

# A CANONICITY SCORE FOR YOUR CORPUS

MERTEN KRÖNCKE      JANA ECKARDT      AGNES HILGER  
Universität Stuttgart      Universität Göttingen      Universität Würzburg

## 01

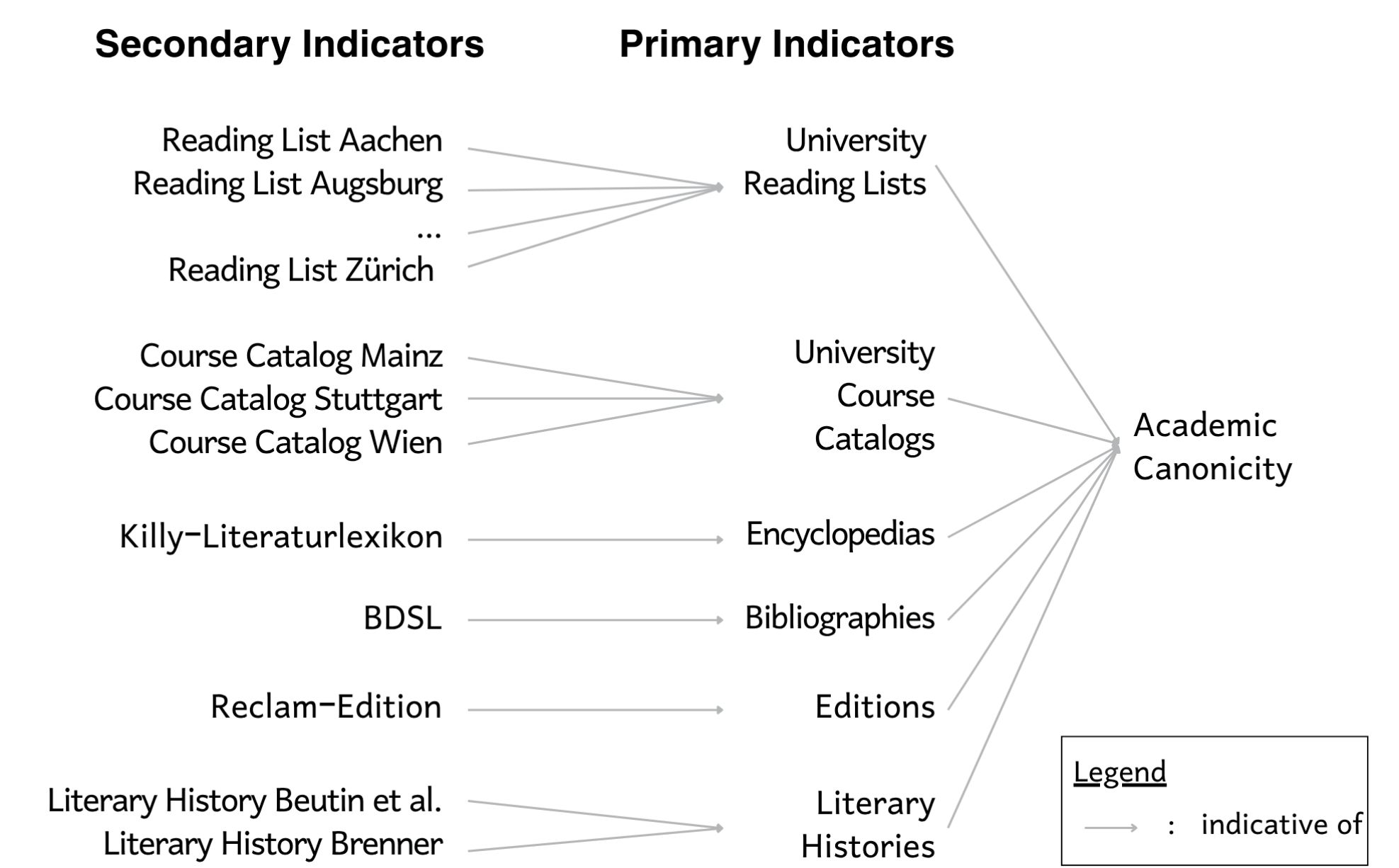
### INTRODUCTION

Canonicity scores are valuable for canon critique, corpus construction, and diachronic comparisons. We present author-based scores for the current academic canonicity of more than 5,000 writers from German-language literature since 1600 – ready to use as a canonicity score for your corpus out of the box.

## 02

### BACKGROUND & METHODS

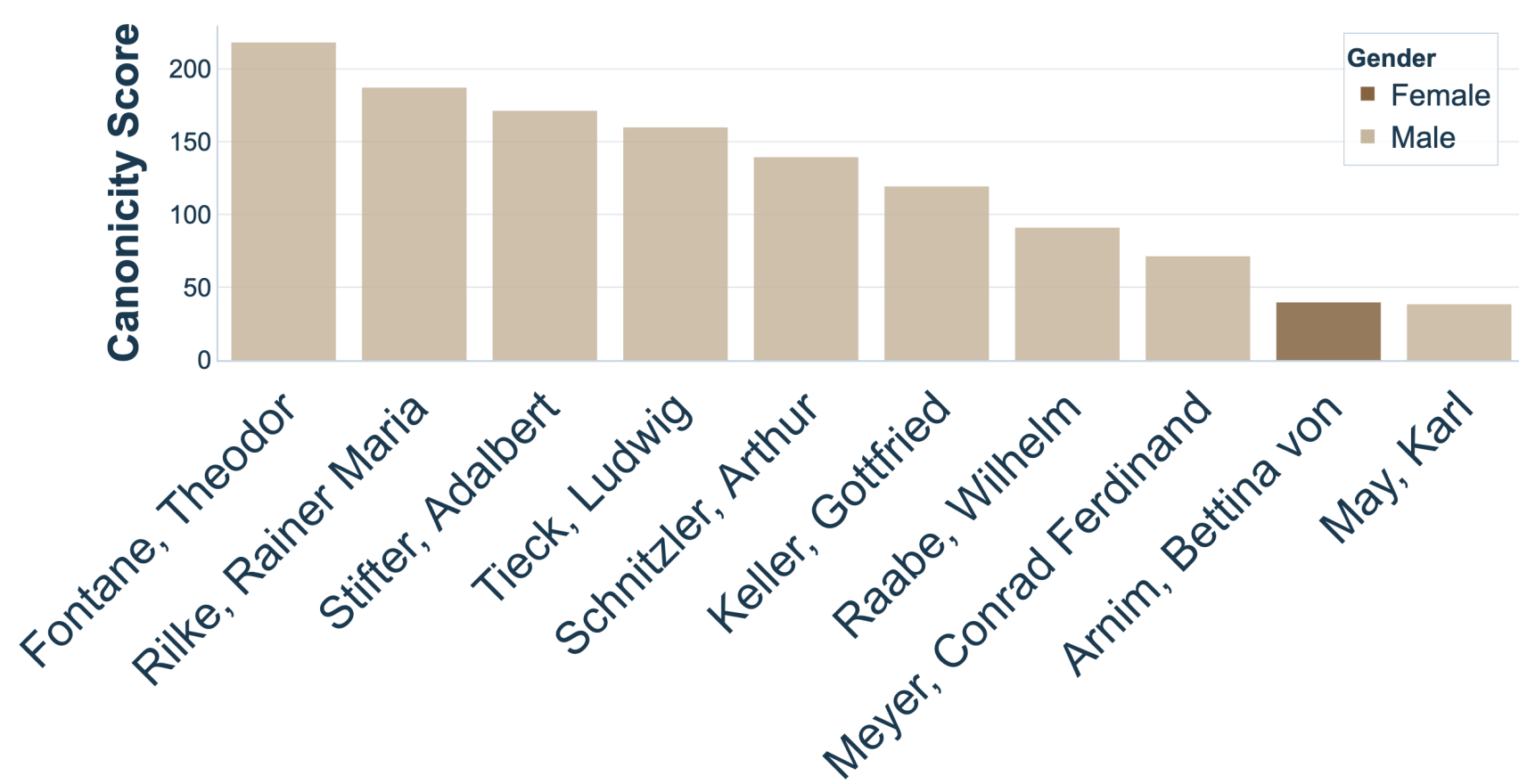
The canonicity scores are based on a forthcoming publication by our group (Eckardt/Hilger/Kröncke 2027). According to our explication, the present-day canonicity of an author among literary scholars is equal to the degree to which these scholars have deemed the author worthy of attention until the present day. Drawing on 40 indicators from 6 source types, we aggregated one score per author, scaled from 1000 downwards. The ranking on the right shows the top of this distribution.



## 03

### USE CASE 1: ELTEC

Mapping our scores to ELTeC-deu via author GND IDs yields a highly significant correlation with the corpus's work-specific canonicity levels (Spearman's  $r = 0.52$ ,  $p < .001$ ). The figure shows the top 10 ELTeC-deu authors according to our scores.



## 04

### USE CASE 2: GERDRACOR

Our Canonicity Scores can be applied to the authors in GerDraCor using OpenRefine to add Wikidata IDs to GND IDs. The figure shows the UMAP for our six primary indicators as vectors for the authors represented in GerDraCor.

