

Jamie LOD-iver: Extracting and Enriching Historical Recipes

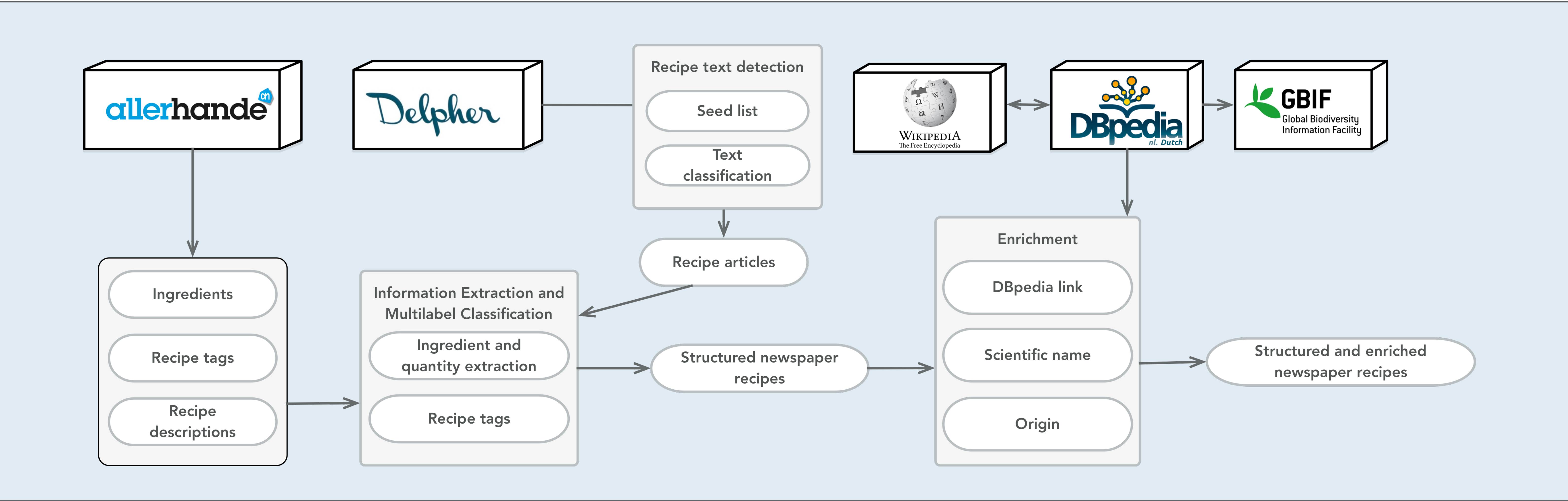
Melvin Wevers, KNAW Humanities Cluster, DHLab
Marieke van Erp, KNAW Humanities Cluster, DHLab
Hugo Huurdeman, University of Amsterdam, University of Oslo
Richard Zijdemans, International Institute for Social History, University of Stirling

What we eat has been an important part of our cultural identity. Our food habits can be distilled from recipes. Food websites are a treasure trove of structured recipes that can easily be analysed automatically through their hRecipe or schema.org markup.

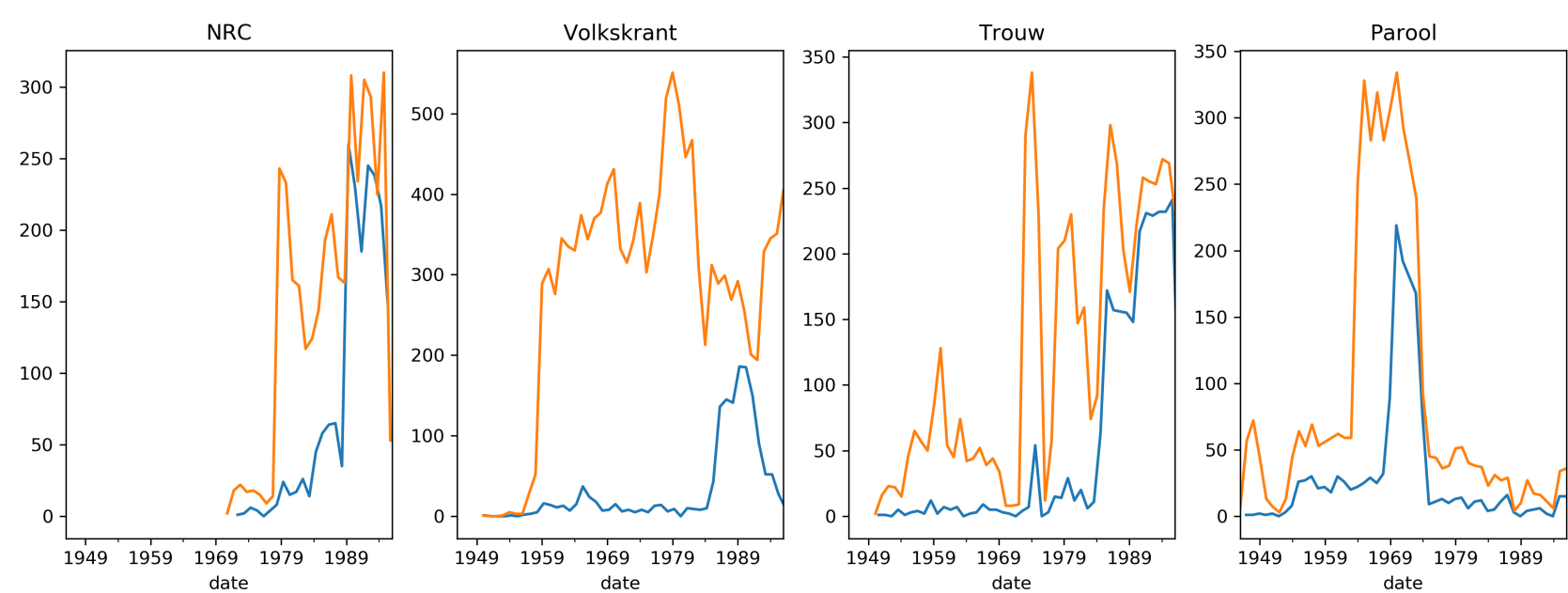
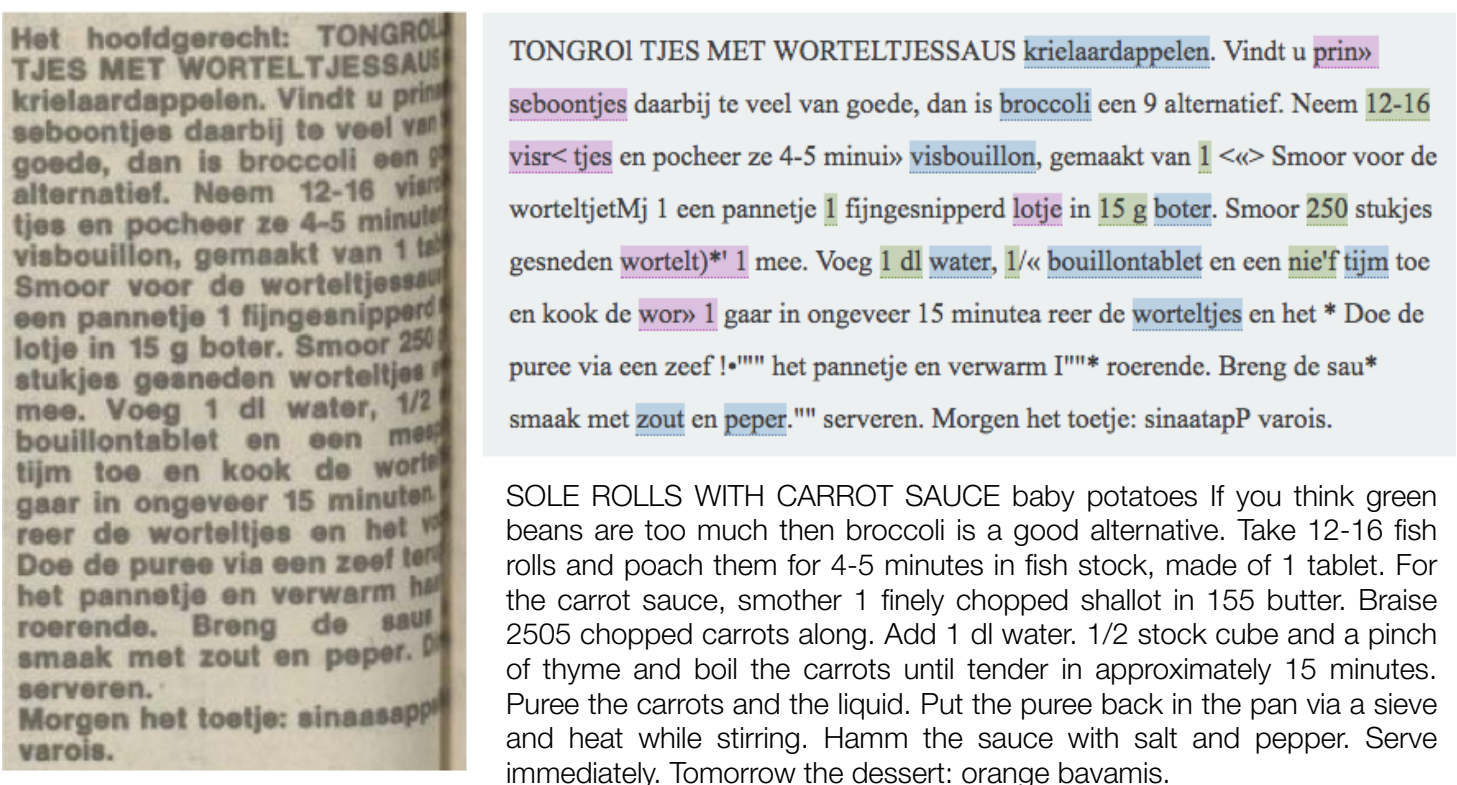
Unfortunately, this is not the case for historical recipes, often available as part of digitized archives. This means that currently we are unable to answer important questions on how food culture has changed over time.

Historical newspapers provide a lens on customs and habits of the past. The challenge is that newspaper data is often unstructured, highly varied, and of fluctuating OCR quality. Therefore, it is difficult to locate and extract recipes from them.

We present a method for extracting and enriching over 24,000 recipes from four digitized historical newspapers (Het Parool, Trouw, De Volkskrant, and NRC Handelsblad) published between 1950 and 1995 via Delpher.



From Newspaper data to Historical Recipe Web



Retrieved articles using seed list (blue) (n = 7,500) and using our classifier (orange) (n = 27,411)

Recipes	27,411
Tags	34,479
Ingredients	365,133
Quantities	107,864
DBpedia	14,900
Scientific	1,657
GBIF	829

Search, Analysis & Visualization

