



Datenexploration und - visualisierung mit TextGrid und Dariah

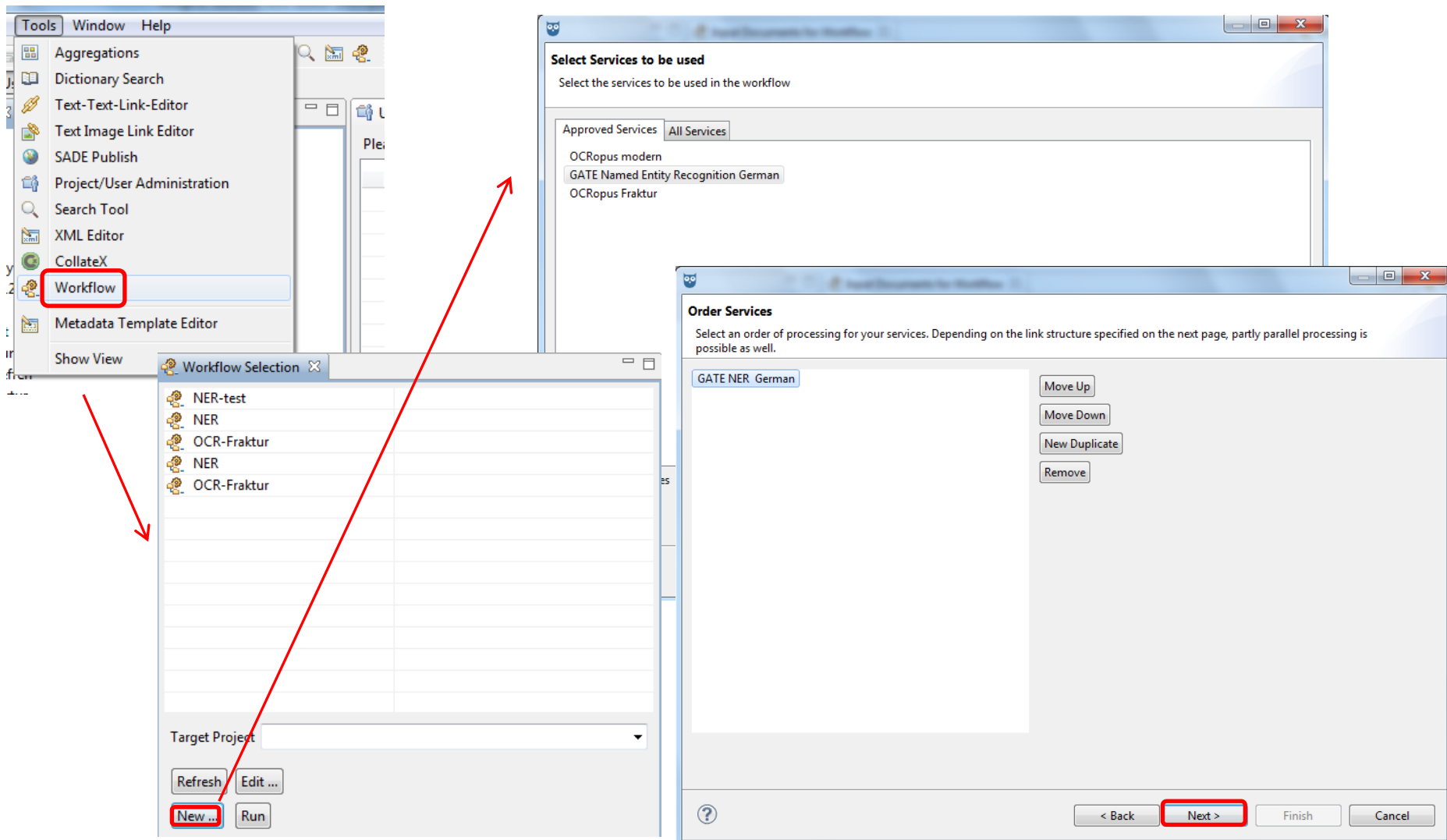
Philipp Vanscheidt

(Universität Trier / Technische Universität Darmstadt)

TextGrid und Dariah bieten verschiedene Möglichkeiten, Daten automatisch zu untersuchen und zu visualisieren. Solche Möglichkeiten sind:

1. Workflow mit Named Entity Recognition
2. Dariah GeoBrowser
3. DigiVoy

Einen Workflow anlegen

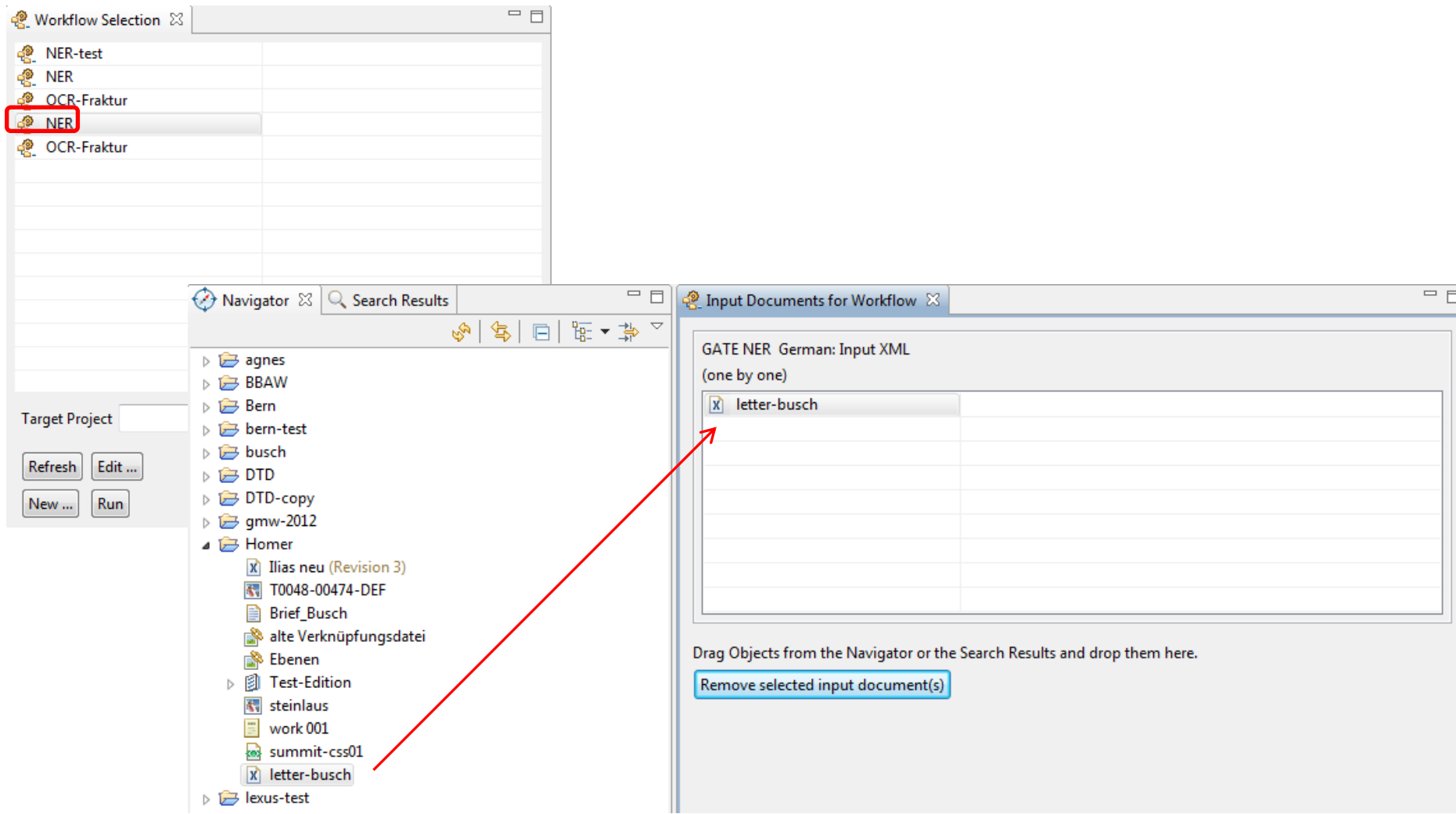


The image shows a sequence of steps in the TextGrid software interface for creating a workflow:

- Tools Menu:** The 'Workflow' option is highlighted with a red box in the 'Tools' menu.
- Workflow Selection Dialog:** A dialog box titled 'Workflow Selection' is open, showing a list of existing workflows: 'NER-test', 'NER', 'OCR-Fraktur', 'NER', and 'OCR-Fraktur'. The 'New ...' button at the bottom is highlighted with a red box.
- Select Services Dialog:** A dialog box titled 'Select Services to be used' is open, showing a list of approved services: 'OCROPUS modern', 'GATE Named Entity Recognition German', and 'OCROPUS Fraktur'.
- Order Services Dialog:** A dialog box titled 'Order Services' is open, showing a list of services with 'GATE NER German' selected. The 'Next >' button at the bottom is highlighted with a red box.

Red arrows indicate the flow from the 'Workflow' menu item to the 'Workflow Selection' dialog, then to the 'Select Services' dialog, and finally to the 'Order Services' dialog.

Einen Workflow vorbereiten

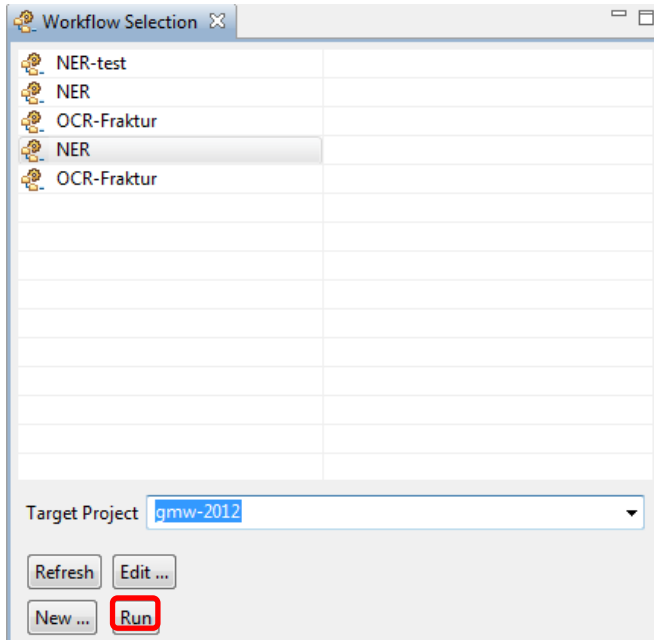


The screenshot displays three windows from the TextGrid application:

- Workflow Selection:** A list of workflows including 'NER-test', 'NER', 'OCR-Fraktur', and 'NER'. The 'NER' entry is highlighted with a red box.
- Navigator:** A tree view of the project structure. The 'letter-busch' file is selected under the 'Homer' folder.
- Input Documents for Workflow:** A window titled 'GATE NER German: Input XML (one by one)'. The 'letter-busch' document is listed in the input area. A red arrow points from the selected 'letter-busch' file in the Navigator to this input area.

Below the input list, there is a text prompt: "Drag Objects from the Navigator or the Search Results and drop them here." and a button labeled "Remove selected input document(s)".

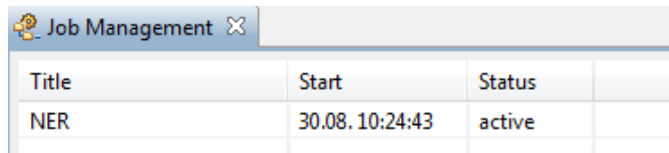
Einen Workflow nutzen



Workflow Selection dialog box showing a list of workflows: **NER-test**, **NER**, **OCR-Fraktur**, **NER**, and **OCR-Fraktur**. The **NER** workflow is selected. Below the list, the **Target Project** is set to **gmw-2012**. Buttons for **Refresh**, **Edit ...**, **New ...**, and **Run** are visible. A red arrow points from the **Run** button to the Job Management window below.

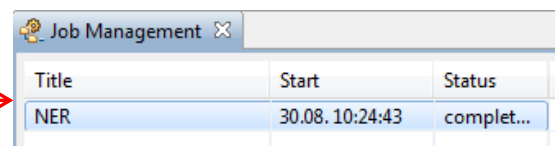


```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <TEI xmlns="http://www.tei-c.org/ns/1.0" xmlns:gate="http://www.gate.ac.uk" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instar
3 <teiHeader>
4 <fileDesc>
5 <titleStm
6 <title>Brief an Nanda Kessler</title>
7 <title>Edition des Briefs an Nanda Kessler</title>
8 <title>Brief an Nanda Kessler</title>
9 <respStm
10 <resp>editor</resp>
11 <name key="pnd:135414032"><name type="place" ref="#819689539c6852ae">Rettelbach</name>, Simon</name>
12 </respStm
13 <respStm
14 <resp>author</resp>
15 <name key="pnd:118517880">Busch, Wilhelm</name>
16 </respStm
17 <respStm
18 <resp>Data Contributor</resp>
19 <name>Simon.<name type="place" ref="#819689539c6852ae">Rettelbach</name>@textgrid.de</name>
20 </respStm
21 </titleStm
22 <publicationStm
23 <idno type="textgrid">textgrid:h4xv.0</idno>
24 </publicationStm
25 <sourceDesc
26 <biblFull
27 <titleStm
28 <title>Briefe 1841 bis <date>1892</date></title>
29 <author key="pnd:118517880">Busch, Wilhelm</author>
30 </titleStm
31 <extent>367</extent>
32 <publicationStm
33 <date when="1968">1968</date>
34 <pubPlace><name type="place">Hannover</name> | <name type="place">Niedersachsen</name> | <name type="p
35 </pubPlace>
36 </publicationStm
37 </biblFull
38 </sourceDesc
```



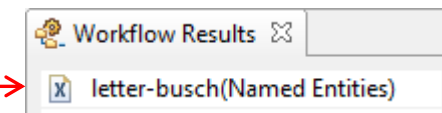
Title	Start	Status
NER	30.08.10:24:43	active

A red arrow points from the Job Management window to the Workflow Results window on the right.



Title	Start	Status
NER	30.08.10:24:43	complet...

A red arrow points from the Job Management window to the Workflow Results window on the right.



Workflow Results window showing the workflow **letter-busch(Named Entities)** with a document icon next to it. A red arrow points from the Job Management window to this window.