Capturing the Semantics of Smell: The Odeuropa Data Model for Olfactory Heritage Information

Pasquale Lisena[1(✉)], Daniel Schwabe[2], Marieke van Erp[3], Raphaël Troncy[1], William Tullett[4], Inger Leemans[3], Lizzie Marx[5], and Sofia Colette Ehrich[3]

[1] EURECOM, Sophia Antipolis, France
{pasquale.lisena,raphael.troncy}@eurecom.fr
[2] Jožef Stefan Institute, Ljubljana, Slovenia
daniel.schwabe@ijs.si
[3] KNAW Humanities Cluster, Amsterdam, The Netherlands
marieke.van.erp@dh.huc.knaw.nl, {inger.leemans,sofia.ehrich}@huc.knaw.nl
[4] Anglia Ruskin University, Cambridge, UK
william.tullett@aru.ac.uk
[5] University of Cambridge, Cambridge, UK
esm38@cam.ac.uk

Abstract. Smells are a key sensory experience. They are part of a multi-billion euro industry and gaining traction in different research fields such as museology, art, history, and digital humanities. Until now, a semantic model for describing smells and their associated experiences was lacking. In this paper, we present the Odeuropa data model for olfactory heritage information. The model has been developed in collaboration with olfactory and art historians. Our model can express the various stages in a smell's lifetime – creation, being experienced, deodorisation – and their relation to locations, times and the agents that interact with them.

Keywords: Smell · Ontology · Cultural heritage · Vocabularies · Sensory mining

1 Introduction

Smells are a key sensory experience. As olfactory information goes straight from the nose, from the olfactory bulb, to the limbic system, the amygdala and hyppocampus, smells often evoke strong emotions and memories [32]. Throughout history, these emotive and mnemonic qualities of smelling have been recognised and described, for example in John Louis-Francois Ramond's *Travels in the Pyrenees* (French 1789; English translation 1813):

> There is a somewhat in **perfumes** which powerfully awakens the *memory* of the past. Nothing so soon recalls to the mind *a beloved spot*, *a regretted*

P. Groth et al. (Eds.): ESWC 2022, LNCS 13261, pp. 387–405, 2022.
https://doi.org/10.1007/978-3-031-06981-9_23

> *situation*, or *moments* whose passage has been deeply recorded in the heart, though lightly in the memory. The **fragrance** of a *violet* restores us to the *enjoyment* of many *springs*.

Senses such as vision and hearing are largely studied in signal processing and computer science, while others are underrepresented in scientific research. The sense of olfaction can be found in this latter group. However, the domain of smell, which is often perceived as a fringe one, is in fact quite broad and relevant by humanities and social science scholars [24,25,59]. A new interest in the odours of the past and how past odours are perceived in the present has stimulated research in cultural heritage, or more specific: into olfactory heritage [5] – where scents and smellscapes are understood as a form of both *Tangible* and *Intangible Heritage*[1] – and where scholars have become interested in past ways of smelling and historical smell scapes [28,47]. In addition to the perfume-making industry, we acknowledge the interest of GLAMs [23,62], urban design [22], tourism and environment preservation,[2] human-computer interfaces and 'computer nose' devices [8]. Emerging research has also been triggered by olfactory dysfunctions due to the COVID-19 pandemic, which make the preservation of past olfactory experience more urgent [40].

To capture information about historical smells, olfactory practices, and smell scapes, we can reach out to the rich digital heritage collections that have been developed over the last decades. References to smells and olfactory practices can be found in a large variety of digital texts and images: normative texts, medical texts and perfume handbooks, for instance, offer information about the production and usage of perfumes, or smell management [30,33]. Novels, poems and travel literature reveal connections between odours and identities, testifying to cultural sensitivities around smelling. They may also describe fragrant places, such as churches, parks, or sewers [29,59]. Olfactory clues, gestures and allegories can also be found in paintings, prints and other visual sources [56].

Cultural heritage data collections pose both an opportunity and a challenge. Up to now, most effort in olfactory mining has been put into mapping and classifying fragrances and malodours (specifically in the perfume and odour industries) and in computing the nose - the act of smelling and its effect on the body. Smells are notoriously hard to predict. Thusfar, olfactory informatics has been focused on computing what a molecule smells like based on its chemical structure [36,49,63]. Heritage texts and images however, provide a different type of information. They offer rich data about odour perception and valuation, about the cultural experience of smelling, including subjective interpretations of the perceived odours. To capture this information, different computer science technologies are required, such as image recognition, text mining, and semantic web technologies. Odeuropa[3] is the first major research project to combine these technologies to capture smell experiences in their historical and cultural con-

[1] It is relevant the inclusion of the perfumes of Grasse in the UNESCO list. Source: https://bit.ly/3opPRin. Last visited: 15/03/2022.

[2] Examples in Japan: https://bit.ly/3u4ySFD and in France: https://bit.ly/3rYpv7Q.

[3] https://www.odeuropa.eu.

text. Our goal is not so much to represent smells *per se* but rather to represent the historical and social aspects of smell perception and olfactory practices.

In this paper, we introduce the Odeuropa data model for representing odours and their experiences from a cultural heritage perspective. The data model reuses and extends established ontologies such as CIDOC CRM [14], to represent the relevant information as a set of interconnected events. The model is completed by a set of controlled vocabularies for representing crucial elements such as olfactory objects and gestures.

The ontology is developed using web technologies and is intended as a structure for realising an olfactory Knowledge Graph (KG) in the context of the Odeuropa project. The KG will include olfactory-related data extracted from text and images from the 17th to the early 20th century. This KG is intended to serve as a base for supporting heritage professionals, historians and scent designers in including Artificial Intelligence (AI) techniques in their daily practice [39].

This ontology is contributing to the domain in two ways:

- Offering a structure for representing smell-related information, a necessary step for the preservation of this intangible heritage. The ontology can potentially serve in all the previously mentioned areas of the olfactory domain;
- For the first time, closing a gap in the representation between objective observations and subjective experiences (in particular, sensory ones). This aspect is targeted also going beyond the olfactory domain, using a 2-level structure – with a first layer targeting senses in general and a second one focusing specifically on smells – enabling the description of sensorial experiences in fields such as history, literature, art, and cultural heritage.

The remainder of this paper is organised as follows. In Sect. 2, we describe related work on olfactory and semantic web modelling. In Sect. 3, we present our approach to designing the Odeuropa olfactory model with domain experts and the model requirements. In Sect. 4, we present our model, detailing the modelling decisions underneath, while in Sect. 5 are described the vocabularies that followed from the knowledge elicitation from domain experts. Section 6 reports about the evaluation of the data model. We showcase our model's use and expressivity in Sect. 7 with an example describing a smell experience. We conclude with a discussion and plans for future work in Sect. 8.

2 Related Work

The olfactory domain has typically been the purview of perfumers [27], psychologists [16], and sanitary scientists [61]. In the past decade, museologists, chemists, and historians have become interested in researching and preserving heritage smells [6] and curating smell archives, such as the Osmothèque in Versailles.[4] While there is a subfield of computer science that concerns itself with

[4] https://www.osmotheque.fr/en/the-collection/.

olfactory informatics, these studies are mostly focused on predicting or mapping olfactory characteristics of molecules [36,49,63], and semantic modelling of smells is as yet an under-researched area.

Previous work has shown that Knowledge Graphs are suitable to represent and exploit the domain information in cultural heritage [11,14,26], history [31], and art [1,13], as well as complex intangible domains such as event modelling [53, 60], biomedicine [54], ecological networks [58], and chemistry [41]. Whilst smells have a measurable component, namely their molecular composition, they remain a largely intangible and subjective concept as most smell discourse is based on personal observations. An olfactory model therefore needs to be able to deal with subjective observations that are anchored to a place and time.

In the past decade, (digital) humanities researchers have started working with semantic web researchers to develop ontologies and knowledge graphs for their domain such as [1,7,50,51]. For visual information, various ontologies and knowledge graphs are available such as [55]. Auditory information is well covered by for example [42], and a taste model was developed by [44]. The IoT community has been working on a digital senses model [12] and the concept of an artificial nose [19]. Recently, [52] combined odours, odorants, olfactory receptors and odorant-receptor interactions in a single MySQL database. However, this model focuses on the chemical compositions of odours, is not open data, and is less focused on the sensory impacts of olfactory experiences and heritage than Odeuropa. To the best of our knowledge, there is a lack of ontologies that specifically aim to represent sensory experiences.

3 Design Methodology and Model Requirements

As there is a large gap between the everyday practices of computer science and humanities researchers in which this model is to fit, we opted for a user-centred design methodology in which the olfactory and (art) historian experts were closely involved. In a series of meetings and hands-on exercises, the requirements of the model were elicited whereby the overall Odeuropa project goals were kept in mind as end-goal.[5] A core instrument in this process was the formulation of 74 competency questions for the model to answer.

With each step, the intermediate results were shared and progress on the design was measured according to the competency questions formulated and results of these were used to steer the next development iteration. A visual overview of our method is provided in Fig. 1.

The desired ontology will serve for storing together olfactory information from structured resources, as well as information from texts and images. Furthermore, the resulting knowledge graph is to be used to research smells through time and related to places. Historians are furthermore particularly interested in what people did with smells or how they created smells and what emotions these evoked. These requirements led to the following 7 categories of competency questions, all declinable in time and space:

[5] https://odeuropa.eu/objectives-timeline/.

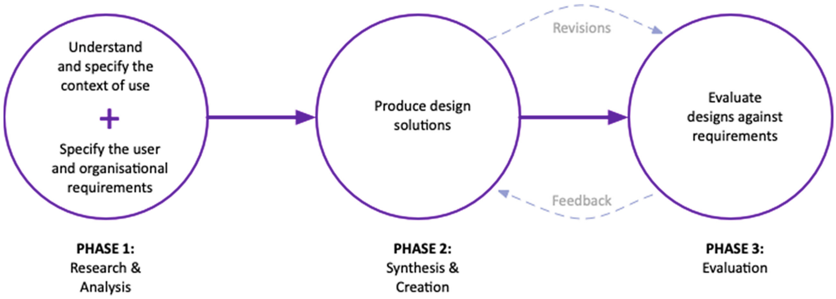

Fig. 1. User-centred design

Smells: About core properties of smell (source, carrier).
- *What are the most frequent smell sources in London in the 18th century?*
- *Which smells were perceived during spring?*

Noses and Gestures: Involving the actors perceiving the odours.
- *Which professions are more present in smelling experience descriptions?*
- *Which smelling gestures are described more frequently by tea-merchants?*

Identities: About the meaning of smells and their capability of being representative of something/someone.
- *Which flavours did people associate with femininity in Asia?*
- *What are the odours most associated with Ashkenazi Jewish practices?*

Emotions: Focusing on the interaction between olfaction and feelings.
- *What odours disgusted upper-class Europeans most?*
- *Which smell triggers memories of childhood?*

Practices: About smell-producing practices.
- *What types of cooking produce a bad smell?*
- *Which practice can reduce a smell intensity?*

Sites and contexts: About the presence of odours in particular places.
- *Which smells are associated with ships?*
- *Which smell could be perceived during the Crimean War?*

Texts and images: About how smells are represented in texts and images.
- *What are the adjectives used for orange aroma in the 15th century?*
- *Which smells can be found in paintings of the Rijksmuseum?*

4 The Odeuropa Data Model

This section describes the olfactory data model, highlighting our core modelling decisions and the main structure of the resulting ontology.

4.1 Extending Established Ontologies

Following best practices in the ontology development [10], we aim to re-use existing data models as base and extend them to represent domain-specific classes and properties. Given the lack of sensory-centered ontology (Sect. 2), we chose **CIDOC CRM** [14] as our core ontology for the following reasons:

- **It is a bridge to other cultural and heritage objects:** CIDOC CRM can be used to describe objects in museum and creative works [38], including paintings and textual resources. This makes it more natural to describe the relations between olfactory information and those elements;
- **It is already familiar to museums and digital libraries:** This can be an advantage when creating interlinking with existing collections and for eventual adoption by these institutions;
- **It is event-based:** [46] Due to the intangible nature of smells and the inevitable subjectivity in their usual descriptions, we decided to focus on the representation of olfactory events rather on odours themselves. In CIDOC CRM, events are the fundamental building blocks: the existence of anything is implying an event that generated it or made someone aware of it. An event can be described in relation to time, space, and involved participants can be linked to other events, including sub-events such as actions and gestures;
- **It is expressive and flexible:** The information to be represented may vary significantly, ranging from highly detailed olfactory experiences to brief mentions to a particular smell. The modularity of CIDOC CRM – itself made of events as building blocks which may be freely interconnected – provides the required flexibility in the representation. In particular, it allows to independently represent the event which generated (or transformed) the smell and the olfactory experience(s), giving the possibility of describing both or only one of them.

CIDOC CRM is extended by **CRMsci** [15], which adds properties about the scientific observation and description of natural phenomena. Here, the *observation* concept has to be understood in the broad sense of *experiencing something*, such that it can also be applied to sensory experiences beyond sight.

As a derivation of CIDOC CRM, the Odeuropa model follows the naming convention to prefix classes and property names with a number and a letter: CIDOC CRM uses E (for classes) and P (for properties); CRMSci uses S (classes) and O (properties); Odeuropa uses L (classes) and F (properties), taking two letters from "o**lf**action". In the text of this paper, we will omit these codes and letters for readability, while keeping them in the figures.

In addition, parts of the following ontologies are used:

- The READ-IT ontology, to represent emotions triggered by events [3];
- The PROV-O Ontology, for representing data provenance [34];
- The FOAF vocabulary, for describing people [18];
- Schema.org [20], e.g. to describe the genre and author of a text or painting.

4.2 A Three-Layered Model

Due to the complexity of the phenomena related to odours, we adopted a layered approach to construct the data model. We identified abstraction levels that

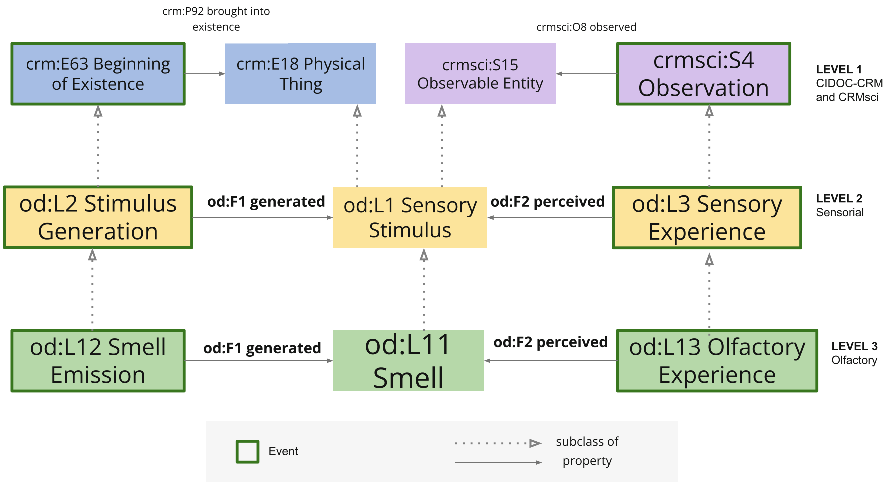

Fig. 2. The core of the Odeuropa data model

roughly correspond to the different aspects of interest. Accordingly, the Odeuropa Data model is organised in the following three levels:

- **Level 1** consists of the CIDOC CRM and CRMsci classes and properties that were used and/or extended. It represents an observation of a phenomenon;
- **Level 2** is an extension of Level 1 for representing sensorial experiences, not limited to olfaction. This level was developed because we identified commonalities shared by all senses and decided to provide more general classes and properties. This will help future extension of the model, including the representation of synaesthetic experiences;
- **Level 3** extends Level 2 by specifically targeting olfactory information.

The three levels are shown in Fig. 2, representing the core of the model. *Smell* (Sensory Stimulus) plays a central role, directly connected to two main types of events, namely *Smell Emission* (Stimulus Generation) and *Olfactory Experience* (Sensory Experience).

In this model, we consider a smell as a unique and non-repeatable entity, with defined time and space coordinates. By way of example, two roses have two distinct (but similar) smells, and the "smell of roses" exists only as a generalisation of the smells of all roses. A given smell can be generated by a unique *Smell Emission event*, but can be experienced multiple times, in distinct situations, by multiple people. This captures the fact that each person can perceive and describe the same smell differently [2].

Figure 3 reports all elements that are part of the data model. The information is organised around the three main events, directly linked to the Smell:

- The **Smell Emission** allows us to describe the smell generation from a smell source (e.g. tobacco) and the carrier of the smell (e.g. a pipe). These elements

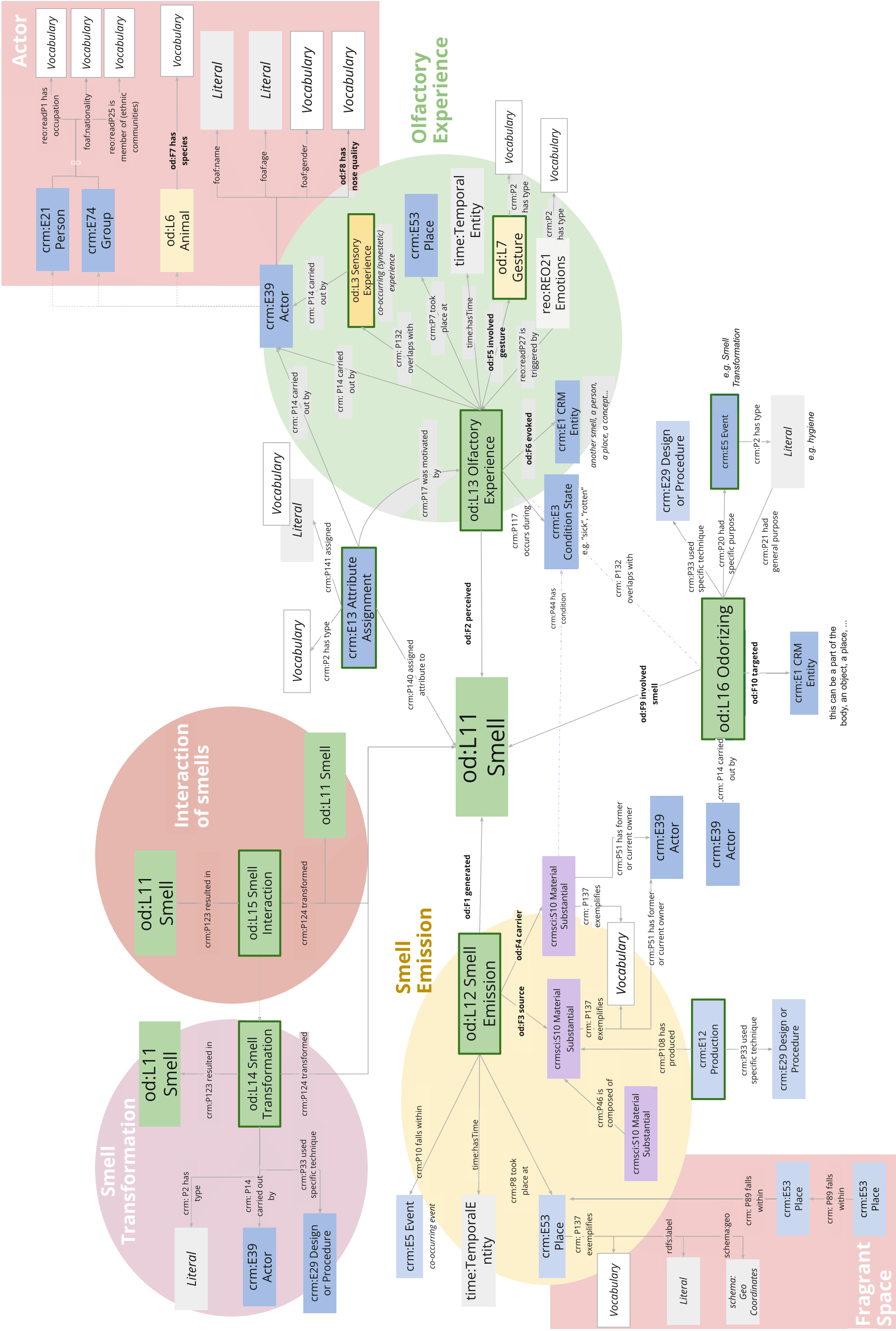

Fig. 3. The Odeuropa data model

can be further described through their components and/or the production process which creates them;
- The **Olfactory Experience** allows us to describe the perception of smell – who perceived the smell, their eventual emotions and gestures. In addition, it records the description that the perceiver makes of the smell, be it through adjectives (typed and linked to vocabularies using the Attribute Assignment class) or through the mention of (i.e., association with) evoked entities such as other smells, people, places, etc.;
- The **Odorizing** class allows to describe how a specific smell was used. For instance, it is possible to specify the purpose for which an odour was used – e.g. covering another smell, medical reason, etc. –, who was using it on what the smell is being used – e.g. a room, a part of the body, etc.

Both event classes inherit from CIDOC CRM some common properties to specify the time and space of the event and eventual co-occurring events. Given the subjective nature of the words used for describing smell, we preferred to model them as *Attribute Assignment* connecting the word (*assigned*) to the smell (*assigned attribute to*), with a direct link to the original person (*carried out by*) and the possibility to include the attribute (*has type*) in a category (e.g. hedonic, intensity, character, state, etc.).

Furthermore, the model includes also classes such as *Stimuli/Smell Transformation* – to represent events that modify a smell, e.g. opening a window – and *Stimuli/Smell Interaction* – to represent smells that are perceived as a combination of different smells, e.g. different foods in a dining room. Special care was devoted to model perceivers (i.e. the agents perceiving smells), by employing and extending the class *Actor* to represent people, groups and animals. Similarly, fragrant spaces are also represented, capturing those attributes that allow us to aggregate them – by type of place or by geographical contiguity.

4.3 Provenance Information

As we intend to trace smells through time, we need to keep track of the sources from which statements in our knowledge graph are derived and through what process. Furthermore, to anchor statements in time and place, we want to keep track of when they were published, and if possible who published them to – for example – map a debate on cultural differences with respect to a particular odour. To keep track of this in the KG, we apply the following strategy:

CIDOC CRM enables us to represent that a text (*Linguistic Object*) or image (*Visual Item*) contains a reference (*refers to*) to an entity, which is, in our domain, a Smell or an Olfactory Event. To include the information without drastically increasing the number of triples in the KG, these *refers to* links are instantiated on a subset of the graph, containing at least the core.

PROV-O [34] is used to record the ways this information was extracted from textual and visual sources, including the agent and/or software/algorithm which extracted the information and a confidence score in case automatic processes

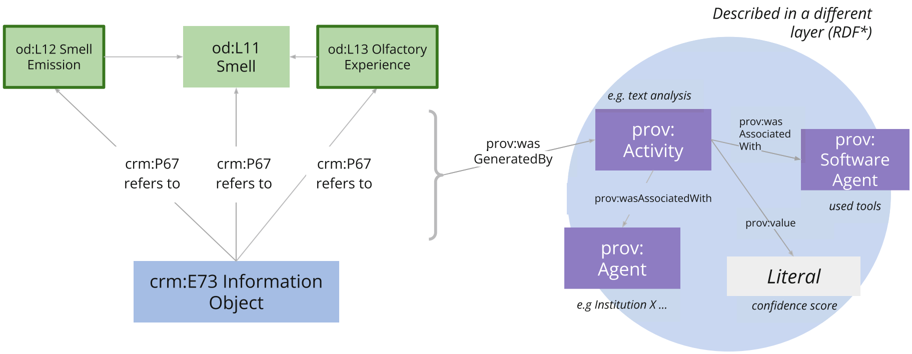

Fig. 4. The provenance of information represented in the Odeuropa data model

were involved. To keep the graph clean, we include this information in a second layer. This is realised by applying RDF* [21] and linking the provenance information to the relevant *refers to* properties, as shown in Fig. 4. In this way, the information and the meta-information are kept distinct, while it is always possible (when needed) to retrieve the provenance of a data excerpt.

5 Controlled Vocabularies

For the description of some fundamental olfactory-related concepts, a collection of controlled vocabularies was created. The use of vocabularies helps to better disambiguate entities, grouping synonyms and labels in different languages under a single identifier (URI). Our vocabularies are represented in SKOS [43], a format that allows us to define, for each concept, preferred and alternate labels, descriptions, broader, narrower and related terms. In this way, we can construct a hierarchy of terms, grouping the related ones and to instantiate bridges between concepts belonging to different vocabularies.

Following previous experiences in constructing controlled vocabularies in Digital Humanities [35,37], our collection is composed of previously-existing taxonomies – which are expressed using SKOS – and vocabularies built from scratch through the collaboration of domain experts and computer scientists.

The list of olfactory vocabularies converted in SKOS format is reported in Table 1 and consists of:

- The Fragrance Circle by Edward Drom, a smell wheel used in perfumery; [9]
- Michael Edwards' Fragrance Wheel, including 4 families and 14 subfamilies of olfactory groups used in modern perfumery;[6]
- The Odour wheel of historical books, for smell heritage preservation [5];
- The Nose-first classification of iconographies realised by Ehrich et al. for linking smells and their representation in art [17];

[6] https://en.wikipedia.org/wiki/Fragrance_wheel Last visited: 07/12/2021.

Table 1. Vocabularies converted in SKOS. Some classification systems have a second level which consists of smell sources rather than smell classes (reported as 1+1).

Name	Type	Levels	Top level concepts	Total concepts
Drom's fragrance circle	Odour wheel	2	16	77
Michael Edwards' scent wheel	Odour wheel	2	4	18
Odour wheel of historical books	Odour wheel	2	8	43
Nose-first classification of iconographies	Classification	1+1	25	168
Flavornet and human odour space	Classification	1+1	25	495
Zwaardemaker smell system	Classification	1+1	9	9

- The Flavornet odour space, the compilation of aroma compounds found in human odour space [4];
- The Linnaeus/Zwaardemaker smell system developed in 1895 [48].

These vocabularies were manually converted to a common format based on CSV and then processed and converted to SKOS. In addition, 3 multi-language vocabularies were developed in a collaboration between knowledge engineers and domain experts, representing:

- **Fragrant spaces**, listing interesting (from an olfactory point of view) such as churches, buildings, natural places, etc. These concepts are intended to be linked to instances of type `E53 Place` through *P137 exemplifies*;
- **Olfactory gestures**, simple actions which possibly occur during olfactory experiences, e.g. sniffing, covering the nose, etc. The included concepts are intended to be linked to instances of type `L13 Olfactory Experience` through *F5 involved gesture*;
- **Olfactory objects**, including entities (natural or human made) which are particularly relevant because emitting odours – e.g. a flower – or potentially carrying odour sources – e.g. a perfume bottle or a pomander. The included concepts are intended to be linked to instances of type `L12 Smell Emission` through *F3 source* or *F4 carrier*.[7]

The realisation of these vocabularies was carried out with synchronised spreadsheet tabs – one for each language – to collect the translations of each term. In addition, semantic relationships between terms inside the same vocabulary were instantiated – e.g. "Rose" *skos:broader* "Flower" or "Pipe" *skos:related* "Tobacco – and between vocabularies – e.g. "Library" *skos:related* "Book". An overview of the available languages is shown in Table 2. Please note that a given concept does not always have an appropriate translation in all languages.

[7] While some of these are clearly carriers (*wind*, *bottle*) and other smell sources (*jasmine*, *sulphur*), some specific elements can embody any of the two role depending on the context (*smoke*). For this reason, we decided to have a single vocabulary including all terms, reporting the preferred role when possible.

Table 2. Multi-language vocabularies for English (EN), German (DE), French (FR), Italian (IT), Dutch (NL), and Slovene (SL)

Name	Total concepts	EN	DE	FR	IT	NL	SL
Fragrant spaces	110	110	4	108	106	110	4
Olfactory gestures	35	35	0	33	32	16	0
Olfactory objects	417	400	172	378	381	390	402

6 Evaluation with Competency Questions

To guide the design of the data model and to provide a way to evaluate it, we used the set of 74 Competency Questions (CQ) [45] collected in Sect. 3 before the development of the model. These CQ were proposed by domain experts – historians and scholars with expertise in olfactory heritage – and are organised in 7 categories, reported in Table 3. These questions allowed the team to iteratively improve versions of the data model, in sequences of development and check. We considered the process of designing the ontology complete only when each CQ could be expressed with a proper SPARQL query, making sure that all the components and relations necessary to answer this question are in place.

In the final version of the model, we distinguish 4 different cases:

- The vast majority of questions can be answered with a SPARQL query. Example: *What are the most frequent smell sources in London in the 18th century?*
- A few questions cannot be answered by simple SPARQL queries, but require more AI methods to find a proper solution.
  Example: *Was muck perceived as more disgusting than smog?*
- 4 questions are answerable with SPARQL, but require external information that are outside the scope of the model – e.g. with the addition of knowledge bases such as WikiData.
  Example: *Which smell could be perceived during a war?*
- 1 question requires an extension of the model. Given that the challenging element of this question is not directly related to the olfactory/heritage information but to time representation, we decide to keep this issue open for future work.
  Example: *Which smells were perceived during morning?*

Apart from the last group, we consider the other cases satisfied by the model. Our results are summarised in Table 3.

Some of the CQs require additional AI techniques to be solved in a more exhaustive way. For example, when searching for bad smells, we are not only interested in the result for a query exactly matching the word *bad*, but we are interested in all kinds of *malodours*, smells described as *stinking*, *terrible*, *awful*, etc. We identified 2 possible strategies to address this situation:

Table 3. The number of competency question per category, together with the number of answerable one with the sole model (OK), in combination with AI techniques (AI), with the addition of external data (ExtData) and only with a further extension of the model (Extension)

Category	OK	AI	ExtData	Extension	Total
A. Smells	10	0	0	1	11
B. Noses and gestures	6	0	0	0	6
C. Identities	6	0	0	0	6
D. Emotions	6	0	0	0	6
E. Practices	8	5	0	0	13
F. Sites and contexts	9	0	2	0	11
G. Texts and images	19	0	2	0	21
TOTAL	64	5	4	1	74

- Sentiment detection on the words used for describing the smell (when we search for *good*/*bad* or *pleasant*/*unpleasant* smells);
- Rely on word embeddings and compute the similarity between the word in the graph (e.g. *reluctant*, *fetid*) and the searched one (e.g. *disgusting*).

7 Showcase: Modelling the Smell of a Location

To better understand and appreciate the expressivity and flexibility of the proposed data model, we showcase a modeling example. In Fig. 5, we model the olfactory information contained in a passage from Vita Sackville-West's *Knole and the Sackvilles* (1922). In this book, the author describes the house she grew up in but could not inherit due to aristocratic inheritance customs:

> "They [**galleries** of **Knole**, ed.] have the **old**, **musty** smell which to **me**, whenever I met it, would bring back Knole. I suppose it is really the smell of all **old houses** - a mixture of **woodwork**, **pot-pourri**, **leather**, **tapestry**, and the little **camphor bags** which keep away the moth, and specifically about the pot pourri: **bowls of lavender** and **dried rose-leaves** stand on the window-sills; and if you **stir them** up you get the quintessence of the smell, a sort of **dusty** fragrance, **sweeter** in the under layers where it has held the damp of the spices."

The different olfactory sources mentioned are not physically combined together as in a recipe,[8] but they separately emit different smells which are combined (*Smell Interaction*) in the galleries of Knole. The author perceives this ensemble smell and describes this ensemble smell as *old* and *musty*. In the

[8] In that case, there will not be a Smell Interaction, but a single Smell Emission having as source the union of the different ingredients.

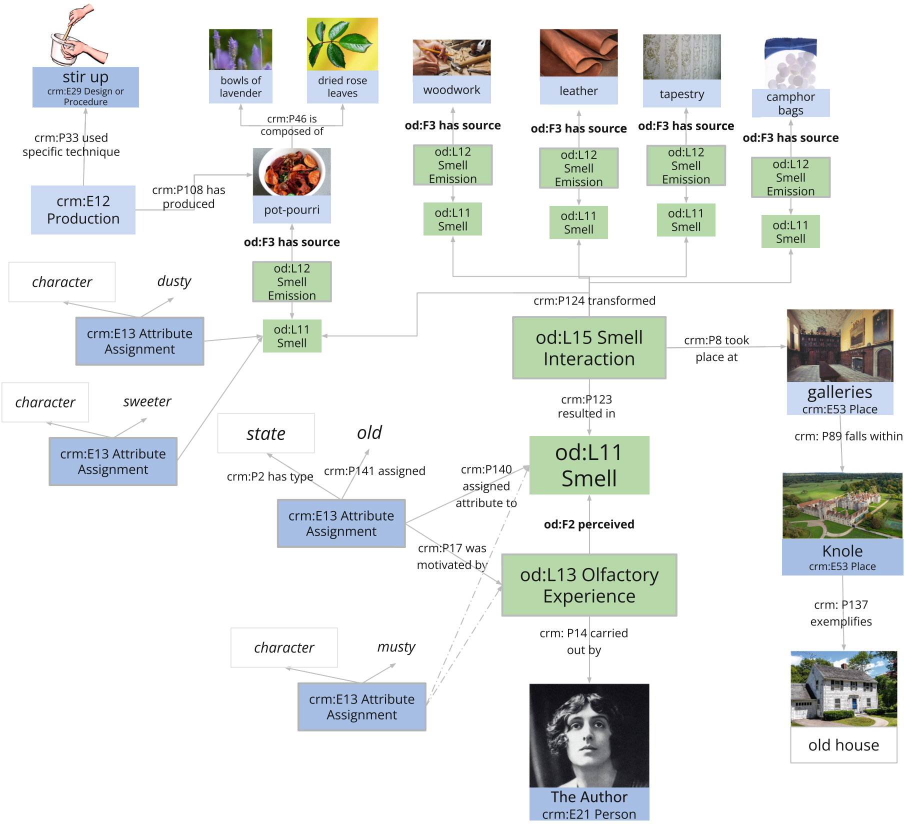

Fig. 5. An example of data representation in the Odeuropa model from a passage of Vita Sackille-West's *Knole and the Sackvilles* (1922)

text, one of the member smells, emitted by the pot-pourri is described, as *dusty* and *sweeter*, also mentioning the procedure of its realisation from lavender and rose leaves. Graph nodes can be interlinked with the controlled vocabularies of olfactory objects – e.g. leather, camphor bags – and fragrant spaces – e.g. old house. Further examples can be found in the Data Model presentation (see Table 5).

8 Conclusions and Future Work

In this paper, we introduced the Odeuropa Data Model, an olfactory extension of CIDOC CRM and CRMsci. The model can represent smell-related information, in particular describing the emission, use and experience of a given odour. The data model is accompanied by a set of multi-language controlled vocabularies for disambiguating of crucial olfactory information elements, such as the odour source or associated gestures. The model is implemented in OWL format and

Table 4. Re-used classes and properties in the Odeuropa data model

Ontologies	Reused classes	Reused properties
CIDOC CRM	10	16
CRMsci	1	0
FOAF	0	4
PROV-O	3	3
READ-IT	1	3
Schema.org	1	4
Time	1	1

Table 5. Resource table

Resource	URL
Data model complete presentation	https://bit.ly/3GuIHzL
Ontology (OWL)	https://github.com/Odeuropa/ontology
Ontology (Documentation)	http://data.odeuropa.eu/ontology/
Competency questions	https://bit.ly/odeuropa-cq
Vocabularies (RDF)	https://github.com/Odeuropa/vocabularies
Vocabularies (SKOSmos)	http://vocab.odeuropa.eu/
Vocabulary API	http://data.odeuropa.eu/api/vocabulary Doc: https://github.com/D2KLab/vocabulary-api
Odeuropa KG	http://data.odeuropa.eu/

published at http://data.odeuropa.eu/ontology under a Creative Commons 4.0 CC-BY License, along with its documentation. Odeuropa proposes 13 new classes and 10 new properties to capture olfactory information, defined as subclassed and subproperties of CIDOC CRM and CRMsci. To these, classes and properties from other models have been reused, as reported in Table 4.

Table 5 lists the pointers to all resources that we developed and published in the context of this work, available as resources to the whole community. In addition to the ontology and the competency questions, some olfactory controlled vocabularies are available via different access points in RDF (Turtle format) using SKOS, in a wide-public visualisation based on SKOSmos [57], through a HTTP API which can be used for interlinking.

The ontology and the vocabularies are part of the Odeuropa Knowledge Graph, hosted at http://data.odeuropa.eu/. At the time of writing, we are populating this graph with data extracted from text and images. This will constitute a multifaceted playground for the data model and for olfactory heritage research. Use of the knowledge graph may also inspire further extensions, validations and improvements of the Odeuropa Data Model.

In future work, we intend to further extend the data model. In particular, we aim to close the gap between the smell heritage domain and the perfume industry, for example by including the representation of chemical compounds and olfactory notes. In addition, we want to better investigate the capability of the model to represent synaesthetic experiences, i.e. the connections people perceive between different sensory experiences such as seeing colours when smelling fragrances. We also intend to extend the vocabularies by including new terms and translations and by adding new thesauri and classifications to our list.

Acknowledgements. This work has been partially supported by European Union's Horizon 2020 research and innovation programme within the Odeuropa project (grant agreement No. 101004469). Smells that helped get this paper out: citrus (to boost our energy levels), rosemary (to keep us alert) and the smell of hell (to keep us on our toes).

References

1. Achichi, M., Lisena, P., Todorov, K., Troncy, R., Delahousse, J.: DOREMUS: a graph of linked musical works. In: Vrandečić, D., et al. (eds.) The Semantic Web – ISWC 2018. LNCS, vol. 11137, pp. 3–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00668-6_1
2. Almagor, U.: Odors and private language: observations on the phenomenology of scent. Hum. Stud. **13**(3), 253–274 (1990). https://doi.org/10.1007/BF00142757
3. Antonini, A., et al.: Understanding the phenomenology of reading through modelling. Semant. Web **12**, 191–217 (2021). https://doi.org/10.3233/SW-200396
4. Arn, H., Acree, T.: Flavornet: a database of aroma compounds based on odor potency in natural products. Dev. Food Sci. **40**, 27–28 (1998)
5. Bembibre, C., Strlič, M.: Smell of heritage: a framework for the identification, analysis and archival of historic odours. Herit. Sci. **5**(1), 2 (2017). https://doi.org/10.1186/s40494-016-0114-1
6. Bembibre Jacobo, C.: Smell of Heritage. Ph.D. Thesis, UCL (University College London) (2020)
7. de Boer, V., van Doornik, J., Buitinck, L., Marx, M., Veken, T., Ribbens, K.: Linking the kingdom: enriched access to a historiographical text. In: Proceedings of the Seventh International Conference on Knowledge Capture, pp. 17–24 (2013)
8. Brooks, J., et al. (eds.): STT21: Smell, Taste, and Temperature Interfaces workshop. Yokohama, Japan (2021). https://stt21.plopes.org/
9. Brud, W.: Words versus odours, how perfumers communicate. Perfum. Flavorist **11**, 27–44 (1986)
10. Carriero, V.A., et al.: The landscape of ontology reuse approaches. In: Applications and Practices in Ontology Design, Extraction, and Reasoning, pp. 21–38. IOS Press (2020)
11. Carriero, V.A., et al.: ArCo: the Italian cultural heritage knowledge graph. In: Ghidini, C., et al. (eds.) The Semantic Web – ISWC 2019. LNCS, vol. 11779, pp. 36–52. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30796-7_3
12. Datta, S.K., Coughlin, T.: An IoT architecture enabling digital senses. In: 2016 IEEE 6th International Conference on Consumer Electronics-Berlin (ICCE-Berlin), pp. 67–68. IEEE (2016)
13. Dijkshoorn, C., et al.: The Rijksmuseum collection as linked data. Semant. Web J. **9**, 221–230 (2018)

14. Doerr, M.: The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. AI Mag. **24**(3), 75 (2003)
15. Doerr, M., Kritsotaki, A., Rousakis, Y., Hiebel, G., Theodoridou, M.: Definition of the CRMsci an extension of CIDOC-CRM to support scientific observation. Forth-Institution of Computer Science (2015)
16. Dravnieks, A.: Atlas of Odor Character Profiles. American Society for Testing and Materials, Philadelphia (1985)
17. Ehrich, S., Verbeek, C., Zinnen, M., Marx, L., Bembibre, C., Leemans, I.: Nose first. Towards an olfactory gaze for digital art history. In: First International Workshop on Multisensory Data and Knowledge (MDK), Zaragoza, Spain (2021)
18. Graves, M., Constabaris, A., Brickley, D.: FOAF: connecting people on the semantic web. Cat. Classif. Q. **43**(3–4), 191–202 (2007). https://doi.org/10.1300/J104v43n03_10
19. Guest, C., et al.: Feasibility of integrating canine olfaction with chemical and microbial profiling of urine to detect lethal prostate cancer. PLoS ONE **16**(2), e0245530 (2021)
20. Guha, R.V., Brickley, D., Macbeth, S.: Schema. org: evolution of structured data on the web. Commun. ACM **59**(2), 44–51 (2016)
21. Hartig, O.: The RDF* and SPARQL* approach to annotate statements in RDF and to reconcile RDF and property graphs. In: W3C Workshop on Web Standardization for Graph Data, Berlin, Germany (2019)
22. Henshaw, V.: Urban Smellscapes: Understanding and Designing City Smell Environments. Routledge/Taylor & Francis Group, New York (2014)
23. Howes, D.: Introduction to sensory museology. Senses Soc. **9**(3), 259–267 (2014). https://doi.org/10.2752/174589314X14023847039917
24. Howes, D.: Empire of The Senses: The Sensual Culture Reader. Routledge, London (2021)
25. Howes, D., Classen, C.: Ways of Sensing: Understanding the Senses in Society. Routledge, London (2013)
26. Isaac, A., Haslhofer, B.: Europeana linked open data - data.europeana.eu. Semant. Web J. **4**, 291–297 (2013)
27. Jaubert, J.N., Tapiero, C., Dore, J.: The field of odors: toward a universal language for odor relationships. Perfum. Flavorist **20**, 1 (1995)
28. Jenner, M.S.: Follow your nose? Smell, smelling, and their histories. Am. Hist. Rev. **116**(2), 335–351 (2011)
29. Kettler, A.: The Smell of Slavery: Olfactory Racism and the Atlantic World. Cambridge University Press, Cambridge (2020)
30. Kiechle, M.A.: Smell Detectives: An Olfactory History of Nineteenth-century Urban America. University of Washington Press, Seattle (2017)
31. Koho, M., Ikkala, E., Leskinen, P., Tamper, M., Tuominen, J., Hyvönen, E.: WarSampo knowledge graph: Finland in the second world war as linked open data. Semant. Web J. **12**, 265–278 (2021)
32. Krusemark, E.A., Novak, L.R., Gitelman, D.R., Li, W.: When the sense of smell meets emotion: anxiety-state-dependent olfactory processing and neural circuitry adaptation. J. Neurosci. **33**(39), 15324–15332 (2013)
33. Le Guérer, A.: Parfum. Le), Des origines à nos jours. Odile Jacob (2005)
34. Lebo, T., et al.: PROV-O: the PROV ontology. Technical report, World Wide Web Consortium (2013)

35. Leon, A., Gaitán, M., Insa, I., Sebastián, J., Alba, E.: SILKNOW. designing a thesaurus about historical silk for small and medium-sized textile museums. In: Ortiz Calderón, P., Pinto Puerto, Verhagen, P., Prieto, A. (eds.) Science and Digital Technology for Cultural Heritage. CRC Press, London (2020). https://doi.org/10.1201/9780429345470-34
36. Licon, C.C., et al.: Chemical features mining provides new descriptive structure-odor relationships. PLoS Comput. Biol. **15**(4), e1006945 (2019)
37. Lisena, P., et al.: Controlled vocabularies for music metadata. In: 19th International Society for Music Information Retrieval Conference (ISMIR), Paris, France (2018). http://ismir2018.ircam.fr/doc/pdfs/68_Paper.pdf
38. Lisena, P., Troncy, R.: Representing complex knowledge for exploration and recommendation: the case of classical music information. In: Cota, G., Daquino, M., Pozzato, G.L. (eds.) Applications and Practices in Ontology Design, Extraction, and Reasoning, Studies on the Semantic Web Series (SSWS), vol. 49, pp. 107–123. IOS Press (2020). https://doi.org/10.3233/SSW200038
39. Lisena, P., van Erp, M., Bembibre, C., Leemans, I.: Data mining and knowledge graphs as a backbone for advanced olfactory experiences. In: Brooks et al. [8] (2021). https://stt21.plopes.org/wp-content/uploads/2021/05/STT2021_Odeuropa.pdf
40. Mathis, S., et al.: Olfaction and anosmia: from ancient times to COVID-19. J. Neurol. Sci. **425**, 117433 (2021). https://doi.org/10.1016/j.jns.2021.117433, https://www.sciencedirect.com/science/article/pii/S0022510X21001271
41. de Matos, P., et al.: ChEBI: a chemistry ontology and database. J. Cheminform. **2**(1), 1 (2010). https://doi.org/10.1186/1758-2946-2-S1-P6
42. Meroño-Peñuela, A., et al.: The midi linked data cloud. In: International Semantic Web Conference, vol. 10588, pp. 156–164. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68204-4_16
43. Miles, A., Pérez-Agüera, J.R.: SKOS: simple knowledge organisation for the web. Cat. Classif. Q. **43**(3–4), 69–83 (2007)
44. Naravane, T., Lange, M.: Ontological framework for representation of tractable flavor: food phenotype, sensation, perception. In: ICBO (2018)
45. Noy, N.F., Hafner, C.D.: The state of the art in ontology design: a survey and comparative review. AI Mag. **18**, 53 (1997). https://doi.org/10.1609/aimag.v18i3.1306, https://ojs.aaai.org/index.php/aimagazine/article/view/1306
46. Pasin, M., Motta, E.: Ontological requirements for annotation and navigation of philosophical resources. Synthese **182**, 235–267 (2009). https://doi.org/10.1007/s11229-009-9660-3
47. Perkins, C., McLean, K.: Smell walking and mapping, chap. 10. Manchester University Press, Manchester (2020). https://doi.org/10.7765/9781526152732.00017, https://www.manchesterhive.com/view/9781526152732/9781526152732.00017.xml
48. Philpott, C.M., Bennett, A., Murty, G.E.: A brief history of olfaction and olfactometry. J. Laryngol. Otol. **122**(7), 657–662 (2008). https://doi.org/10.1017/S0022215107001314
49. Sanchez-Lengeling, B., Wei, J.N., Lee, B.K., Gerkin, R.C., Aspuru-Guzik, A., Wiltschko, A.B.: Machine learning for scent: learning generalizable perceptual representations of small molecules. arXiv preprint arXiv:1910.10685 (2019)
50. Schleider, T., et al.: The SILKNOW knowledge graph. Semant. Web 1–16 (2021)

51. Schouten, S., de Boer, V., Petram, L., van Erp, M.: The wind in our sails: developing a reusable and maintainable Dutch maritime history knowledge graph. In: Proceedings of the 11th on Knowledge Capture Conference, K-CAP 2021, pp. 97–104. Association for Computing Machinery, New York (2021). https://doi.org/10.1145/3460210.3493548
52. Sharma, A., Saha, B.K., Kumar, R., Varadwaj, P.K.: OlfactionBase: a repository to explore odors, odorants, olfactory receptors and odorant-receptor interactions. Nucleic Acids Res. (2021). https://doi.org/10.1093/nar/gkab763
53. Shaw, R., Troncy, R., Hardman, L.: LODE: linking open descriptions of events. In: Gómez-Pérez, A., Yu, Y., Ding, Y. (eds.) The Semantic Web, ASWC 2009. LNCS, vol. 5926, pp. 153–167. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-10871-6_11
54. Smith, B., et al.: The OBO foundry: coordinated evolution of ontologies to support biomedical data integration. Nat. Biotechnol. **25**(11), 1251–1255 (2007)
55. Stamou, G., van Ossenbruggen, J., Pan, J.Z., Schreiber, G., Smith, J.R.: Multimedia annotations on the semantic web. IEEE Multimedia **13**(1), 86–90 (2006)
56. van Suchtelen, A.: Fleeting Scents in Colour. Mauritshuis, Den Haag, the Netherlands (2021)
57. Suominen, O., et al.: Publishing SKOS vocabularies with Skosmos. Manuscript submitted for review (2015)
58. Torta, G., Ardissono, L., La Riccia, L., Savoca, A., Voghera, A.: Representing ecological network specifications with semantic web techniques. In: KEOD-International Conference on Knowledge Engineering and Ontology Development, vol. 2, pp. 86–97. SCITEPRESS-Science and Technology Publications, Lda. (2017)
59. Tullett, W.: Smell in Eighteenth-Century England: A Social Sense. Oxford University Press, Oxford (2019)
60. Van Hage, W.R., Malaisé, V., Segers, R., Hollink, L., Schreiber, G.: Design and use of the Simple Event Model (SEM). J. Web Seman. **9**(2), 128–136 (2011)
61. Van Harreveld, A.P., Heeres, P., Harssema, H.: A review of 20 years of standardization of odor concentration measurement by dynamic olfactometry in Europe. J. Air Waste Manag. Assoc. **49**(6), 705–715 (1999)
62. Verbeek, C., van Campen, C.: Inhaling memories. Senses Soc. **8**(2), 133–148 (2013). https://doi.org/10.2752/174589313X13589681980696
63. Wu, D., Luo, D., Wong, K.Y., Hung, K.: POP-CNN: predicting odor pleasantness with convolutional neural network. IEEE Sens. J. **19**(23), 11337–11345 (2019)