Context model for olfactory mentions in texts

Deliverable D4.5

Version DRAFT

NEGOTIATING OLFACTORY AND SENSORY EXPERIENCES IN CULTURAL HERITAGE PRACTICE AND RESEARCH

The Odeuropa project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004469. This document has been produced by the Odeuropa project. The content in this document represents the views of the authors, and the European Commission has no liability in respect of the content.

Grant Agreement No.	101004469
Project Acronym	ODEUROPA
Project full title	Negotiating Olfactory and Sensory Experiences in Cultural Heritage Practice and Research
Funding Scheme	H2020-SC6-TRANSFORMATIONS-2020
Project website	http://odeuropa.eu/
Project Coordinator	Prof. Dr. Inger Leemans KNAW Humanities Cluster Email: inger.leemans@huc.knaw.nl
Document Number	Deliverable D4.5
Status & version	DRAFT
Contractual date of delivery	30 June 2023
Date of delivery	30 June 2023
Type	Other
Security (distribution level)	Public
Number of pages	47
WP contributing to the deliverable	WP4
WP responsible	WP4
EC Project Officer	Hinano Spreafico

Authors: Inna Novalija,[1] Raphaël Troncy,[2] Dunja Mladenić,[1] Luis Rei,[1] Pasquale Lisena,[2] Stefano Menini,[3] Youssra Rebboud,[2]

Internal reviewers: Sara Tonelli,[3] Victoria-Anne Michel,[4]

Affiliations (1) JSI , (2) EUR, (3) FBK, (4) ARU

Keywords: olfactory context, causality analysis, question answering, deep learning, semantic web

Abstract: Deliverable D4.5 describes the methodology and results related to obtaining contextual information for olfactory mentions in text. We investigate identifying causal relations from historical documents, presenting three different approaches: i) a supervised approach for effects and circumstances extraction, ii) a question-answering approach for causes and iii) an effects extraction and transfer learning approach targeted at causality. In our research, we apply zero-shot classification methods for data enrichment with concepts from Odeuropa vocabularies and present examples that demonstrate the usefulness of the application of semantic technologies for detection of olfactory causality in text. Furthermore, in this document, we provide a historical analysis of causes and effects for olfactory mentions, extracted for different historical periods within a number of analyzed corpora.

Table of Revisions

Version	Date	Description and reason	By	Affected sections
0.1	23 February 2023	ToC	Inna Novalija	Table of Contents
0.2	31 March 2023	First Draft	Inna Novalija, Raphaël Troncy	All
0.3	30 April 2023	Second Draft	Inna Novalija, Raphaël Troncy, Dunja Mladenić, Luis Rei, Stefano Menini, Pasquale Lisena	All
0.4	13 June 2023	Version for Review	Inna Novalija, Raphaël Troncy, Dunja Mladenić, Luis Rei, Stefano Menini, Pasquale Lisena	All
0.5	21 June 2023	Revision after Review	Inna Novalija, Raphaël Troncy, Dunja Mladenić, Stefano Menini, Pasquale Lisena	All
1.0	30 June 2023	Final check and approval by project manager and PI	Marieke van Erp, Inger Leemans	-

Executive Summary

Summary table	
Challenges	There are multiple challenges associated with the extraction of contextual information from historical data, in particular with obtaining the information about pre-conditions and post-condictions/causes and effects related to olfactory mentions in text. Challenges are associated with different types of causality that might be present in text (such as explicit or implicit causality, intra-sentential or inter-sentential causality). Different data sources might produce different results, and causes and effects might vary in different historical periods.
Barriers	The main barrier to overcome in context modeling for olfactory mentions is the specificity of the task. The topic of context modelling for olfactory information combines together the task of smell extraction along with extraction of causal relations (causes and effects, pre-conditions and post-conditions). The specificity of the topic reflects the lack of annotated datasets that can be used for training models and benchmarking.
Practices	The practices for obtaining a context model for olfactory mentions in text rely on different methods, including knowledge based, rule based, statistical, machine learning and deep learning based approaches. In this deliverable, we assess different technologies for extracting causes and effects of smells, including a supervised approach, an approach based on question answering and a transfer learning approach that leverage transformer-based models.
Guidelines	In this deliverable, we provide the methodology for obtaining context for olfactory mentions in text using state-of-the-art techniques for semantic annotation. The developed methodology allows for enrichment of textual content with Odeuropa vocabularies concepts, grouping the extracted causes and effects in context-related layers by historical periods, and using them in the digital cultural heritage domain.

Layman's Summary

What causes a smell and what do smells cause? Smells can occur nearly everywhere, affecting the world around them. But what are the causes and effects of smells, and how do they differ over time and contexts? Deliverable D4.5 'Context model for olfactory mentions in texts' describes a new methodology for connecting contextual information to smell occurrences in historical, olfactory-rich information sources.

We investigate the possibility to model complex causal relationships (cause/effect, circumstances/-effect) related to pre-conditions and post-conditions associated with smell events.

State-of-the-art machine learning and deep learning methods are used to extract causes and effects connected to smells for defined historical periods. The results obtained are furthermore grouped into specific layers based on inputs from historians.

We observe that causes and effects related to smell events differ between various historical periods, as well as vary in relation to analyzed data sources.

The complex contextual information obtained presents a basis for the enrichment process of the European Olfactory Knowledge Graph and can be used by experts in the olfactory area for research and development purposes.

Contents

1 Introduction

This Deliverable D4.5 ”Context Model for Olfactory Mentions in Texts” presents the main results obtained inthe task T4.4 ”Harvesting Context Related to Olfactory Cultural Heritage”, which uses semantic technologies and text mining to analyse the context in which references to smell occur and investigate potential changes over time. Specifically, in Task 4.4, we investigate the time period from 1600 until 1925.

Within this work, we have developed methods to identify the causal relationships connected to the appearance of smell in texts. The survey on the extraction of causal relations from natural language text [Yang et al., 2021] specifies cause/effect analysis as the detection of a relationship between two entities e1 and e2, such that the occurrence of e1 results in the occurrence of e2.

We first present a brief overview of state-of-the-art technologies and methods for dealing with causality in text, including knowledge based approaches and dependency patterns techniques, statistical and machine learning methods and finally deep learning approaches. In our research work, we have investigated some approaches to model complex contextual relationships, such as causal relationships, describing what led to a situation where and odour appears and what kind of consequences or effects were caused by presence of that odour. We present three different methods for causality extraction, including supervised approach for circumstances/effects mining, question answering approach for cause/effect detection and transfer learning approach targeted at causality. We have used machine learning methods for analysis of causality over time within different corpora containing olfactory mentions. Domain experts have been involved for validating a set of groups (or layers) on top of extracted causal relations from text. The results of this task constitute the basis for enriching the European Olfactory Knowledge Graph.

This remaining of the document is structured as follows:

- Section 2 provides a brief overview of related initiatives in the area of causality analysis.
- Section 3 describes three different approaches we developed for causality analysis in the domain of smell. The developed methods show how causality analysis can be addressed with supervised methods, question answering methods and transfer learning methods. In addition, this section presents a list of datasets used for development and application purposes.
- Section 4 explains how semantically enriched text can be used for smell causality analysis. In this research work, we consider that ”context” represents pre-conditions and post-conditions (”causes”/”effects”, ”circumstances”/”effects”) associated with smell events.
- Section 5 details the detection of ”causes” and ”effects” layers. The results obtained with causality extraction methods, semantically enriched with Odeuropa taxonomy concepts, are used to define the specific groups for causes and effects over time.

Finally, we present some conclusion and outline future work in Section 6.

2 Related Work

There are different methods for cause-effect detection, including knowledge-based methods, statistical methods and deep learning methods. However, the current related initiatives are not focused on historical text analysis. In comparison with related approaches, in the Odeuropa project, we are dealing with extraction of causes and effects for a specific topic (olfactory events), and analyze the historical development of smell causality.

Past works [Yang et al., 2021] highlight a number of definitions and challenges associated with the extraction of causes and effects from text:

- explicit and implicit causality [Yang et al., 2021], where explicit causality represents the relations that are connected by explicit causal connectives, such as causal links, causative verbs, adverbs adjectives, conditional phrases. Implicit causality on the other hand does not involve any connectives, but relies on the background knowledge and reasoning mechanisms; [Blanco et al., 2008, Hendrickx et al., 2010, Sorgente et al., 2013]

- intra-sentential causality, where cause and effect are located in the same sentence, and inter-sentential causality where cause and effect are distributed between different sentences.

In this work, we aim at covering different causality aspects, focusing mainly on sentences that contain olfactory (or smell words) mentions.

2.1 Knowledge-Based Methods and Rule-Based Approaches

The knowledge-based and rule-based methods are basic common methods for causality detection, which, however require extensive domain knowledge and have poor cross-domain applicability [Yang et al., 2021]. Expressing cause-effect relations in text can be obtained via specifying rules, using:

- causal links to sentences;
- specific language constructions expressing for results;
- causative verbs, adjectives and adverbs;
- constructions in the form $if \longrightarrow then$

Altenberg [Altenberg, 1984] defined a number of causal link types, such as adverbial links, prepositional links, subordination links, clause-integrated links that can be used for detecting cause/effect or pre-conditions/post-conditions for a specific event in text. Another variant of causal structures are conditional causal structures, such as ”$If \longrightarrow then$” that explicitly represent the ”cause” and ”effect” for some events and reflect the ”causal” relation between them.

2.2 Dependency Patterns

Natural language processing and dependency patterns approaches are frequently used for cause-effect detection [Garcia, 1997, Khoo et al., 1998, Radinsky et al., 2012, Ittoo and Bouma, 2013]. Patterns can be obtained via sentence structure analysis, lexico-semantic or syntactic analysis.

Inter-sentence causality is explored by Garcia et al. [Garcia, 1997] and Khoo et al. [Khoo et al., 1998] for identification of explicit causal relations. A tool for the extraction of causal relations in French based on 23 explicit causal verbs is presented in [Garcia, 1997].

In the domain of public media, Radinsky et al. [Radinsky et al., 2012] proposed a Pundit algorithm to generate causality pairs from news articles. A specific generalizations ⟨ Pattern, Constraint, Priority ⟩ was suggested in order to achieve high automation. Verbal and non-verbal patterns for the identification of implicit causality have been used by Ittoo and Bouma [Ittoo and Bouma, 2013], who developed a minimally supervised method in order to identify three pre-defined types of implicit causality in an iterative way. The approach included different pattern types:

- result verbal patterns, such as ”increase”, ”reduce”, ”kill”, ”become”;
- patterns for detecting inseparable causes and effects;
- nonverbal patterns, like ”rise in” and ”due to”.

2.3 Statistical/Machine Learning Approaches

Although statistical and machine learning-based approaches require less predefined patterns and hence less manual work, this type of causality analysis requires some feature engineering. The approaches often employ NLP tools (e.g., Spacy, Stanford CoreNLP, Stanza) [Vasiliev, 2020, Manning et al., 2014, Qi et al., 2020] for features generation based on labeled data. Different algorithms, such as Support Vector Machines (SVM), NaÏve Bayes (NB), and Logistic Regression (LG) are utilized in classification tasks. We briefly describe below a number of related works that apply machine learning techniques and publicly available datasets to causality detection.

Blanco et al. [Blanco et al., 2008] use different kinds of features and machine learning algorithms with well-known TREC dataset [Craswell et al., 2020]. SemEval-2010 Task corpus [Hendrickx et al., 2010] has been frequently utilized for model evaluation, see for instance Pakray and Gelbukh [Pakray and Gelbukh, 2014], Sorgente et al. [Sorgente et al., 2013], and Zhao et al. [Zhao et al., 2016]. The sentence syntactic structure is explored by Zhao et al.[Zhao et al., 2016] for causality detection. Pechsiri et al. [Pechsiri et al., 2006] train a NB and SVM classifier based on verb-pair rules in order to obtain implicit causality from Thai texts.

From the perspective of data availability, causality detection for olfactory mentions in text is challenging due to the lack of olfactory related data. Odeuropa is therefore the first project that provides a benchmark annotated dataset specifically for the olfactory domain.

2.4 Deep Learning Approaches

Deep learning methods have recently become the new state-of-the-art methods for causality detection. Using deep learning methods allows for high productivity, but at the same time they are computationally costly and have lower explainability. Neural networks (NNs) are basic algorithms for deep learning (DL). Different researchers have tackled the causality detection problem with deep learning methods [Kyriakakis et al., 2019, Wang et al., 2016, Zhang et al., 2018].

Long Short-Term Memory (LSTM) models with word embeddings [Martínez-Cámara et al., 2017] were proposed for causality mining, while CNN was used [Jin et al., 2020] for detecting deeper contextual semantic information between causes and effects. In [Dasgupta et al., 2018], the authors proposed a linguistically informed recursive neural network architecture for automatic extraction of cause-effect relations from text. The extracted causal events and their relations are used to build a causal-graph after clustering and appropriate generalization, which is then used for predictive purposes.

3 Extracting Causes and Effects of Smells

In this section, we present the various approaches that Odeuropa partners explored for context detection and causality analysis. including a supervised approach for the extraction of effects and circumstances, a question-answering approach for the detection of causes and effects and a transfer learning approach for characterizing causes. As a preamble, the following dataset subsection describes the datasets used for the analysis, and refers to the Odeuropa benchmark [Menini et al., 2022b] as standardized dataset for comparison.

3.1 Datasets

3.1.1 Historical Datasets

We have used a number of historical datasets that contain olfactory mentions in text:

- Project Gutenberg [Project Gutenberg, 2023] is a library of over 70,000 free eBooks. The project contains digitized historical texts of different genres, including fiction, poetry, scientific studies, travel literature etc.

- The Old Bailey Corpus [Clarin: Old Bailey, 2023] is a socio-linguistically, pragmatically and textually annotated corpus based on the Proceedings of the Old Bailey. The Old Bailey document corpora present collection of proceedings from London's Central Criminal Court. The Proceedings of the Old Bailey (2163 volumes contain almost 200,000 trials) were published from 1674 to 1913, containing the spoken word of the period.

- The Royal Society Corpus (RSC) [Clarin: Royal Society, 2023] is a diachronic corpus of scientific English covering hundreds (1665–1996) years of scientific writing. The corpus comprises 47 837 texts, primarily scientific articles, and is based on publications/proceedings of the Royal Society of London.

3.1.2 Odeuropa Benchmark

The Odeuropa benchmark consists of a selection of documents from a pool of documents covering different time periods (from 1620 to 1925) and topics (e.g. medicine, law, literature), for 7 languages. The annotation was carried out using the INCEpTION annotation platform [Klie et al., 2018], following the guidelines presented in [Tonelli et al., 2021a].

Whereas the Project Gutenberg dataset, The Old Bailey Corpus and Royal Society Corpus contain texts predominantly in English, the Odeuropa benchmark contains multilingual and specifically annotated olfactory texts. We have used Odeuropa benchmark to test different approaches for extracting causes and effects. In the historical analysis, we have applied question answering approach with semantic enrichment (as one of novel approaches for detection smell causes and effects) on the Project Gutenberg dataset, The Royal Bailey Corpus and Royal Society Corpus.

Table 1 provides the datasets summary information for data used for extraction of effects/circumstances and in causality analysis.

Table 1: Datasets Summary Information

Dataset	Size	Genres	Covered Years
Odeuropa benchmark (EN, IT, FR, NL, SL, GE, LA)	On average, 10 documents per domain, per language	Different genres, domains (Household and Recipes, Law, Literature, Medicine and Botany, Perfumes and Fashion, Public health, Religion, Science and Philosophy, Theatre, Travel and Ethnography.)	1620 to 1925
Gutenberg collection	70,000 free electronic books	Different genres (fiction, poetry, scientific studies etc.)	from 1003 (actively from 1500)
Old Bailey collection	200,000 trials	Legal, criminal	1674 to 1913
Royal Society collection	47 837 texts	Science	1665 to 1996

3.2 Extracting Effects and Circumstances Using a Supervised Approach

For the supervised classifier, we adopt a multi-task learning [Caruana, 1993, Caruana, 1997] approach. We trained a neural network to learn different tasks in parallel while using a shared representation, so that each task updates the model's shared parameters with respect to every task in the network, ideally leading to a more robust representation with less over-fitting. In this configuration each task corresponds to the classification of a single frame element. We classify [Tonelli and Menini, 2021a] 10 frame elements, namely Smell Word, Smell Source, Quality, Odour Carrier, Evoked Odorant, Location, Perceiver, Time, Circumstances, Effect.

We adopted a multitask approach since it was more effective than a single multiclass classifier (see Table 2 for the comparison), and because simpler tasks, as can be smell words detection, can act as auxiliary task and share information for the classification of more difficult frame elements.

To fine-tune the models, we used MaChAmp [van der Goot et al., 2021], a toolkit for fine-tuning in multi-task settings, and the classification of each frame element was configured as BIO tasks, since they are usually spans of multiple words.

The dataset used for fine-tuning the models is the Odeuropa benchmark [Menini et al., 2022a] more extensively presented in Deliverable D3.2 and created with the guidelines from Deliverable D3.1 [Tonelli and Menini, 2021b]. All the results reported in Table 2 are the average of the experiments done with 10 different data splits, with each data split having 80% of the smell words and related frame elements (FE) as training data, 10% for validation and 10% as test. The splits are not completely random as we sought to keep the same temporal and domain distribution in every run.

MaCHAmp can be configured with a different loss weight parameter for each task to define the main/auxiliary tasks. For each task, we compare two different values of loss weight: 1 and 0.75, testing different combinations over the 10 tasks. A hyperparameter search was applied to one of the splits with the following search space: learning rate $[1e-3, 1e-4, 1e-5]$, batch size $[16, 32]$ and number of training epochs $range(1, 10)$. All configurations reported in Table 2 use a learning rate of $1e-4$ and a batch size of 32, and all the loss weight set to 1, which yield the best performance.

The performance of the classifier on the English language is displayed in Table 2. The table compares models obtained by fine-tuning bert-base-cased[1] [Devlin et al., 2019] on English data only with the models models obtained from fine-tuning bert-base-multilingual-cased [2] [Devlin et al., 2019] using olfactory annotations from 6 different languages (English, French, Italian, Dutch, German, Slovene). We also compare the results obtained by using the *span F1* and *token F1* as training metrics.

Table 2: Results of the English frames classifiers on the 10 frame elements used in Odeuropa Project. Each result is the average of 10 different runs done on 10 different data splits)

Approach	Training Metric	Smell Word	Smell Source	Quality	Odour Carrier	Evoked Odorant	Location	Perceiver	Time	Circ.	Effect
Multitask-mono	span-f1	0,871	0,571	0,758	0,482	0,572	0,542	0,510	0,434	0,461	0,405
Multitask-mono	token-f1	0,864	0,571	0,759	0,483	0,535	0,535	0,484	0,417	0,480	0,365
Multitask-multi	span-f1	0,865	0,574	0,759	0,462	0,517	0,546	0,488	0,528	0,480	0,339
Multitask-multi	token-f1	0,783	0,536	0,745	0,479	0,552	0,508	0,489	0,465	0,452	0,347
BERT-Multiclass	span-f1	0,821	0,461	0,652	0,361	0,295	0,349	0,365	0,37	0,215	0,115

3.3 Extracting Causes and Effects Using a Question Answering Approach

Question Answering models are used in the machine learning and artificial intelligence community for retrieving answers to questions from a given text based on specific content. For Odeuropa purposes, we formulate questions related to “causes” and “effects” of smell events and explore the sentences where we have predefined smell occurrences. We used a RoBERTa-Large QA Model [Bartolo et al., 2021] trained from a RoBERTa large model [Bartolo et al., 2020a].

The utilized QA model is originally trained on synthetic adversarial data generated using a BART-Large question generator on Wikipedia passages from SQuAD as well as Wikipedia passages external to SQuAD, and then it is trained on SQuAD and AdversarialQA [Bartolo et al., 2020b]. In QA setting, we have provided cause/effect and smell-related prompts to the model and then analyzed the obtained outputs. For instance, the extraction of causes and effects related to olfactory mentions in text has been based on the following prompts (used in question answering model):

- “What is the cause of the smell word?”
- “What is the effect of the smell word?”

where ”smell word” would represent the smell word occurring in the analyzed sentence (for instance, ”scent”, ”odour”, ”odor”, ”stench”, ”stink”, ”stunk”, ”perfume”, ”aroma”, ”reek”, ”fragrance”,

[1] https://huggingface.co/bert-base-cased
[2] https://huggingface.co/bert-base-multilingual-cased

”whiff” etc.). When we find a smell word occurring in text, we substitute ”smell word” with a smell word value, in such way we can obtain questions such as:

- ”What is the cause of the stench?”
- ”What is the effect of the stench?”

Figure 1 displays the methodology behind the question answering approach along with causality analysis. In particular, in order to test the question answering approach, we have performed a sample selection of 200 sentences by periods of 100 years from the explored data sources described above (Project Gutenberg data collection [Project Gutenberg, 2023], Old Bailey data collection[Clarin: Old Bailey, 2023], Royal Society data collection [Clarin: Royal Society, 2023]). We have utilized the QA models for causality detection based on smell prompts and analyzed the answers obtained from the model. The question answering example with results of cause-effect extraction is presented in Table 3.

We have performed a manual validation for each data source, defining a confidence threshold for causes and for effects for each data source. In particular, we have found that on average, the confidence score for correctly identified causes is 0.27 (for Old Bailey and for Royal Society data collections) and 0.38 (for Project Gutenberg data collection). On average, the confidence score for correctly identified effects is 0.32 (for Project Gutenberg data collection), 0.27 (for Old Bailey data collection) and 0.31 (for Royal Society data collection). In comparison, the incorrectly identified causes got the average confidence score of: 0.20 (Project Gutenberg data collection), 0.05 (Old Bailey data collection) and 0.19 (Royal Society data collection). The incorrectly identified effects obtained the average confidence scores of: 0.18 (Project Gutenberg data collection), 0.12 (Old Bailey data collection), 0.12 (Royal Society data collection).

Following that, Figure 1 displays a number of steps reflecting how to use results, obtained with question answering methodology, for semantic enrichment and for historical applications (Sections 4 and 5).

Figure 1: QA Methodology with Semantic Annotation and Validation

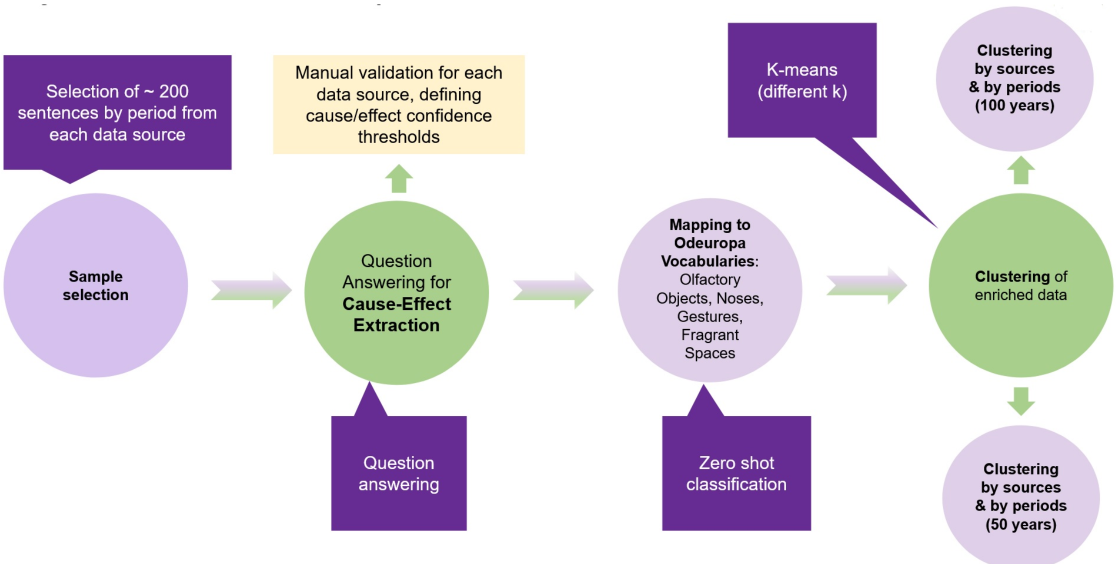


Table 3: Question Answering Cause-Effect Extraction Example

<table>
<tr><td colspan="3">Book: Fynes Moryson, An itinerary vvritten by Fynes Moryson Gent. First in the Latine tongue, and then translated by him into English: containing his ten years travel through the twelves dominions of Germany, Bohmerland, Sweitzerland, Netherland, Denmarke, Poland, Italy, Turky, France, England, Scotland, and Ireland [...]. London: Iohn Beale, 1617
File:
<code>"044E Fynes Moryson 1617 An Itinerary__58_distribution__8_distribution_TRAV.txt"</code></td></tr>
<tr><td colspan="3">Sentence:
”It is compassed with high hils on all sides, but onely where the Sea enters on the South-side at a passage fifty paces broad, and the forme of it is round, and the hils that compasse it now seeme pleasant, but of old were all coued with a thicke wood, which shutting vp the aire, and by the shadow drawing many birds to it, was thought to be the cause that these birds stifled with the smell of brimstone, fell suddenly dead, till the Emperour Augustus”</td></tr>
<tr><td>Extracted with Question Answering
Cause: ”brimstone”,
Cause_Score”:
0.11893732100725174,
Effect: ”fell suddenly dead”,
Effect_Score:
0.6697907447814941</td><td>Extracted by Supervised Methods
Smell_Word: ”smell”,
Smell_Source: ”of brimstone”
Effect: ”these birds | fell suddenly dead”</td><td>Annotated in Odeuropa Benchmark
Odour_Carriers: []
Circumstances: []
Smell_Word: ”smell”
Qualities: []
Effects: [”fell suddenly dead”]
Sources: [”of brimstone”]
Evoked_Odorants: []
Other: []</td></tr>
</table>

To evaluate the applicability of question answering methods to the olfactory domain, we have performed evaluations mapping extracted causes and effects to annotations from the Odeuropa benchmark. The mapping results are presented in Table 4. In particular, we mapped the spans of ”causes” and ”effects” obtained with question answering to spans from different frame elements, manually annotated by annotators in the Odeuropa benchmark. For instance, on Figure 3, it is possible to notice that annotators have annotated ”fell suddenly dead” as effect while the QA model also identified ”fell suddenly dead” as the effect (in this case, the spans completely intersect).

Table 4: Cause/Effect Tests

Element	**From benchmark (same paragraph)**	**Mapped causes**	**Mapped effects**
Quality	1004	131	230
Source	1131	364	95
Odorant	80	24	7
Carrier	239	66	52
Circumstances	225	96	35
Effect	171	24	115
Other	732	87	101
All	3582	792	635

3.4 Characterizing Causes Using a Transfer Learning Approach

Our starting point is the Odeuropa text annotation benchmark described in Deliverable D3.2 [Tonelli et al., 2021b]. In particular, we are looking for the ‘Effects’ and ‘Circumstances’ which have been manually annotated. In Deliverable D3.4 [Hürriyetoğlu et al., 2023], we described a system that bundles these annotations in semantically-different short textual sentences, including:

- co-occurring events (Circumstances)
- causes or conditions for the smells (Circumstances)
- gestures or conditions triggered by the smells (Effects)
- the intention for which the smell was used to (e.g. removing a bad smell) (Circumstances or Effects)

We posit that distinguishing among these cases could improve the interpretation of these sentences and the overall data quality. In particular, we argue that it would allow us to precisely represent this part of the olfactory information in the EOKG as already foreseen in the development of the Odeuropa Ontology [Lisena et al., 2022]. For example, we would like to identify the specific purposes or consequences of some gestures.

In [Rebboud et al., 2023], a model for extracting fine-grained relationships from text – namely cause, intend, enable, prevent – has been trained on a synthetic-augmented dataset, showing good results in predicting the right relation. The model consists in a set of transformer layers that feed a softmax, for predicting the most probable relation – including a "no relation" tag. Our intuition is that the model can be successfully applied to those sentences in which Circumstances and/or Effects are present, to give us more information about them.

On the 425 English sentences from the benchmark including at least one *Effect* or *Circumstances*, two English speaking, but not native language speakers, performed a manual annotation choosing among the different classes {cause, intend, enable, prevent, no relation}, in addition to the null relation 0. These annotated data are used as ground truth for our experiments.

We then applied the model described above, as-is, to predict the relations from the sentence, the *effect* frame element or the *circumstances* frame element. For the latter two, we decided to additionally apply some different processing:

(1) prepend the frame element with a short text, namely "A smell happened", in order to form a complete sentence – e.g. "A smell happened " + "During the ballet". We have experimented with different texts to prepend without significantly changing the results;

(2) backtrack the frame element until the closer conjunction, adverb or preposition in the sentence, in order to help the machine to better understand the relation – e.g. the effect "he fainted" is backtracked until the previous adverb "so", to compose "so he fainted". This backtracking has not been applied if the frame was already starting with the searched part of speech (POS). A standard NLTK pipeline has been applied for detecting the POS;

(3) a combination of (1) and (2), so backtracking and prepending with a small text.

In Table 5, we report the results for each input. The support for each input is 425 for the full sentence, 205 for the effect, and 287 for the circumstances.

Table 5: Macro average of the results of the Transfer Learning Approach

Input	**Precision**	**Recall**	**F1 Score**	**Accuracy**
Full sentence	0.29	0.34	**0.24**	0.45
Effects	0.39	0.30	0.20	0.35
Effects (prepended by short text)	0.39	0.26	0.21	0.48
Effects (backtracked)	**0.50**	**0.32**	0.23	0.38
Effects (backtracked and prepended)	0.40	0.28	0.22	**0.49**
Circumstances	0.22	0.22	0.16	0.34
Circumstances (prepended by short text)	0.29	0.20	0.13	0.44
Circumstances (backtracked)	0.25	0.25	0.19	0.37
Circumstances (backtracked and prepended)	0.19	0.20	0.13	0.44

The F1 score is quite low in all experiments. However, we observe that when the *Circumstances* frame elements are used in input to the classifier, the results are even lower. This can be due to the ambiguity of such frame, which is indeed a container for quite different contents, from situation to status to events. It appears also that adding the prepending text is making the prediction harder. The best results are obtained with the full sentence (F1 score), and Effect backtracked to the conjunction, for which we included per-class detailed scores in Table 6. These results are unbalanced among classes, and suggest that different strategies should be applied to detect relations such as prevention.

Table 6: Detailed results for the two better performing input

Input	Class	Precision	Recall	F1 Score	Support
Full Sentence	0	0.38	0.11	0.16	114
	cause	0.49	0.83	0.62	197
	enable	0.33	0.02	0.03	60
	intend	0.00	0.00	0.00	35
	prevent	0.25	0.74	0.38	19
Effects (backtracked)	0	0.55	0.13	0.21	45
	cause	0.55	0.62	0.58	101
	enable	0.33	0.07	0.12	27
	intend	1.00	0.04	0.08	24
	prevent	0.08	0.75	0.15	8

3.5 Discussion

In the previous sections, we have presented a number of approaches that can be utilized for extracting context for olfactory mentions in text. In this section, we present several examples (see Table 7, Table 8 and Table 9) that provide a qualitative overview for supervised, question answering and transfer learning methods.

We observe that causes extracted by question answering methods are regularly mapped to smell sources. Effects extracted by question answering and supervised methods might differ in spans, but are also regularly mapped from the context point of view. The transfer learning approach demonstrates predicted causality relations on similar spans as well.

Table 7: Context Extraction Example 1

Supervised approach
Sentence: ”And He would really have to take care where He walked , because the place was in a really terrible state , and He would have to keep his hand on the halter because horses , even stallions , were most foolishly upset at the scent of lion .”
Smell_word: ”scent”
Smell_Source: ”of lion”
Perceiver: ”horses|stallions”
Effect: ”most foolishly upset”

Continued on next page

Table 7: Context Extraction Example 1 (Continued)

QA approach Sentence: ”And He would really have to take care where He walked , because the place was in a really terrible state , and He would have to keep his hand on the halter because horses , even stallions , were most foolishly upset at the scent of lion .” Cause: ”lion” Cause_start: 232 Cause_end: 236 Cause_score: 0.8108424544334412 Effect: ”horses , even stallions , were most foolishly upset” Effect_start: 164 Effect_end: 215 Effect_score: 0.17993047833442688
Transfer learning approach Sentence: ”And he would really have to take care where he walked, because the place was in a really terrible state, and he would have to keep his hand on the halter because horses, even stallions, were most foolishly upset at the scent of lion.” Smell_word: ”scent” Effect: ”most foolishly upset” Predicted relation (full sentence): cause (score: 0.99, correct) Predicted relation (backtracked effect): cause (score: 0.96, correct)

Table 8: Context Extraction Example 2

Supervised approach Sentence: ”Perhaps all these remedies may be good, saith the Grand Mother but they are not for our turns; for alas a day, the very smell of salve makes her fall into a swoon; neither can she suffer the least motion of sucking, for the very pain bereaves her of her senses.” Smell_Word: ”smell” Smell_Source: ”of salve” Effect: ”makes her fall into a swoon”
QA approach Sentence: ”Perhaps all these remedies may be good, saith the Grand Mother but they are not for our turns; for alas a day, the very smell of salve makes her fall into a swoon; neither can she suffer the least motion of sucking, for the very pain bereaves her of her senses.” Cause: ”salve” Cause_start: 134 cause_end: 139 Cause_score: 0.7884625196456909 Effect: ”makes her fall into a swoon” Effect_start: 140 Effect_end: 167 Effect_score: 0.3024158477783203

Continued on next page

Table 8: Context Extraction Example 2 (Continued)

<table>
<tr><td>Transfer learning approach
Sentence: ”Perhaps all these remedies may be good, saith the Grand Mother but they are not for our turns; for alas a day, the very smell of salve makes her fall into a swoon; neither can she suffer the least motion of sucking, for the very pain bereaves her of her senses.”
Effect: ”makes her fall into a swoon”
Predicted relation (full sentence): cause (score: 0.96, correct)
Predicted relation (backtracked effect): cause (score: 0.99, correct)</td></tr>
</table>

Table 9: Context Extraction Example 3

<table>
<tr><td>Supervised approach
Sentence: ”The effect of the foul odors of the ship may be combatted by the use of aromatic electuaries, ” which comfort the heart, the brain and the stomach.”
Smell_Word: ”odors|aromatic|aromatic”
Smell_Source: ”of the ship”
Quality: ”foul|aromatic|aromatic”
Effect: ”which comfort the heart, the brain and the stomach”</td></tr>
<tr><td>QA approach
Sentence: ”The effect of the foul odors of the ship may be combatted by the use of aromatic electuaries, \”\”\”\” which comfort the heart , the brain and the stomach.”
Cause: ”the ship”
Cause_start: 32
Cause_end: 40
Cause_score: 0.7127965092658997
Effect: ”comfort the heart, the brain and the stomach”
Effect_start: 106
Effect_end: 151
Effect_score: 0.4783854782581329</td></tr>
<tr><td>Transfer learning approach
Sentence: ”The effect of the foul odors of the ship may be combatted by the use of aromatic electuaries, which comfort the heart, the brain and the stomach.
Effect: – absent in the annotation –
Predicted relation (full sentence): enable (score: 0.99, should be intend)
Predicted relation (backtracked effect): n.a.</td></tr>
</table>

4 Semantic Annotation for Enhancing Context Modeling

Since Odeuropa Task 4.4 aims at harvesting context related to olfactory cultural heritage, we introduce different approaches allowing for obtaining additional contextual information from text snippets. In particular, in this section we present i) a zero-shot classification approach for enriching textual content with terms from the Odeuropa taxonomy and ii) the use of the Wikifier tool with embedded Odeuropa vocabularies for semantic annotation.

4.1 Zero-Shot Approach

Zero-shot text classification is an NLP task where a model is trained on a set of labeled examples but is then able to classify new examples from previously unseen classes [HuggingFace.co, 2023]. For zero-shot classification in the Odeuropa project, we have used the ”facebook/bart-large-mnli” model [Lewis et al., 2020] and the Odeuropa taxonomies (with over 700 concepts from Olfactory objects, Noses, Gestures and Fragrant spaces vocabularies) as set of labeled examples.

Zero-shot annotation allows obtaining additional contextual information and text enriching with olfactory related context. Table 11 presents a zero-shot annotation example at sentence level, where a sentence is enriched with concepts defined in some of the Odeuropa taxonomies. Table 12 shows a zero-shot annotation example, where extracted ”cause” from a sentence is enriched with concepts from Odeuropa taxonomies.

Table 10 presents an evaluation of zero-shot annotation with Accuracy Top 1 and Accuracy Top 3 method (in comparison to Odeuropa benchmark gold standard). In this experiment, we have used 316 sentences from Odeuropa benchmark and compared zero-shot annotations on these sentences with gold standard annotations stored in the European Olfactory Knowledge Graph. We have tried four different experimental settings described below:

- Accuracy top 1 and Threshold 0.05. Accuracy top 1 presents the metric, where the correct label is the top concept predicted. In this setting, zero-shot predicted concepts should have confidence score over 0.05.
- Accuracy top 3 and Threshold 0.05. Accuracy top 3 presents the metric, where the correct label is among the top 3 labels predicted. In this setting, zero-shot predicted concepts should have confidence score over 0.05.
- Accuracy top 1 and Threshold 0.03. Accuracy top 1 presents the metric, where the correct label is top concept predicted. In this setting, zero-shot predicted concepts should have confidence score over 0.03.
- Accuracy top 3 and Threshold 0.03. Accuracy top 3 presents the metric, where the correct label is among the top 3 labels predicted. In this setting, zero-shot predicted concepts should have confidence score over 0.03.

Table 10: Zero-shot Annotation Evaluation

Evaluation method	Threshold	Accuracy
Accuracy top 1	0.05	0.646302251
Accuracy top 3	0.05	0.768488746
Accuracy top 1	0.03	0.697749196
Accuracy top 3	0.03	0.848874598

Table 11: Zero-Shot Annotation Example (Sentence Level)

Sentence: ”Prawns and Shrimps, if they are hard and stiff, of a pleasant scent, and their tails turned strongly inward, are new; but if they are limber, their colour faded, of a faint smell, and feel slimy, they are stale.”
Linked Entity (EOKG): ¡http://data.odeuropa.eu/vocabulary/olfactory-objects/470¿
Linked Label (EOKG): ”Prawn”
Zero-Shot Annotations:

Continued on next page

Table 11: Zero-Shot Annotation Example (Sentence Level) (Continued)

Annotation:	"Prawn"
Score:	0.148075745
Annotation:	"Animal product"
Score:	0.100027852
Annotation:	"Seafood"
Score:	0.062754557

Table 12: Zero-Shot Annotation Example (Cause Level)

Sentence: "I could tell by their smell that they contained either spirits of gin or brandy, much above proof"	
Cause (QA): "spirits of gin or brandy"	
Zero-Shot Annotations:	
Annotation:	"Alcohol"
Score:	0.848037243
Annotation:	"Substance"
Score:	0.02533154
Annotation:	"Person"
Score:	0.020983342

4.2 Wikifier Approach

The JSI Wikifier [Brank et al., 2017] is a web service that takes a text document as input and annotates it with links to relevant Wikipedia concepts. JSI Wikifier allows to enrich the text with appropriate semantic concepts related to smell.

Within Odeuropa project, we have integrated Odeuropa vocabularies into the general Wikifier text processing pipeline and as an outcome of Odeuropa project, all Wikifier users are now able to semantically annotate text not only with Wikipedia concepts, but also with Odeuropa vocabularies.

Figure 2: Wikifier Results (Main Page, Text 1)

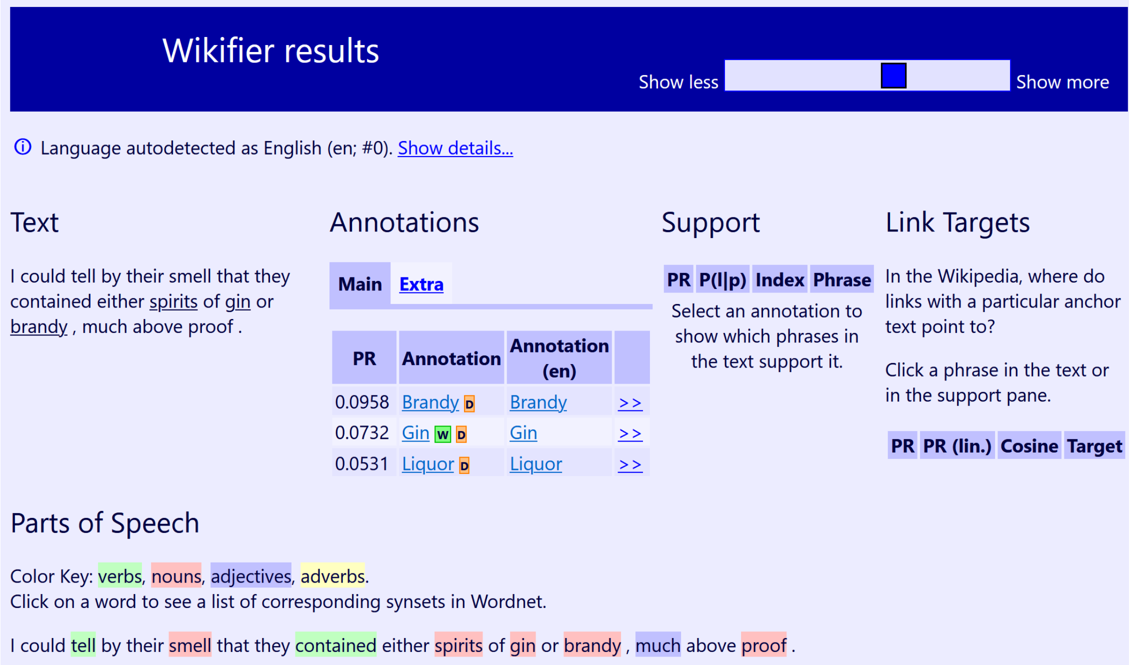
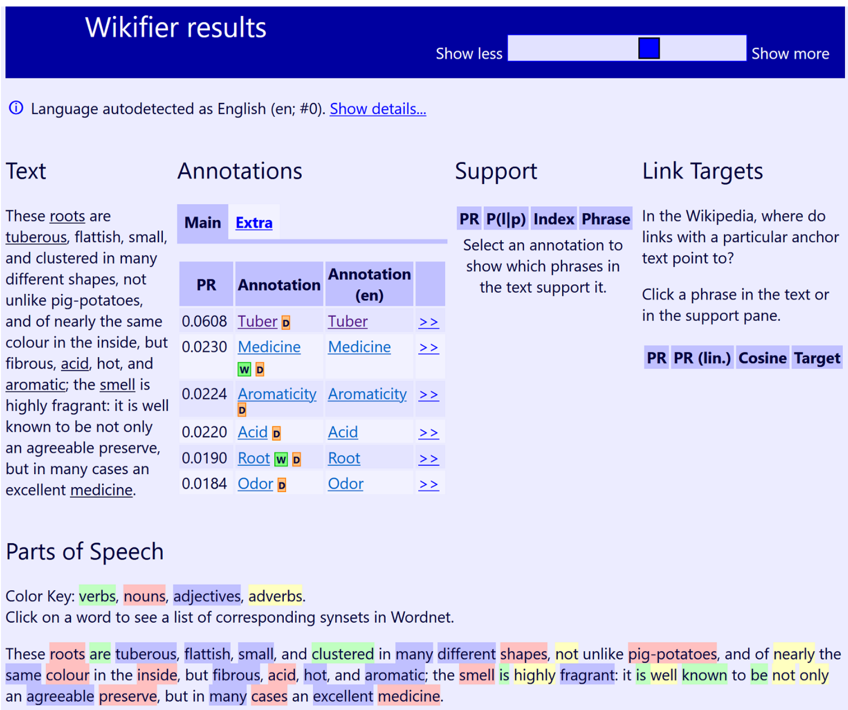


Figure 3: Wikifier Results (Main Page, Text 2)



Figures 2, 3 and 4 present examples of text annotated with Wikipedia concepts. Wikifier Results (Main) show the Wikipedia annotations, while Wikifier Results (Extra) provide annotations with Odeuropa taxonomies.

Semantic annotation with Wikifier tool (including wikification process and semantic enrichment with embedded Odeuropa vocabularies) has been utilized in Odeuropa project as part of Task T4.4 "Harvesting Context Related to Olfactory Cultural Heritage" and Task T3.4 "Multilingual emotion recognition". We have used Wikifier for providing additional semantic context in the task of mining emotions related to smell [Massri et al., 2022]. In [Massri et al., 2022], we have applied the Wikifier tool for annotating texts using Wikipedia concepts related to specific emotions and olfactory terms occurring in fairy tales. Wikifier is therefore a tool with semantic annotation functionalities that enables to bring contextual information from Wikipedia and Odeuropa vocabularies.

Figure 4: Wikifier Results (Extra)

5 Application to Historical Analysis

In this section, we present possible ways to analyze causes and effects from the perspective of cultural heritage experts. We introduce an approach for grouping causes and effects by periods of 50 and 100 years and observe the differences in time and over different data sources. Following that, we suggest a number of layers (where layer represents a group of causes/effects similar to each other) for validation to cultural heritage experts (Section **??**). Expert historians have assessed the bottom-up grouping results obtained from the data and provided the possible naming for the layers. In Section 5.2, we show the visualizations of causes and effects independently of historical periods, using Ontogen tool[Fortuna et al., 2007] for semi-automatic ontology construction.

5.1 Grouping Causes and Effects

The goal of cause and effects grouping is to perform exploratory historical analysis of smell-related causes and effects extracted from different sources (Project Gutenberg data collection [Project Gutenberg, 2023], Old Bailey data collection [Clarin: Old Bailey, 2023], Royal Society corpus [Clarin: Royal Society, 2023]), group causes and effects by clusters (layers) and observe how clustering results change over time.

The Methodology for historical data analysis has been based on the following steps:

(1) From each data source (Project Gutenberg corpus, Old Bailey corpus, Royal Society corpus) select a sample of sentences for analysis (1000 sentences);

(2) Identify causes/effects for the sentences using QA technique (on top of ”mbartolo/roberta-large-synqa-ext” model).

(3) Semantically annotate the results with Odeuropa taxonomies (Olfactory objects vocabulary, Gestures vocabulary, Fragrant spaces vocabulary and Noses vocabulary) using zero-shot classification task (with ”facebook/bart-large-mnli” model).

(4) Cluster the obtained data with K-means clustering (for different periods and different k values: 3,5,7,10). Tested time periods:

   – 1600-1649
   – 1650-1699
   – 1700-1749
   – 1750-1799
   – 1800-1849
   – 1850-1899
   – 1900-1925

   and

   – 1600-1699
   – 1700-1799
   – 1800-1899
   – 1900-1925

(5) Observe the results over time and utilize expert opinion for definition of potential layers names.

We have used K-means clustering [MacQueen, 1967] as one of the simplest and well-known unsupervised machine learning algorithms. K-means tends to group similar data points together and discover underlying patterns by looking at the predefined number of groups or clusters (k) in the data. The cluster or group of data point aggregates similar data points together. In the area of causality detection, K-means allows for identifying groups of similar causes or groups of similar effects for olfactory related texts.

Based on the observations of groupings for causes and effects, we have identified the following potential layers (potential groups of smell-related causes/effects) presented in Table 13.

The following Tables present different groups of causes and effects extracted from Gutenberg, Old Bailey and Royal Society corpora within 50-years and 100-years periods, while the following Figures display the graphical representation of grouping for causes or effects – the distinct groups are observed within specific historical periods for different corpora. In the details, Table 14 shows 3 clusters of causes extracted from Project Gutenberg data collection in different periods of 50 years. It is possible to observe the clusters specifically related to ”Flowers” layer, ”Animals” layer, ”Filth/Pollution” layer etc.

Table 13: Layers for Causes and Effects

Layer	Found in Causes	Found in Effects
Animals	+	+
Person/Body	+	+
Flowers	+	+
Fragrance	+	+
Alcohol/Drink	+	
Fire	+	+
Smoke	+	
Filth/Pollution	+	+
Kitchen	+	
Religious	+	+
Prison	+	
Substance	+	
Material	+	
Chemical process	+	
Radiation	+	
Nature	+	
Science	+	
Materia Medica	+	+
Occupational	+	
Chemistry	+	+
Commerce	+	+
(Infanticide) Trials	+	
Public Health	+	
Theft		+
Coining offences		+

Table 14: Gutenberg Causes, k=3, by 50 years

Clusters	Period
Cluster 1 (Flowers) folded,zamorin,god,makes,drooping,field,blossoms,orange,flowers,flower Cluster 2 (Kitchen) coal,kitchen,religion,body,lovage,food,drink,blood,animal,person Cluster 3 (Filth/Pollution) dew,matter,waste,beastly,stink,wretched,excrement,flesh,putrid,filth	1800-1849
Cluster 1 water,rose,wind,tobacco,wood,filth,person,fragrance,plant,air Cluster 2 (Flowers) lilac,wild,petals,faded,trees,passion,plant,blossoms,flowers,flower Cluster 3 (Animals) wolf,wolves,fishes,sheep,creature,mammal,dog,fish,product,animal	1850-1899
Cluster 1 (Nature) footsteps,food,fresh,tufts,blue,small,summer,orange,flowers,flower Cluster 2 (Person) jerome,healthy,strong,youthful,story,life,man,woman,hair,person Cluster 3 plant,alcoholic,fragrance,substance,chemical,breath,sweet,drink,filth,animal	1900-1925

Figure 5 shows how the points from 3 clusters of causes in Project Gutenberg data for olfactory mentions in historical period for years 1850-1899 are grouped on plot. Non-intersecting points from different clusters can be observed on this figure.

Figure 5: Gutenberg Causes, k=3, 1850-1899

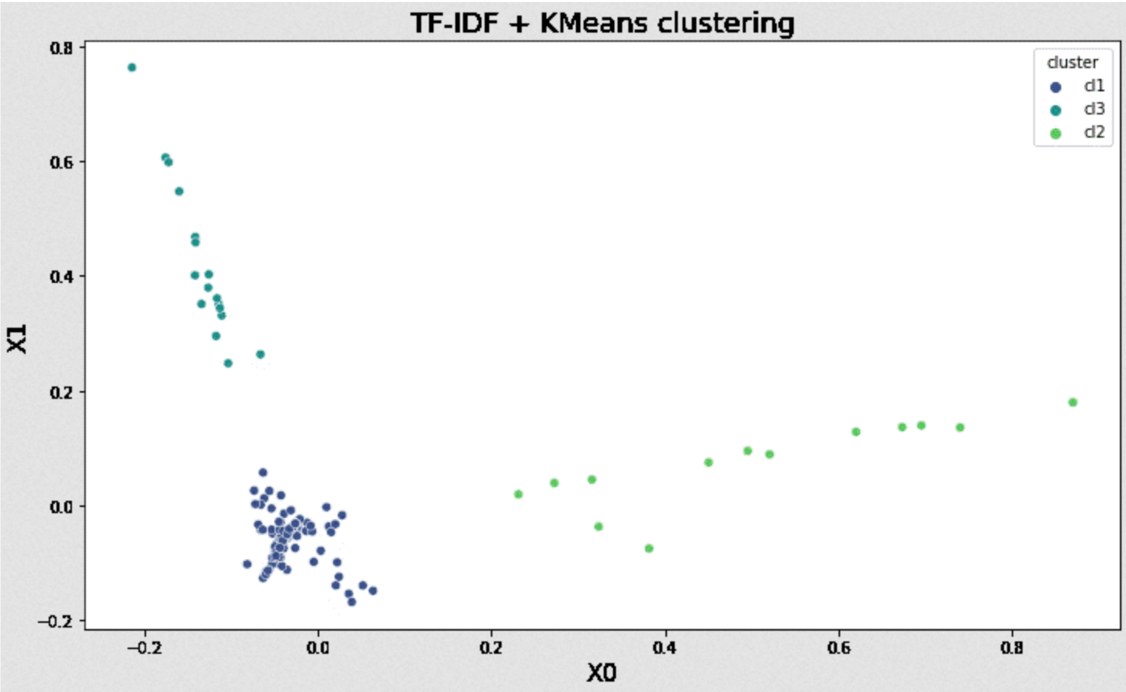


Table 15 presents 3 clusters of causes extracted from Old Bailey data collection in different periods of 50 years. It is possible to observe the clusters specifically related to ”Alcohol” layer, ”Chemistry” layer, ”Substance” layer etc. Since Old Bailey corpus reflects the content and materials from London's Central Criminal Court, in the grouping process we can observe cluster of words related to prison, prisoner etc.

Table 15: Old Bailey Causes, k=3, by 50 years

Cluster	Period
Cluster 1 substance,chemical,woman,nose,tobacco,snuff,fire, animal,candle,person Cluster 2 filth,extremely,woman,poison,bodily,stomach,fluid, enough,breath,body Cluster 3 (Alcohol) fragrance,fumigation,flame,frequently,drank,much,drinking, brandy,alcoholic,drink	1750-1799
Cluster 1 drink,candle,fragrance,tobacco,animal,box,snuff, scent,person,fire Cluster 2 (Chemistry) ozone,pot,opium,phosphorus,acid,vinegar,tarpit,chemical, substance,turpentine	1800-1849

Cluster 3 (Prison/Alcohol) spirits,spiritous,brandy,person,prisoner,prison,liqueur, liquor,alcoholic,drink	
Cluster 1 (Commerce) premises,man,mr,fragrance,gas,smoke,varnish, person,factory,wind Cluster 2 (Alcohol) lobelia,beer,boiler,fullery,alcoholic,drinking,drunk, fluid,bodily,drink Cluster 3 (Substance/Material) acid,sulphuric,naphtha,drug,sulphur,gunpowder,copper,paraffin, substance,chemical	1850-1899
Cluster 1 draymen,door,dog,dirty,digested,carbolic,chloroform,paraffin, substance,chemical Cluster 2 (Alcohol) breath,influence,bodily,fluid,drunkenness,alcohol,drunk, drinking,alcoholic,drink Cluster 3 holding,woman,blood,vomit,smell,pepper,person, house,room,gas	1900-1925

Figure 6 show the points from 3 clusters of detected causes in Old Bailey data collection. The analyzed time period covers years 1850-1899.

Figure 6: Old Bailey Causes, k=3, 1850-1849

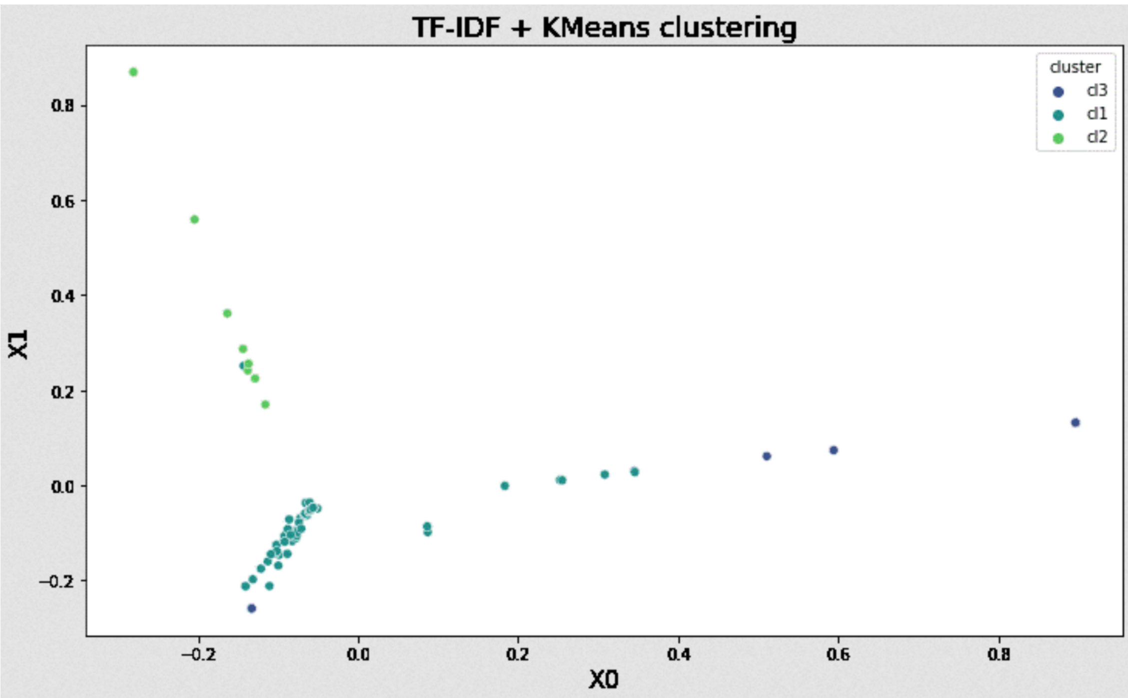


Table 16 presents 3 clusters of causes extracted from Royal Society data collection in different periods of 50 years. It is possible to observe the clusters specifically related to ”Substance” layer, ”Material” layer, ”Radiation”, ”Chemical Process” layer etc.

Table 16: Royal Society Causes, k=3, by 50 years

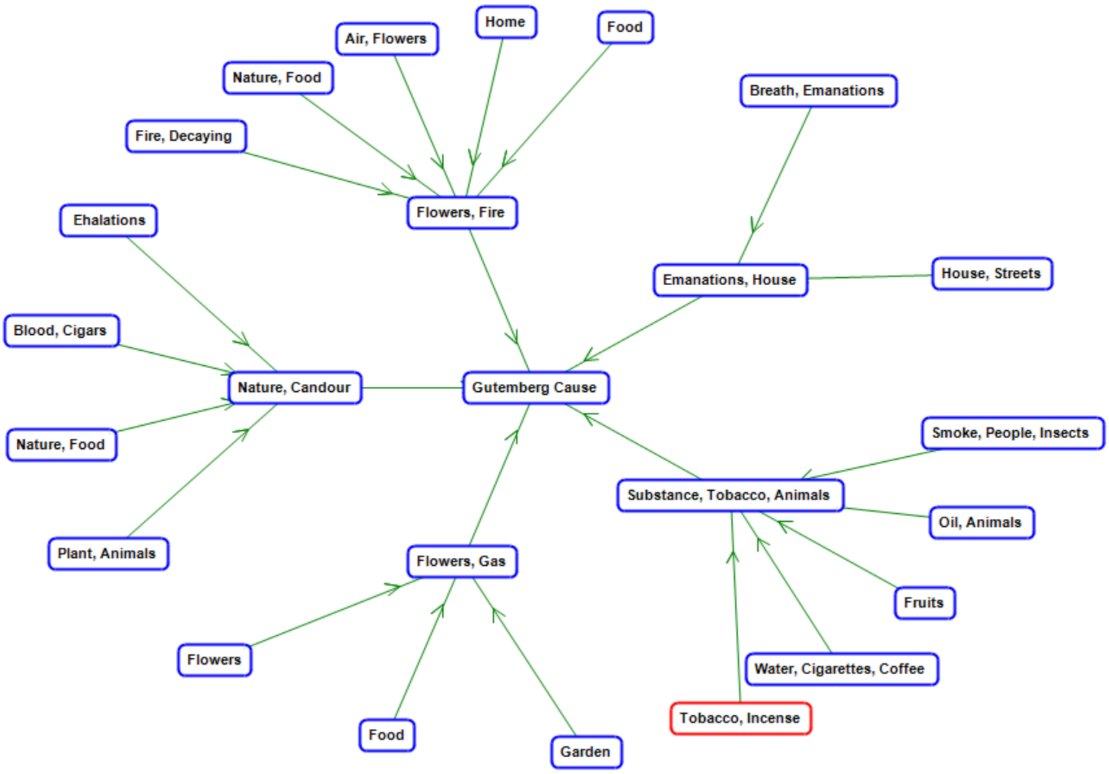
Cluster	**Period**
Cluster 1 water,fumigation,root,alcoholic,filth,fruit,animal,drink,excrement,fig Cluster 2 (Flowers) juice,trees,palm,tuberose,tree,plants,leaves,leaf,vegetable,plant Cluster 3 (Substance) sulphureous,matter,effluviums,filth,salt,air,sulphur, pollution,chemical,substance	1650-1699
Cluster 1 (Substance) ethide,sulphur,carbonate,acetic,chlorine,acid,ethylic,boric,substance,chemical Cluster 2 furze,water,small,fragrance,gas,products,dilligence,acid,pot,ozone Cluster 3 (Pollution/Chemical Process) gases,erupted,entirely,microbial,impurity,impregnation, pollution,putrefaction,filth,excrement	1850-1899
Cluster 1 (Radiation) person,oblite,summer,air,radioactive,lightning,helium,solution,thorium,emanation Cluster 2 (Material/Radiation) hydrogen,hydrocarbon,chloramino,mercury,acid,durian, formaldehyde,radium,substance,chemical Cluster 3 ema,elements,element,durian,dried,dissolved,disintegration,diacetyl,excess,radiation	1900-1925

In addition, in Annex 6, we present a number of tables and figures for analysis of historical causality development in texts with olfactory mentions.

5.2 Visualizing Groups of Causes and Circumstances using OntoGen

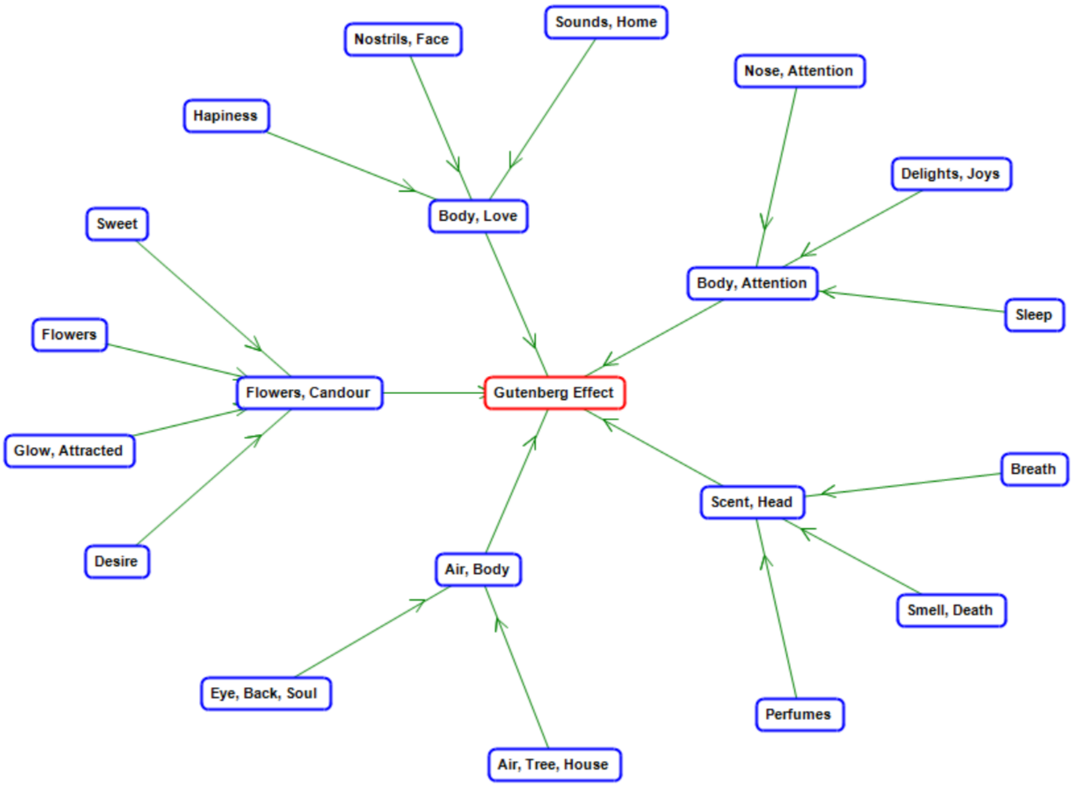
The OntoGen system [Fortuna et al., 2007] is targeted at semi-automatic ontology construction and integrates machine learning and text mining algorithms into an efficient user interface, lowering the entry barrier for users who are not professional ontology engineers. The main features of the systems include unsupervised and supervised methods for concept suggestion and concept naming, as well as ontology and concept visualization. We have used the Ontogen tool for the analysis of causes and effects in different data collections (without splitting data into historical time periods). The analysis of causes extracted from Project Gutenberg data collection with Ontogen tool is displayed in Figure 7.

Figure 7: Analysis of Gutenberg Causes with Ontogen Tool



Analysis of effects extracted from Project Gutenberg data collection with Ontogen tool is provided on Figure 8.

Figure 8: Analysis of Gutenberg Effects with Ontogen Tool



Annex 6 provides additional analysis of causes and effects extracted from Old Bailey and Royal Society data collections.

The analysis of overall causes and effects demonstrates that splitting data into different historical periods (as described in Section 5) might provide additional usefulness for domain experts.

6 Conclusion

In this deliverable, we have presented the related work, methodology and applications for context modeling for olfactory mentioned in text.

In order to tackle the Task 4.4, the WP4 partners developed a number of machine learning approaches allowing to discover causal relationships describing what led to a situation where odour appears and what kind of consequences were caused by an appearance of odour. In particular, the supervised approach for extracting effects and circumstances (based on Odeuropa benchmark annotated data), question answering approach for extracting causes and effects and transfer learning approach characterizing causes. The developed approaches have been assessed in quantitative way.

In addition, state-of-the-art annotation methods have been used for enriching extracted causal textual data with Odeuropa olfactory vocabularies concepts. The enriched content has been used for modeling olfactory causality development over time.

The obtained causality extraction, enrichment and grouping results present a basis for integration of context into European Olfactory Knowledge Graph.

References

[Altenberg, 1984] Altenberg, B. (1984). Causal linking in spoken and written english. *Studia Linguistica*, 38:20–69.

[Bartolo et al., 2020a] Bartolo, M., Roberts, A., Welbl, J., Riedel, S., and Stenetorp, P. (2020a). Beat the ai: Investigating adversarial human annotation for reading comprehension. *Transactions of the Association for Computational Linguistics*, 8:662–678.

[Bartolo et al., 2020b] Bartolo, M., Roberts, A., Welbl, J., Riedel, S., and Stenetorp, P. (2020b). Beat the AI: Investigating adversarial human annotation for reading comprehension. *Transactions of the Association for Computational Linguistics*, 8:662–678.

[Bartolo et al., 2021] Bartolo, M., Thrush, T., Jia, R., Riedel, S., Stenetorp, P., and Kiela, D. (2021). Improving question answering model robustness with synthetic adversarial data generation. *ArXiv*, abs/2104.08678.

[Blanco et al., 2008] Blanco, E., Castell, N., and Moldovan, D. I. (2008). Causal relation extraction. In *International Conference on Language Resources and Evaluation*.

[Brank et al., 2017] Brank, J., Leban, G., and Grobelnik, M. (2017). Annotating documents with relevant wikipedia concepts. In *Informatica*.

[Caruana, 1993] Caruana, R. (1993). Multitask learning: A knowledge-based source of inductive bias1. In *Proceedings of the Tenth International Conference on Machine Learning*, pages 41–48. Citeseer.

[Caruana, 1997] Caruana, R. (1997). Multitask learning. *Machine learning*, 28(1):41–75.

[Clarin: Old Bailey, 2023] Clarin: Old Bailey (2023). Old bailey corpus. https://fedora.clarin-d.uni-saarland.de/oldbailey. Accessed: 2023-04-20.

[Clarin: Royal Society, 2023] Clarin: Royal Society (2023). Royal society corpus. https://fedora.clarin-d.uni-saarland.de/rsc/. Accessed: 2023-04-20.

[Craswell et al., 2020] Craswell, N., Mitra, B., Yilmaz, E., Campos, D. F., and Voorhees, E. M. (2020). Overview of the trec 2020 deep learning track. *ArXiv*, abs/2102.07662.

[Dasgupta et al., 2018] Dasgupta, T., Saha, R., Dey, L., and Naskar, A. (2018). Automatic extraction of causal relations from text using linguistically informed deep neural networks. In *SIGDIAL Conference*.

[Devlin et al., 2019] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*, pages 4171–4186.

[Fortuna et al., 2007] Fortuna, B., Grobelnik, M., and Mladenić, D. (2007). Ontogen: Semi-automatic ontology editor. In *Human Interface and the Management of Information. Interacting in Information Environments, Symposium on Human Interface 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings, Part II*, pages 309–318.

[Garcia, 1997] Garcia, D. (1997). Coatis, an nlp system to locate expressions of actions connected by causality links. In *International Conference Knowledge Engineering and Knowledge Management*.

[Hendrickx et al., 2010] Hendrickx, I., Kim, S. N., Kozareva, Z., Nakov, P., Séaghdha, D. Ó., Padó, S., Pennacchiotti, M., Romano, L., and Szpakowicz, S. (2010). Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. In *International Workshop on Semantic Evaluation*.

[HuggingFace.co, 2023] HuggingFace.co (2023). Zero shot classification. https://huggingface.co/tasks/zero-shot-classification. Accessed: 2023-05-14.

[Hürriyetoğlu et al., 2023] Hürriyetoğlu, A., Tonelli, S., Menini, S., Brank, J., Tullet, W., and van Erp, M. (2023). Odeuropa text processing system with historical corpus analysis version 2. Technical Report Deliverable D3.4, Odeuropa.

[Ittoo and Bouma, 2013] Ittoo, A. and Bouma, G. (2013). Minimally-supervised learning of domain-specific causal relations using an open-domain corpus as knowledge base. *Data Knowl. Eng.*, 88:142–163.

[Jin et al., 2020] Jin, X., Wang, X., Luo, X., Huang, S., and Gu, S. (2020). Inter-sentence and implicit causality extraction from chinese corpus. *Advances in Knowledge Discovery and Data Mining*, 12084:739 – 751.

[Khoo et al., 1998] Khoo, C. S. G., Kornfilt, J., Oddy, R. N., and Myaeng, S.-H. (1998). Automatic extraction of cause-effect information from newspaper text without knowledge-based inferencing. *Literary and Linguistic Computing*, 13:177–186.

[Klie et al., 2018] Klie, J.-C., Bugert, M., Boullosa, B., de Castilho, R. E., and Gurevych, I. (2018). The inception platform: Machine-assisted and knowledge-oriented interactive annotation. In *International Conference on Computational Linguistics*.

[Kyriakakis et al., 2019] Kyriakakis, M., Androutsopoulos, I., i Ametll'e, J. G., and Saudabayev, A. (2019). Transfer learning for causal sentence detection. *ArXiv*, abs/1906.07544.

[Lewis et al., 2020] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., and Zettlemoyer, L. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.

[Lisena et al., 2022] Lisena, P., Schwabe, D., van Erp, M., Troncy, R., Tullet, W., Leemans, I., Marx, L., and Ehrich, S. (2022). Capturing the Semantics of Smell: The Odeuropa Data Model for Olfactory Heritage Information. In *19^{th} Extended Semantic Web Conference (ESWC), Resource Track*, Hersonissos, Greece.

[MacQueen, 1967] MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In Cam, L. M. L. and Neyman, J., editors, *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. University of California Press.

[Manning et al., 2014] Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., and McClosky, D. (2014). The stanford corenlp natural language processing toolkit. In *Annual Meeting of the Association for Computational Linguistics*.

[Martínez-Cámara et al., 2017] Martínez-Cámara, E., Shwartz, V., Gurevych, I., and Dagan, I. (2017). Neural disambiguation of causal lexical markers based on context. In *International Conference on Computational Semantics*.

[Massri et al., 2022] Massri, M. B., Novalija, I., Mladenić, D., Brank, J., Graça da Silva, S., Marrouch, N., Murteira, C., Hürriyetoğlu, A., and Šircelj, B. (2022). Harvesting context and mining emotions related to olfactory cultural heritage. *Multimodal Technologies and Interaction*, 6(7).

[Menini et al., 2022a] Menini, S., Paccosi, T., Tonelli, S., Van Erp, M., Leemans, I., Lisena, P., Troncy, R., Tullett, W., Hürriyetoğlu, A., Dijkstra, G., Gordijn, F., Jürgens, E., Koopman, J., Ouwerkerk, A., Steen, S., Novalija, I., Brank, J., Mladenic, D., and Zidar, A. (2022a). A multilingual benchmark to capture olfactory situations over time. In *Proceedings of the 3rd Workshop on Computational Approaches to Historical Language Change*, pages 1–10, Dublin, Ireland. Association for Computational Linguistics.

[Menini et al., 2022b] Menini, S., Paccosi, T., Tonelli, S., van Erp, M., Leemans, I., Lisena, P., Troncy, R., Tullett, W., Hürriyetoğlu, A., Dijkstra, G., Gordijn, F., Jürgens, E., Koopman, J., Ouwerkerk, A., Steen, S., Novalija, I., Brank, J., Mladenić, D., and Zidar, A. (2022b). A multilingual benchmark to capture olfactory situations over time. In *Workshop on Computational Approaches to Historical Language Change*.

[Pakray and Gelbukh, 2014] Pakray, P. and Gelbukh, A. (2014). An open-domain cause-effect relation detection from paired nominals. In *Mexican International Conference on Artificial Intelligence*.

[Pechsiri et al., 2006] Pechsiri, C., Kawtrakul, A., and Piriyakul, R. (2006). Mining causality knowledge from textual data. In *Artificial Intelligence and Applications*.

[Project Gutenberg, 2023] Project Gutenberg (2023). Project gutenberg. https://www.gutenberg.org. Accessed: 2023-04-20.

[Qi et al., 2020] Qi, P., Zhang, Y., Zhang, Y., Bolton, J., and Manning, C. D. (2020). Stanza: A python natural language processing toolkit for many human languages. *ArXiv*, abs/2003.07082.

[Radinsky et al., 2012] Radinsky, K., Davidovich, S., and Markovitch, S. (2012). Learning causality for news events prediction. *Proceedings of the 21st international conference on World Wide Web*.

[Rebboud et al., 2023] Rebboud, Y., Lisena, P., and Troncy, R. (2023). Prompt-based Data Augmentation for Semantically-precise Event Relation Classification. In *Semantic Methods for Events and Stories workshop (SEMMES)*, Heraklion, Greece.

[Sorgente et al., 2013] Sorgente, A., Vettigli, G., and Mele, F. (2013). Automatic extraction of cause-effect relations in natural language text. In *DART@AI*IA*.

[Tonelli and Menini, 2021a] Tonelli, S. and Menini, S. (2021a). Framenet-like annotation of olfactory information in texts. *Proceedings of the 5th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*.

[Tonelli and Menini, 2021b] Tonelli, S. and Menini, S. (2021b). FrameNet-like annotation of olfactory information in texts. In *Proceedings of the 5th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 11–20, Punta Cana, Dominican Republic (online). Association for Computational Linguistics.

[Tonelli et al., 2021a] Tonelli, S., Menini, S., Sinem Tekiroglu, S., and Novalija, I. (2021a). Annotation scheme and multilingual taxonomy. Technical Report Deliverable D3.1, Odeuropa.

[Tonelli et al., 2021b] Tonelli, S., Menini, S., van Erp, M., Troncy, R., and Novalija, I. (2021b). Multilingual historical corpora and annotated benchmarks. Technical Report Deliverable D3.2, Odeuropa.

[van der Goot et al., 2021] van der Goot, R., Üstün, A., Ramponi, A., Sharaf, I., and Plank, B. (2021). Massive choice, ample tasks (MaChAmp): A toolkit for multi-task learning in NLP. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, pages 176–197, Online. Association for Computational Linguistics.

[Vasiliev, 2020] Vasiliev, Y. (2020). Natural language processing with python and spacy: A practical introduction. In *San Francisco : No Starch Press*.

[Wang et al., 2016] Wang, L., Cao, Z., de Melo, G., and Liu, Z. (2016). Relation classification via multi-level attention cnns. In *Annual Meeting of the Association for Computational Linguistics*.

[Yang et al., 2021] Yang, J., Han, S. C., and Poon, J. (2021). A survey on extraction of causal relations from natural language text. *Knowledge and Information Systems*, 64:1161 – 1186.

[Zhang et al., 2018] Zhang, Y., Qi, P., and Manning, C. D. (2018). Graph convolution over pruned dependency trees improves relation extraction. In *Conference on Empirical Methods in Natural Language Processing*.

[Zhao et al., 2016] Zhao, S., Liu, T., Zhao, S., Chen, Y., and Nie, J.-Y. (2016). Event causality extraction based on connectives analysis. *Neurocomputing*, 173:1943–1950.

Annex I

Table 17 presents clusters of causes extracted from Gutenberg data collection in different periods of 100 years.

Clusters	k	Period
Cluster 1 lovage ,body,flowers,tobacco,fragrance,perfume,tree,fire,chemical,substance Cluster 2 (Animals) steaks,dog,anchovy,rat,rats,muscovy,mutton,meat,product,animal Cluster 3 (Pollution/Filth) putrefaction,bad,cardboard,dust,dirt, pollution,carcass,excrement,filthy,filth	3	1600-1699
Cluster 1 (Science) gone,yellow,filth,pollution,sulphur,vive,sulphure,quicke,mountain,brimstone Cluster 2 (Fragrance/Perfume) odor,censer,sanctity,ropery,incense,rose,brere,perfumes,fragrance,perfume Cluster 3 (Materia Medica) gums,gum,greener,great,good,gone,heat,yellow,brimstone,fire Cluster 4 (Pollution/Filth) excrements,putrefaction,bad,cardboard,dirt,dust,carcass, excrement,filthy,filth Cluster 5 air,product,lovage,body,flowers,tobacco,tree,animal,chemical,substance	5	
Cluster 1 root,stable,vegetable,tobacco,wood,water,drink,filth,animal,plant Cluster 2 (Occupational) garbage,stable,excellent,solid,smeared,lubricating,palm, hop,petroleum,oil Cluster 3 (Religious) finger,censing,perfect,holy,charity,elder,convenience,lovage,seed,sage Cluster 4 stall,twigs,sandal,noxious,poison,treed,fruit,nut,trees,tree Cluster 5 fragrance,fountains,fountain,gale,elated,living,faded,grass,flowers,flower Cluster 6 (Substance) resin,corol,nitre,anthoxanthum,gum,turpentine,vinegar,acid,substance,chemical	7	1700-1799

Cluster 7 fumigation,fountain,forget,food,foetida,abstract,absence, adulteration,wingless,idoform		
Cluster 1 (Pollution/Filth) waste,beastly,substance,stink,wretched,dust,flesh,putrid,excrement,filth Cluster 2 (Flowers) petals,drooping,field,faded,orange,passion,plant,blossoms,flowers,flower Cluster 3 product,water,food,drink,blood,plant,fragrance,air,person,animal	3	1800-1899
Cluster 1 lovage,chemical,smoke,wind,food,rose,blood,plant,air,person Cluster 2 (Fragrance) perfumery,censer,lavender,jasmine,patchouli,sanctity,odor, incense,perfume,fragrance Cluster 3 (Filth/Pollution) waste,beastly,substance,stink,wretched,dust,flesh,putrid,excrement,filth Cluster 4 mammal,seals,dog,coffee,fish,water,alcoholic,product,drink,animal Cluster 5 (Flowers) petals,drooping,field,faded,orange,passion,plant,blossoms,flowers,flower	5	
Cluster 1 (Occupational) pot,opium,solvent,paraffin,gasoline,gunpowder,petrol, petroleum,substance,chemical Cluster 2 (Alcohol/Drink) fruit,clover,smoke,plant,alcoholic,fragrance,breath,sweet,drink,filth Cluster 3 jerome,healthy,strong,youthful,story,life,man,woman,hair,person Cluster 4 (Animals) nature,wild,things,fat,lioness,cattle,goose,lion,product,animal Cluster 5 (Flowers) fried,white,blue,rose,small,roses,bloom,neglected,flowers,flower	5	1900-1925

Table 17: Gutenberg Causes, by 100 years

Table 18 presents clusters of effects extracted from Gutenberg data collection in different periods of 50 years.

Table 19 presents clusters of effects extracted from Gutenberg data collection in different periods of 100 years.

Table 20 presents clusters of causes extracted from Old Bailey data collection in different periods of 100 years.

Table 21 presents clusters of effects extracted from Old Bailey data collection in different periods of 50 years.

Table 22 presents clusters of effects extracted from Old Bailey data collection in different periods of 100 years.

Table 23 presents clusters of causes extracted from Royal Society data collection in different periods of 100 years.

Table 24 presents clusters of effects extracted from Royal Society data collection in different periods of 50 years.

Table 25 presents clusters of effects extracted from Royal Society data collection in different periods of 100 years.

Figure 9 shows how the points from 5 clusters of causes in Project Gutenberg data collection

Table 18: Gutenberg Effects, k=3, by 50 years

Cluster	Period
Cluster 1 (Body) body,sweat,hear,spirit,spirits,melancholy,nourish,person,excrement,sage Cluster 2 (Religious) relish,censing,delightsome,tastes,better,commends, lightsome,smelling,sweet,lovage Cluster 3 winter,end,odious,world,stinking,smell,woman,offended,filth,nose	1600-1649
Cluster 1 (Flowers) flesh,large,nothing,delicious,may,expedition,meadow,grateful,pleasant,lovage Cluster 2 stable,idoform,person,animal,odor,offensive,fragrant,sage,agreeable,fragrance Cluster 3 (Pollution/Filth) increased,vulgar,observing,bad,horrid,stench, altogether,stink,stinking,filth	1750-1799
Cluster 1 fullery,sage,like,filth,woman,gangreen,sweet,idoform,lovage,person Cluster 2 fragrance,taste,actually,made,feel,air,smell,holding,nose,something Cluster 3 fish,scented,smelt,pillow,know,scent,floral,fervour,candour,fragrance	1900-1925

for olfactory mentions in historical period for years 1800-1899 are grouped on plot.

Figure 9: Gutenberg Causes, k=5, 1800-1899

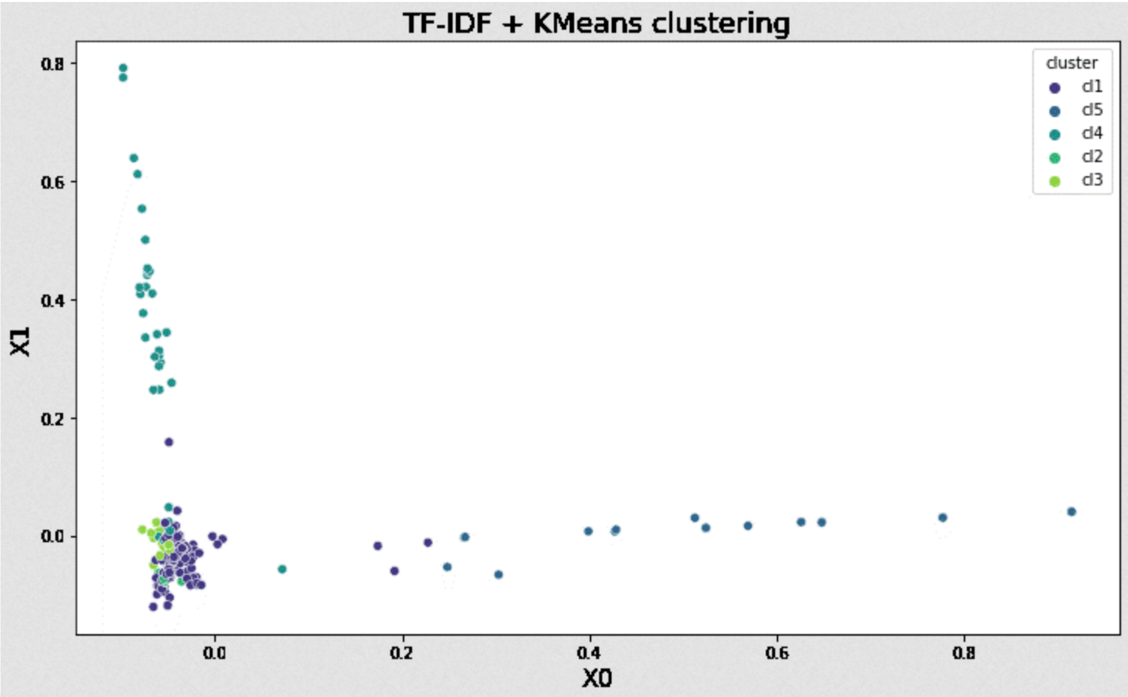


Figure 10 shows how the points from 3 clusters of effects in Project Gutenberg data collection

Table 19: Gutenberg Effects, by 100 years

Cluster	k	Period
Cluster 1 (Materia Medica) make,body,excrement,well,person,savour,stable,sweet,sage,lovage Cluster 2 perfumes,holding,something,offended,filth,aromatise, smell,nose,fragrance,perfume Cluster 3 flora,fleshy,fish,stable,fragrance,bear,scented,scent,dill,ill	3	1600-1699
Cluster 1 root,animal,idoform,filth,agreeable,stable,fragrance,fragrant,sage,lovage Cluster 2 pee,eat,smell,candour,stable,urine,excrement,sanctity,odor,offensive Cluster 3 fragrance,greeted,leaders,bitter,scent,taste,something,holding,nose,person	3	1700-1799
Cluster 1 taste,nose,animal,agreeable,idoform,filth,fragrant,person,sage,lovage Cluster 2 strong,lasting,thereupon,balm,stand,persists,candour,odor,sanctity,stable Cluster 3 holding,something,scented,flavoured,smelling,agreeable, odour,fragrant,perfume,fragrance Cluster 4 (Animals) forced,forget,former,four,fowl,fragrance,earthy,horse,ham,hay Cluster 5 (Animals) forget,former,four,fowl,fragrance,forced,odor,excrement,smell,offensiv e	5	
Cluster 1 flavour,fled,flee,floats,fire,yellow,scene,added,extraordinary,beauty Cluster 2 animal,woman,sage,splendour,idoform,stable,filth,air,person,lovage Cluster 3 smoke,scents,smells,perfumed,fragrant,perfume,fish,candour,smelt,fragrance	3	1800-1899
Cluster 1 fullery,sage,like,filth,woman,gangreen,sweet,idoform,lovage,person Cluster 2 fragrance,taste,actually,made,feel,air,smell,holding,nose,something Cluster 3 fish,scented,smelt,pillow,know,scent,floral,fervour,candour,fragrance	3	1900-1925
Cluster 1 glimmered,girl,garden,gangreen,furze,fur,fullery,gone,youthful,tired Cluster 2 sage,stable,woman,filth,air,gangreen,holding,nose,something,person Cluster 3 like,glorified,kissed,devoutly,drug,richly,cool,glimmered,sweet,lovage Cluster 4 fish,smelt,pillow,know,unexpected,scent,fervour,floral,candour,fragrance Cluster 5 fur,fullery,glorified,dilligence,dismaying,effect, agreeable,intangible,unobtrusive,idoform	5	

Table 20: Old Bailey Causes, by 100 years

Cluster	k	Period
Cluster 1 (Infanticide Trials) filth,nose,woman,excrement,sage,drink,animal,ointment,person,body Cluster 2 frequently,fragrance,one,going,chamber,matter,found,house,flame,fire Cluster 3 young,frequently,fragrance,found,fortnight,fortis,frog,kitchen,itch,insect Cluster 4 (Kitchen) kitchen,burning,burnt,snuffed,press,offering,thrown,door,snuff,candle Cluster 5 (Smoke) fumigation,smoking,equipment,accessory,wig,tobacco, packaging,scent,snuff,box	5	1700-1799
Cluster 1 factory,box,gas,snuff,scent,animal,wind,fragrance,fire,person Cluster 2 (Alcohol) beer,spirits,brandy,spirit,drunk,drinking,liqueur,liquor,alcoholic,drink Cluster 3 (Substance/Material) sulphur,gunpowder,vinegar,acid,copper,tarpit,paraffin, turpentine,substance,chemical	3	1800-1899
Cluster 1 person,slight,veins,effusion,bonnet,woman,straw,shoe,head,blood Cluster 2 (Prison) liquor,washed,dobree,henbane,henry,prison,prisoner,man,mr,person Cluster 3 (Chemistry) sulphur,gunpowder,vinegar,acid,copper,tarpit,paraffin, turpentine,substance,chemical Cluster 4 (Alcohol/Drink) beer,spirits,brandy,spirit,drunk,drinking,liqueur,liquor,alcoholic,drink Cluster 5 (Fire/Smoke) smoke,factory,box,gas,snuff,scent,animal,wind,fragrance,fire	5	
Cluster 1 (Alcohol) door,draymen,dress,drink,drinking,drunk,drunkenness,young,bedroom,room Cluster 2 draymen,door,dog,dirty,digested,carbolic,chloroform, paraffin,substance,chemical Cluster 3 (Alcohol) breath,influence,bodily,fluid,drunkenness,alcohol,drunk, drinking,alcoholic,drink Cluster 4 dress,fire,two,vomited,woman,blood,vomit,person,house,gas Cluster 5 pumpkin,bottle,idoform,forget,could,nose,something,holding,smell,pepper	5	1900-1925

Table 21: Old Bailey Effects, k=3, by 50 years

Cluster	Period
Cluster 1 would,box,kind,agues,filth,person,sage,snuff,itching,humour Cluster 2 going,golden,gone,furze,would,sage,stable,cures,itch,perfectly Cluster 3 going,golden,gone,got,get,thought,fish,smelt,smelling,fragrance	1700-1749
Cluster 1 house,nose,candles,man,could,snuff,idoform,offensive,filth,person Cluster 2 ginger,fish,time,peculiar,candour,fire,strong,smelt,fragrance,smell Cluster 3 fresher,furze,gangreen,ginger,give,good,fragrance,years,neighbours,alarmed	1750-1799
Cluster 1 could,fish,something,holding,nose,drink,smelt,scent,smell,fragrance Cluster 2 go,said,censing,appeared,agitated,got,woman,think,prisoner,person Cluster 3 gas,glad,followed,wrong,firmly,believe,fumigation,candles,snuffing,snuff	1800-1849
Cluster 1 stable,soap,noticed,gangreen,saw,affected,drink,sick,see,person Cluster 2 health,sickly,injurious,annoying,onycha,disagreeable,inconvenience, filth,idoform,offensive Cluster 3 dreadful,tea,could,fish,smelt,holding,nose,something,fragrance,smell	1850-1899
Cluster 1 bodily,stable,forget,blood,faint,idoform,filth,strong,notice,person Cluster 2 forget,four,fragrance,fullery,gait,gangreen,general,flushed,yard,strongly Cluster 3 forget,noticed,tobacco,breath,drink,could,nose,something,holding,smell	1900-1925

Table 22: Old Bailey Effects, by 100 years

Cluster	**k**	**Period**
Cluster 1 (Fragrance) perfumes,perfume,gold,tea,smelled,smell,smelt,scent,fish,fragrance Cluster 2 smelt,saw,offensive,stable,could,prisoner,idoform,drink,snuff,person Cluster 3 could,breath,always,filth,went,forget,nose,holding,something,smell	3	1800-1899
Cluster 1 stable,offensive,idoform,could,holding,something,nose,drink,smell,person Cluster 2 (Commerce) perfumes,perfume,gold,tea,smelled,smell,smelt,scent,fish,fragrance Cluster 3 gangreen,fumigation,fullery,fresh,fragrance,found,fortis,foot,yards,noticed Cluster 4 (Coining Offences) found,fortis,get,yards,firmly,believe,fumigation,candles,snuffing,snuff Cluster 5 (Theft) passed,met,prison,words,angry,ill,pledged,pawned,person,prisoner	5	
Cluster 1 bodily,stable,forget,blood,faint,idoform,filth,strong,notice,person Cluster 2 forget,four,fragrance,fullery,gait,gangreen,general,flushed,yard,strongly Cluster 3 forget,noticed,tobacco,breath,drink,could,nose,something,holding,smell	3	1900-1925

Table 23: Royal Society Causes, by 100 years

Cluster	k	Period
Cluster 1 water,fumigation,root,alcoholic,filth,fruit,animal,drink,excrement,fig Cluster 2 (Flowers) juice,trees,palm,tuberose,tree,plants,leaves,leaf,vegetable,plant Cluster 3 (Substance) sulphureous,matter,effluviums,filth,salt,air,sulphur, pollution,chemical,substance	3	1600-1699
Cluster 1 amphibia,chlorine,boric,sulphuric,resin,sulphur,ammonia, acid,substance,chemical Cluster 2 (Pollution/Filth) impregnation,fumes,white,gas,odorous,particles,putrefaction, pollution,filth,excrement Cluster 3 fragrance,water,sulphur,gas,oil,resin,acid,sulphuretted,animal,hydrogen	3	1800-1899
Cluster 1 (Material/Substance) gum,hot,melted,portion,froths,melts,brittle,substance,chemical,resin Cluster 2 (Substance) ethylic,amphibia,chlorine,boric,sulphuric,sulphur,ammonia, acid,substance,chemical Cluster 3 (Fragrance) ether,ethereal,ethacetone,foug,lost,fumigation,fumigations, bergamot,smell,fragrance Cluster 4 (Public Health) fluid,ozone,water,sulphur,gas,oil,acid,sulphuretted,animal,hydrogen Cluster 5 (Pollution/Filth) impregnation,fumes,white,gas,odorous,particles,putrefaction, pollution,filth,excrement	5	
Cluster 1 (Chemical Process) deposit,active,heating,cooling,durian,excrement,radium,water,thorium,emanation Cluster 2 (Substance) hydrogen,hydrocarbon,chloramino,mercury,acid,durian,formaldehyde, radium,substance,chemical Cluster 3 emanations,silver,cold,ammoniacal,ammonia,reduced,strongly, fehling,reduces,solution Cluster 4 (Radiation/Substance/Material) ema,elements,element,durian,dried,dissolved,disintegration,diacetyl,excess,radiation Cluster 5 (Radiation) gaseous,products,air,lightning,helium,summer,thorium,oblite,person,radioactive	5	1900-1925

Table 24: Royal Society Effects, k=3, by 50 years

Cluster	Period
Cluster 1 (Materia Medica) gangreen,sage,smell,animal,censing,colour,strong,stable,excrement,fragrance Cluster 2 (Pollution/Filth) abominably,sulphureous,sulphur,fetid,render,putrefaction, stinking,excrement,stink,filth Cluster 3 fetid,filth,ac,eh,acgf,de,element,ki,fruit,fig	1650-1699
Cluster 1 (Chemistry) filth,lovage,animal,fragrance,strong,smell,air,offensive,perceived,stable Cluster 2 (Fire) flame,operation,burned,discontinued,sulphur,blackened, fire,burning,offering,burnt Cluster 3 female,felt,fell,feet,zv,ch,bd,ktk,fruit,fig	1750-1799
Cluster 1 colourless,transparent,tar,smelling,change,grayish,dark,coloured,liquid,appearance Cluster 2 (Pollution/Filth) odour,chemical,substance,excrement,pollution,stable,air,idoform,smell,filth Cluster 3 (Chemistry) exist,extent,odour,becomes,quantity,prussiate,peculiar,strongly,censing,perceptible	1800-1849
Cluster 1 (Pollution/Filth) polished,previously,left,excrement,sickening,obnoxious,black,impurity,stain,filth Cluster 2 sharp,substance,odour,burning,nauseous,vomit,holding,something,nose,taste Cluster 3 less,stable,dilligence,liquid,impregnation,pleasant,colourless,acid,idoform,sage	1850-1899
Cluster 1 oblite,noticeable,silicified,summer,radiation,diminution,substance,sage,air,stable Cluster 2 radium,known,leak,since,cooling,unessential,actual, volume,thorium,emanation Cluster 3 fig,whilst,evaporation,even,excess,exposed,extremely, feels,equilibrium,helium	1900-1925

Table 25: Royal Society Effects, by 100 years

Cluster	k	Period
Cluster 1 gangreen,sage,smell,animal,censing,colour,strong, stable,excrement,fragrance Cluster 2 (Pollution/Filth) abominably,sulphureous,sulphur,fetid,render,putrefaction, stinking,excrement,stink,filth Cluster 3 fetid,filth,ac,eh,acgf,de,element,ki,fruit,fig	3	1600-1699
Cluster 1 filth,lovage,animal,fragrance,strong,smell,air, offensive,perceived,stable Cluster 2 (Fire) flame,operation,burned,discontinued,sulphur,blackened, fire,burning,offering,burnt Cluster 3 female,felt,fell,feet,zv,ch,bd,ktk,fruit,fig	3	1700-1799
Cluster 1 sage,chemical,liquid,colourless,acid,smell,excrement, idoform,filth,perceptible Cluster 2 sweetish,substance,odour,unmistakable,sulphuretted,smell, taste,something,holding,nose Cluster 3 shape,fervour,retained,unmixed,earths,fetor,effervescence,smoke,air,stable	3	1800-1899
Cluster 1 oblite,noticeable,silicified,summer,radiation, diminution,substance,sage,air,stable Cluster 2 radium,known,leak,since,cooling,unessential,actual,volume,thorium,emanation Cluster 3 fig,whilst,evaporation,even,excess,exposed,extremely,feels,equilibrium,helium	3	1900-1925

for olfactory mentions in historical period for years 1900-1925 are grouped on plot.

Figure 10: Gutenberg Effects, k=3, 1900-1925

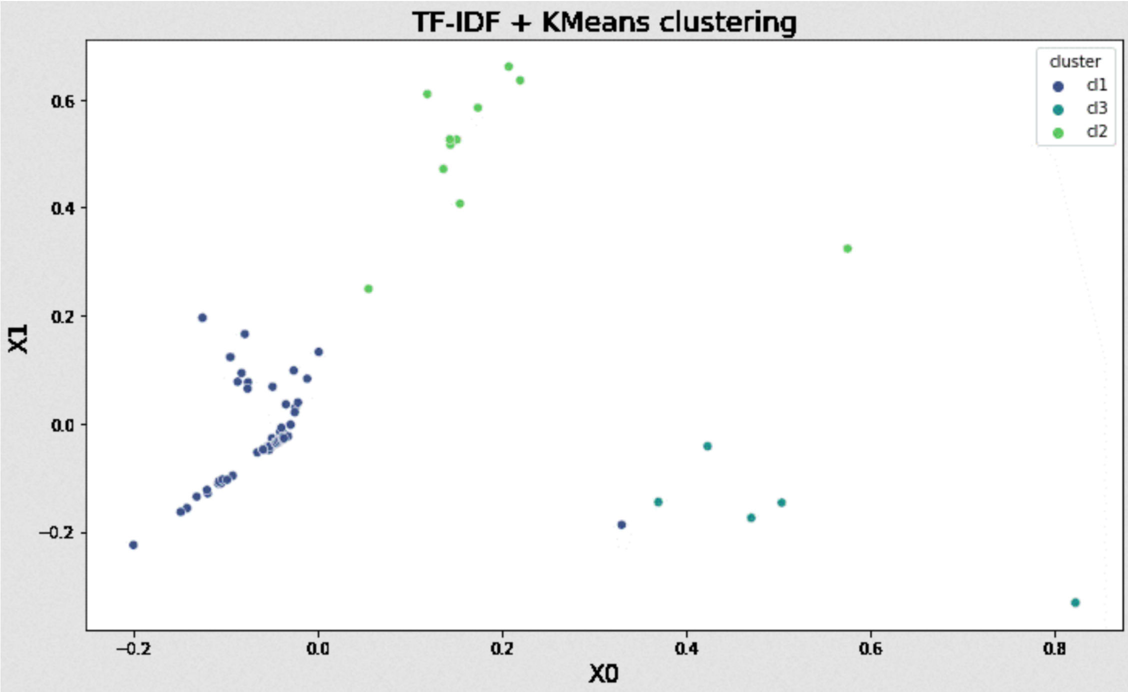


Figure 11 shows how the points from 3 clusters of causes in Old Bailey data collection for olfactory mentions in historical period for years 1800-1899 are grouped on plot.

Figure 11: Old Bailey Causes, k=3, 1800-1899

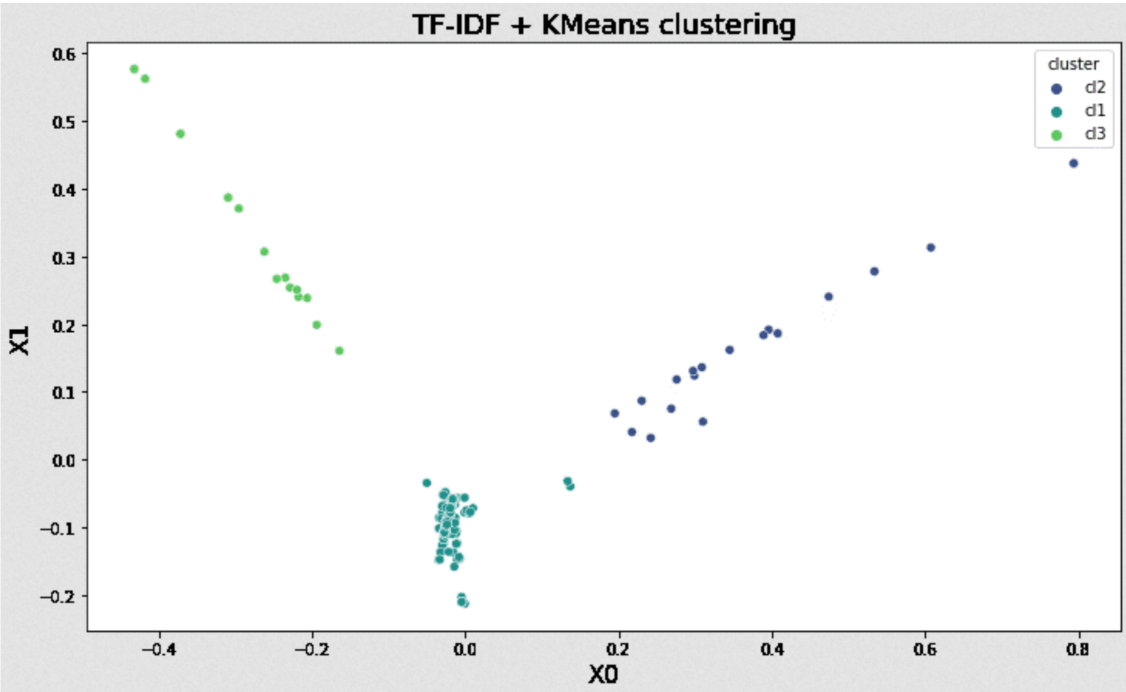


Figure 12 shows how the points from 3 clusters of effects in Old Bailey data collection for olfactory mentions in historical period for years 1800-1899 are grouped on plot.

Figure 12: Old Bailey Effects, k=3, 1800-1899

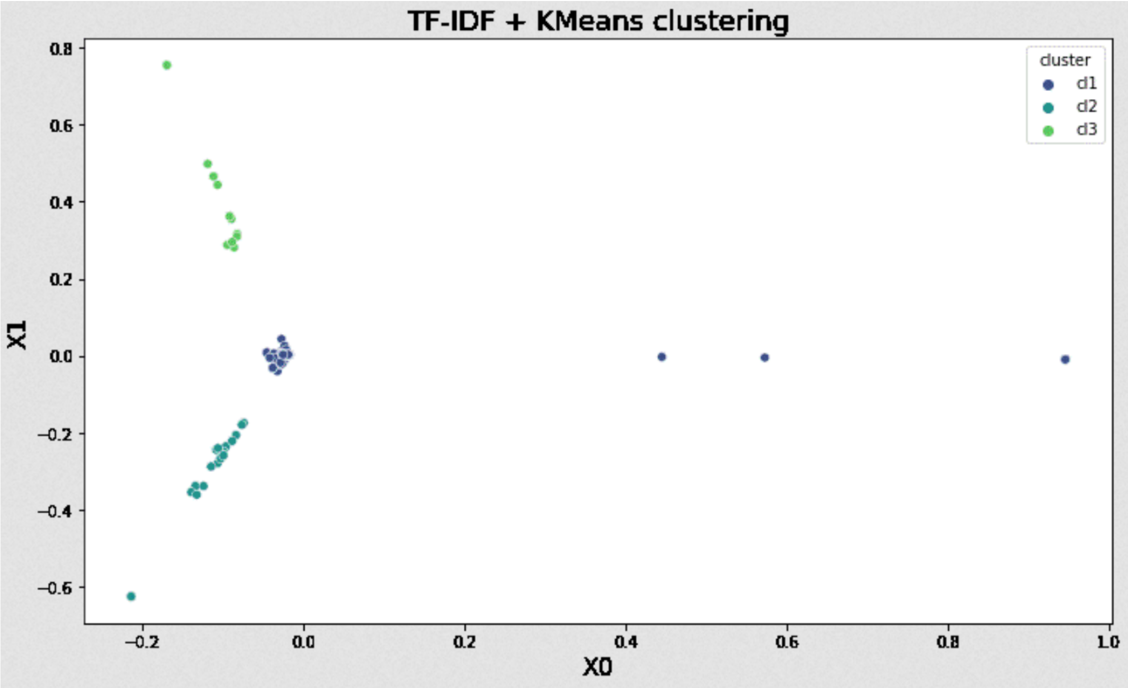
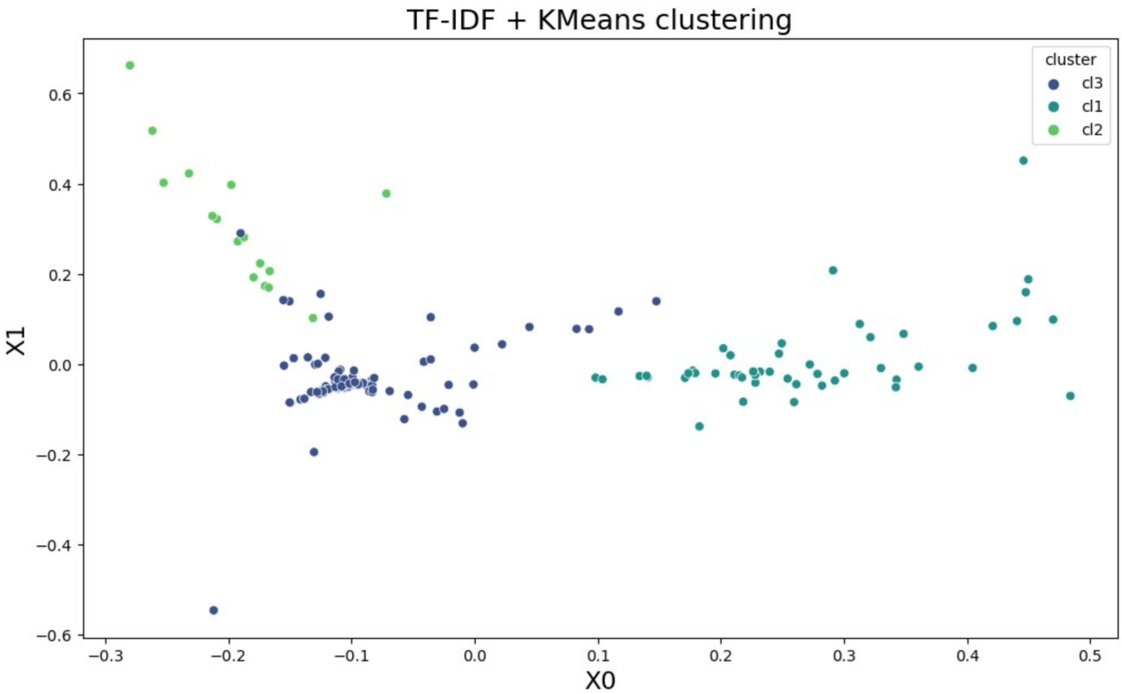


Figure 13 shows how the points from 3 clusters of causes in Royal Society data collection for olfactory mentions in historical period for years 1800-1899 are grouped on plot.

Figure 13: Royal Society Causes, k=3, 1800-1899

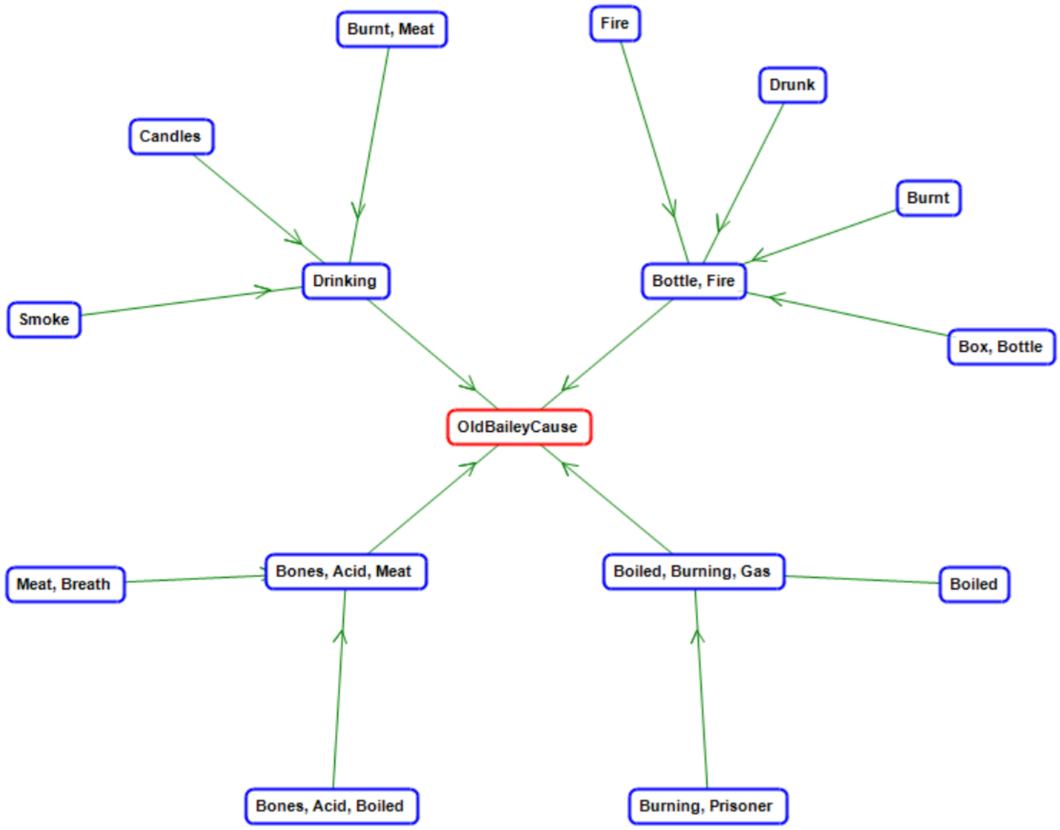


Annex II

Annex II presents an additional cause/effects analysis performed with Ontogen tool.

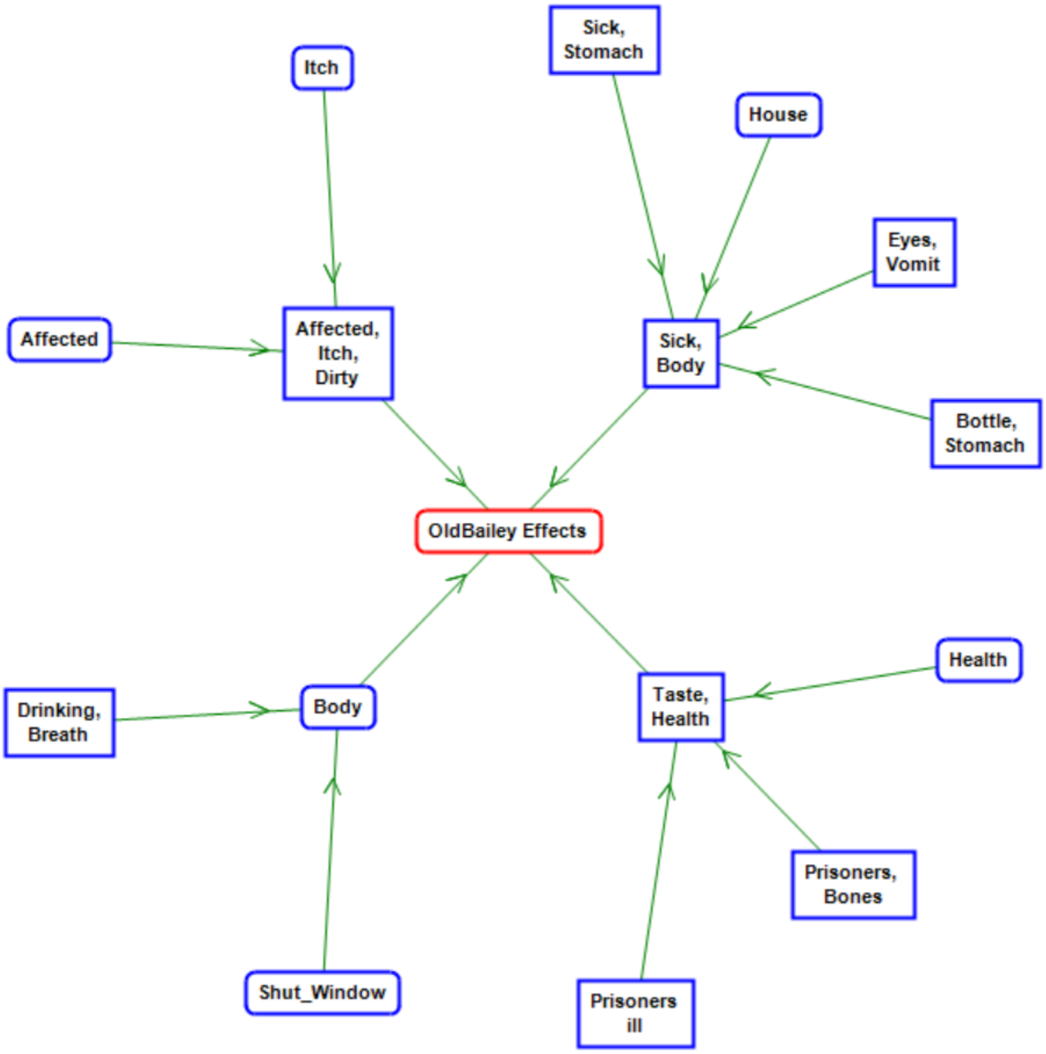
Analysis of causes extracted from Old Bailey data collection with Ontogen tool is provided on Figure 14.

Figure 14: Analysis of Old Bailey Causes with Ontogen Tool



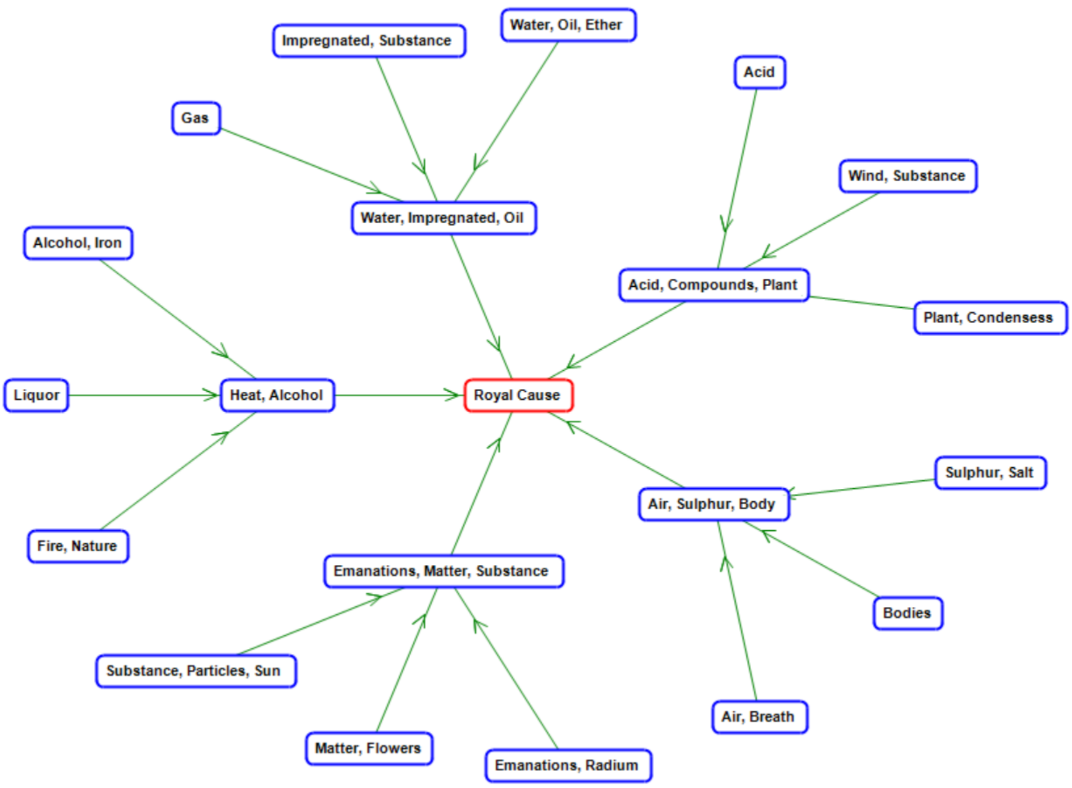
Analysis of effects extracted from Old Bailey data collection with Ontogen tool is provided on Figure 14.

Figure 15: Analysis of Old Bailey Effects with Ontogen Tool



Analysis of causes extracted from Royal Society Bailey data collection with Ontogen tool is provided on Figure 16.

Figure 16: Analysis of Royal Society Causes with Ontogen Tool



Analysis of effects extracted from Royal Society Bailey data collection with Ontogen tool is provided on Figure 17.

Figure 17: Analysis of Royal Society Effects with Ontogen Tool