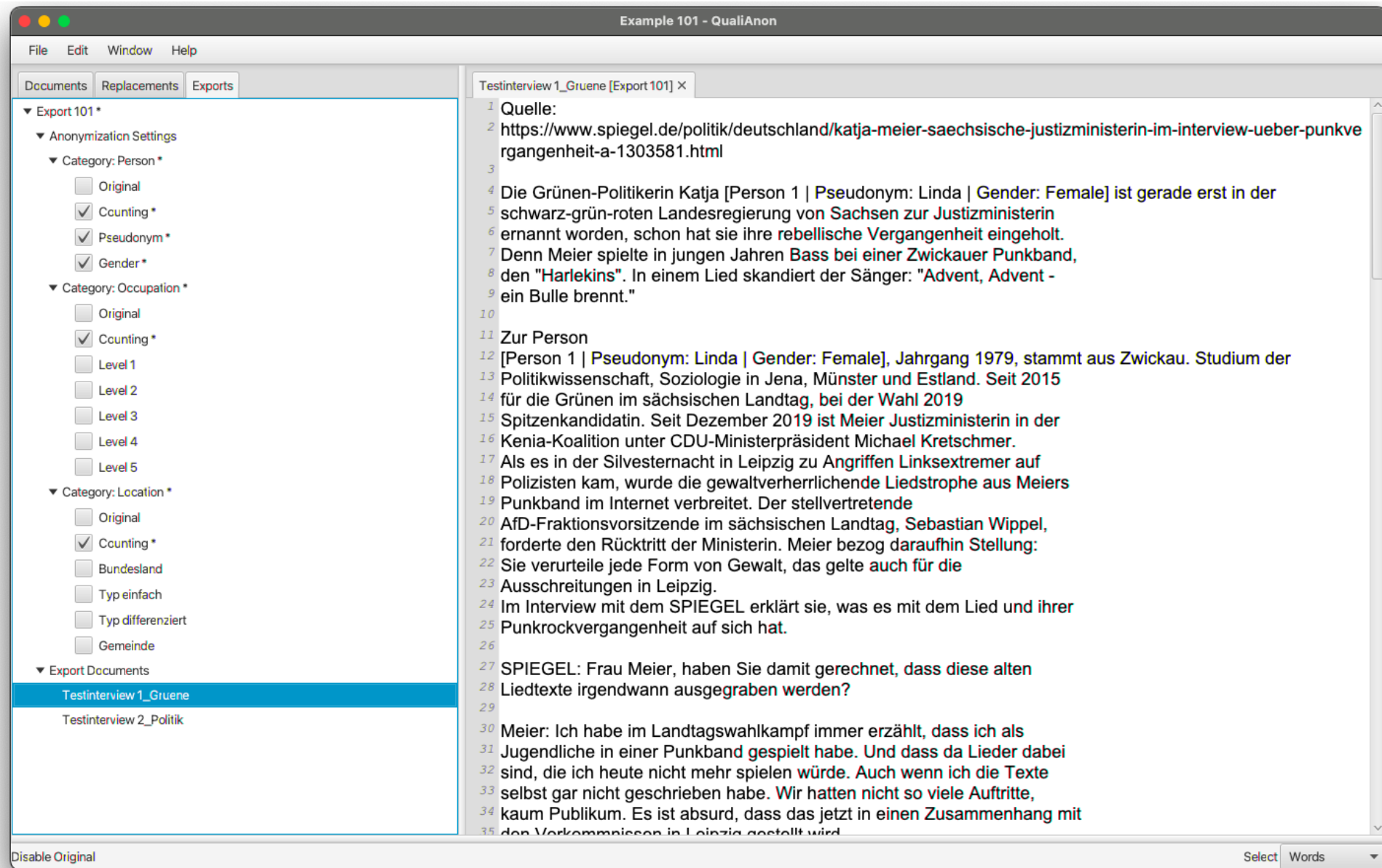
interview-b

- 1: Akira Toriyama Interview
- 3: Ever since the shocking debut of Dr. Slump, the cartoonist Akira
- 8: introduce everything about the "Toriyama school of cartoon thought",
- 15: Toriyama: I think they'll probably keep increasing gradually. If we're
- 27: Toriyama: Well, that in itself is obviously not a problem, as far as

Selected: 3 New replacement Reassign Remove

Saved document interview-b Select Words

Auswahl der Stellen,
die mit dem
„Replacement“
ersetzt werden sollen



The screenshot shows the QualiAnon application window titled "Example 101 - QualiAnon". The interface is divided into a settings panel on the left and a document preview on the right.

Settings Panel (Left):

- Export 101 ***
 - Anonymization Settings**
 - Category: Person ***
 - Original
 - Counting *
 - Pseudonym *
 - Gender *
 - Category: Occupation ***
 - Original
 - Counting *
 - Level 1
 - Level 2
 - Level 3
 - Level 4
 - Level 5
 - Category: Location ***
 - Original
 - Counting *
 - Bundesland
 - Typ einfach
 - Typ differenziert
 - Gemeinde
 - Export Documents**
 - Testinterview 1_Gruene (highlighted)
 - Testinterview 2_Politik

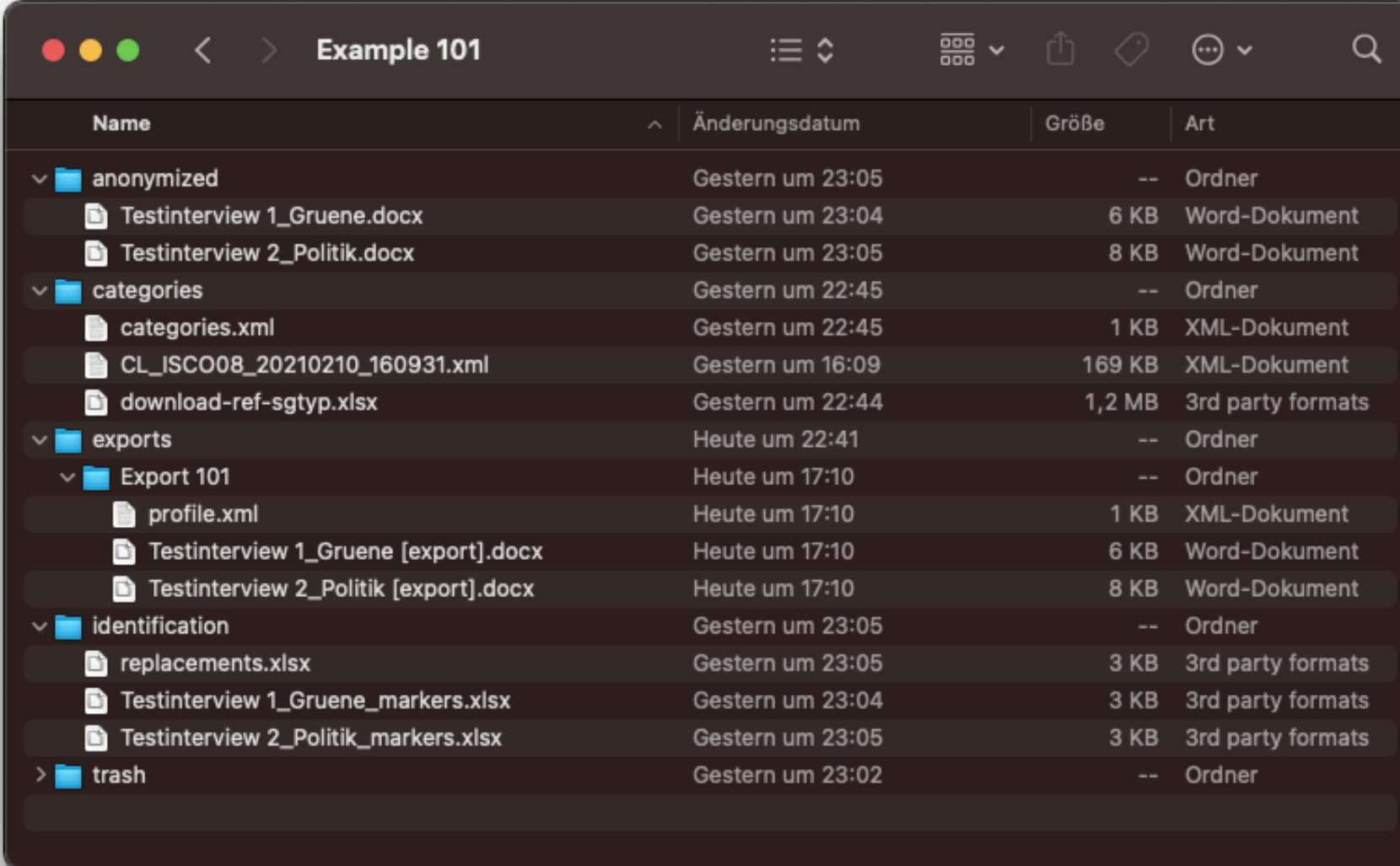
Document Preview (Right):

Testinterview 1_Gruene [Export 101] X

1 Quelle:
2 <https://www.spiegel.de/politik/deutschland/katja-meier-saechsische-justizministerin-im-interview-ueber-punkvergangenheit-a-1303581.html>
3
4 Die Grünen-Politikerin Katja [Person 1 | Pseudonym: Linda | Gender: Female] ist gerade erst in der
5 schwarz-grün-roten Landesregierung von Sachsen zur Justizministerin
6 ernannt worden, schon hat sie ihre rebellische Vergangenheit eingeholt.
7 Denn Meier spielte in jungen Jahren Bass bei einer Zwickauer Punkband,
8 den "Harlekins". In einem Lied skandiert der Sänger: "Advent, Advent -
9 ein Bulle brennt."
10
11 Zur Person
12 [Person 1 | Pseudonym: Linda | Gender: Female], Jahrgang 1979, stammt aus Zwickau. Studium der
13 Politikwissenschaft, Soziologie in Jena, Münster und Estland. Seit 2015
14 für die Grünen im sächsischen Landtag, bei der Wahl 2019
15 Spitzenkandidatin. Seit Dezember 2019 ist Meier Justizministerin in der
16 Kenia-Koalition unter CDU-Ministerpräsident Michael Kretschmer.
17 Als es in der Silvesternacht in Leipzig zu Angriffen Linksextremer auf
18 Polizisten kam, wurde die gewaltverherrlichende Liedstrophe aus Meiers
19 Punkband im Internet verbreitet. Der stellvertretende
20 AfD-Fraktionsvorsitzende im sächsischen Landtag, Sebastian Wippel,
21 forderte den Rücktritt der Ministerin. Meier bezog daraufhin Stellung:
22 Sie verurteile jede Form von Gewalt, das gelte auch für die
23 Ausschreitungen in Leipzig.
24 Im Interview mit dem SPIEGEL erklärt sie, was es mit dem Lied und ihrer
25 Punkrockvergangenheit auf sich hat.
26
27 SPIEGEL: Frau Meier, haben Sie damit gerechnet, dass diese alten
28 Liedtexte irgendwann ausgegraben werden?
29
30 Meier: Ich habe im Landtagswahlkampf immer erzählt, dass ich als
31 Jugendliche in einer Punkband gespielt habe. Und dass da Lieder dabei
32 sind, die ich heute nicht mehr spielen würde. Auch wenn ich die Texte
33 selbst gar nicht geschrieben habe. Wir hatten nicht so viele Auftritte,
34 kaum Publikum. Es ist absurd, dass das jetzt in einen Zusammenhang mit
35 den Verkommenen in Leipzig gestellt wird.

Disable Original | Select | Words

Auswahl der
Level für den
Export



Name	Änderungsdatum	Größe	Art
anonymized	Gestern um 23:05	--	Ordner
Testinterview 1_Gruene.docx	Gestern um 23:04	6 KB	Word-Dokument
Testinterview 2_Politik.docx	Gestern um 23:05	8 KB	Word-Dokument
categories	Gestern um 22:45	--	Ordner
categories.xml	Gestern um 22:45	1 KB	XML-Dokument
CL_ISCO08_20210210_160931.xml	Gestern um 16:09	169 KB	XML-Dokument
download-ref-sgtyp.xlsx	Gestern um 22:44	1,2 MB	3rd party formats
exports	Heute um 22:41	--	Ordner
Export 101	Heute um 17:10	--	Ordner
profile.xml	Heute um 17:10	1 KB	XML-Dokument
Testinterview 1_Gruene [export].docx	Heute um 17:10	6 KB	Word-Dokument
Testinterview 2_Politik [export].docx	Heute um 17:10	8 KB	Word-Dokument
identification	Gestern um 23:05	--	Ordner
replacements.xlsx	Gestern um 23:05	3 KB	3rd party formats
Testinterview 1_Gruene_markers.xlsx	Gestern um 23:04	3 KB	3rd party formats
Testinterview 2_Politik_markers.xlsx	Gestern um 23:05	3 KB	3rd party formats
trash	Gestern um 23:02	--	Ordner

Ordnerstruktur des QualiAnon-Projekts

Anonymisierte Texte, Exporte und sensible Originalinformationen werden nachvollziehbar getrennt in den jeweiligen Ordnern gespeichert

Wie bekomme ich Zugang zu QualiAnon?

- General information: www.qualiservice.org
- Access to software:
E-mail to: qualianon@uni-bremen.de
We will register you for the GitHub-project
<https://github.com/pangaea-data-publisher/qualianon>
- Access to manual as google doc:
https://docs.google.com/document/d/1fLLYvsgXjh_p9p_E1fhikkIPQb19VUiltbRgaWQoD-M/edit

