Bridging aroma chemistry and sensory perception: a text analysis approach

Beatriz Quintanilla Casas

Postdoctoral researcher
Chemometrics and Machine Learning research group
Department of Food Science, University of Copenhagen

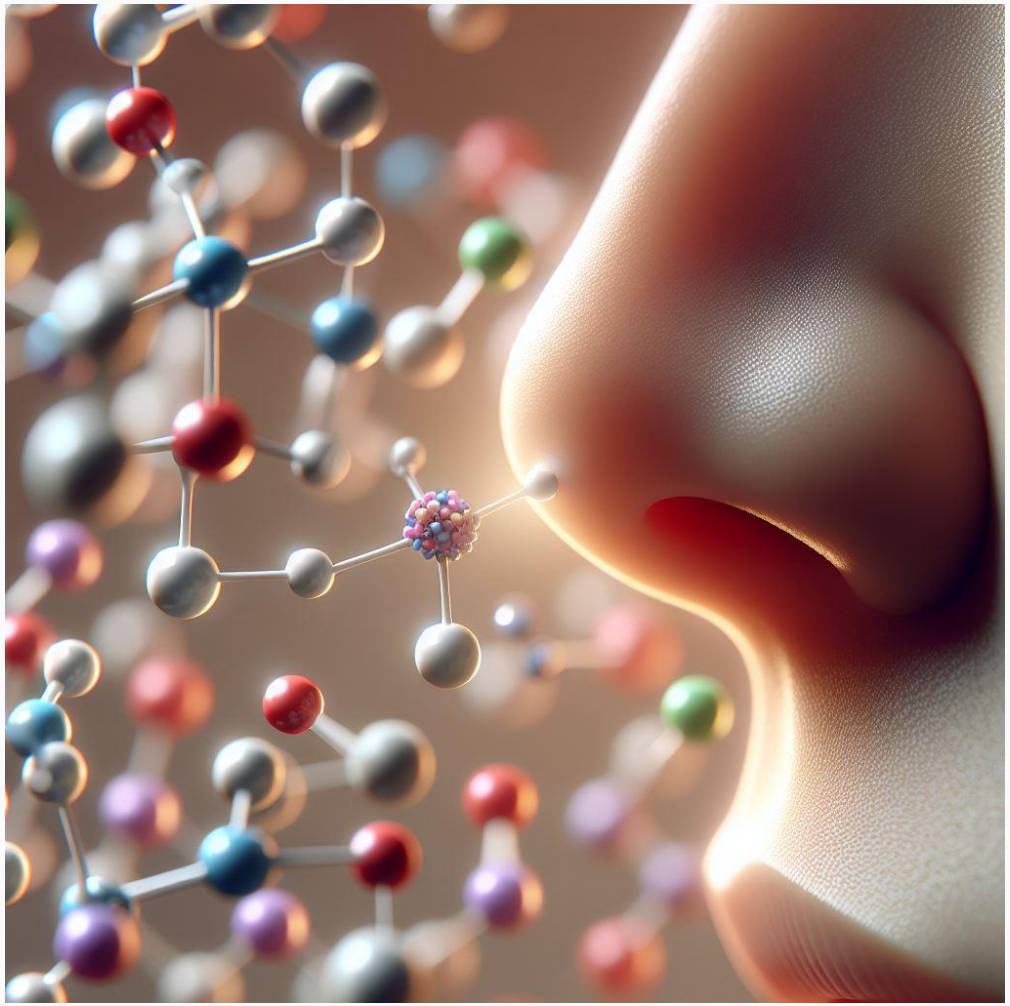
KØBENHAVNS UNIVERSITET

Danish Data Science Academy

Supervisor: Prof. Rasmus Bro

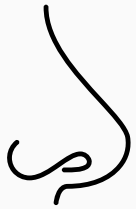
November 8th, 2024
DSK.2024 (Middelfart, Denmark)

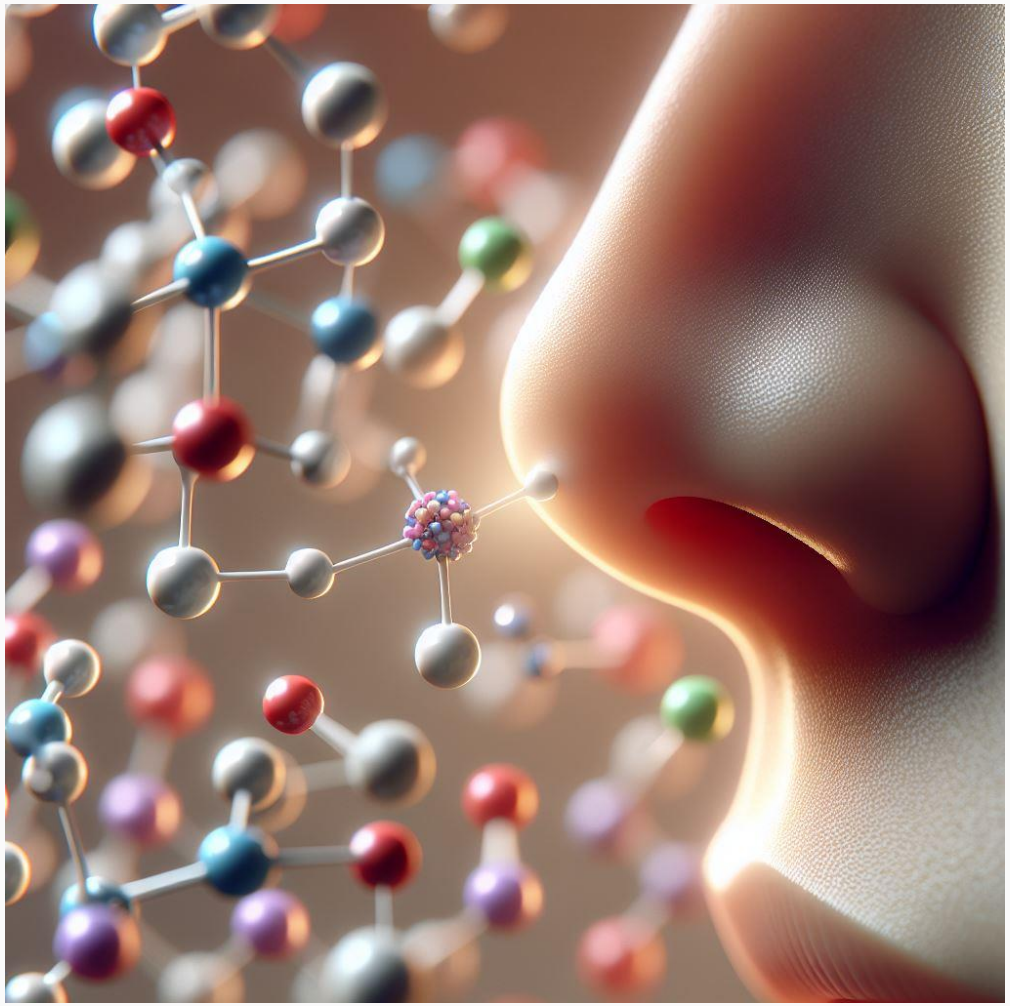
FOOD FLAVOUR



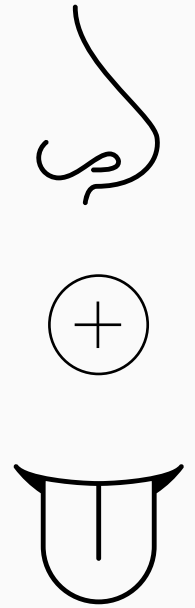
FOOD FLAVOUR

Aroma

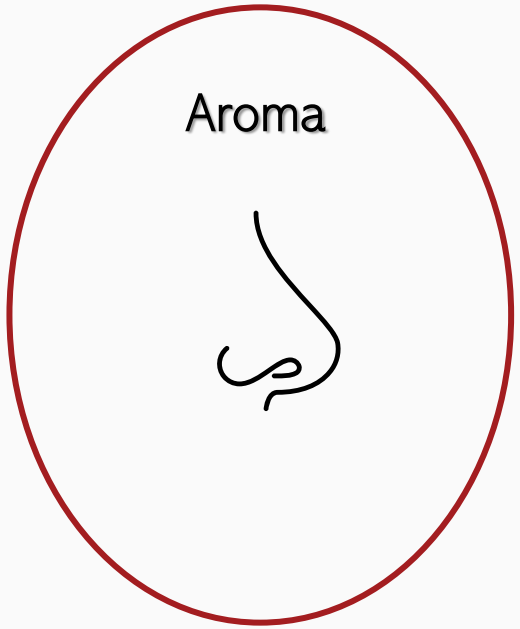




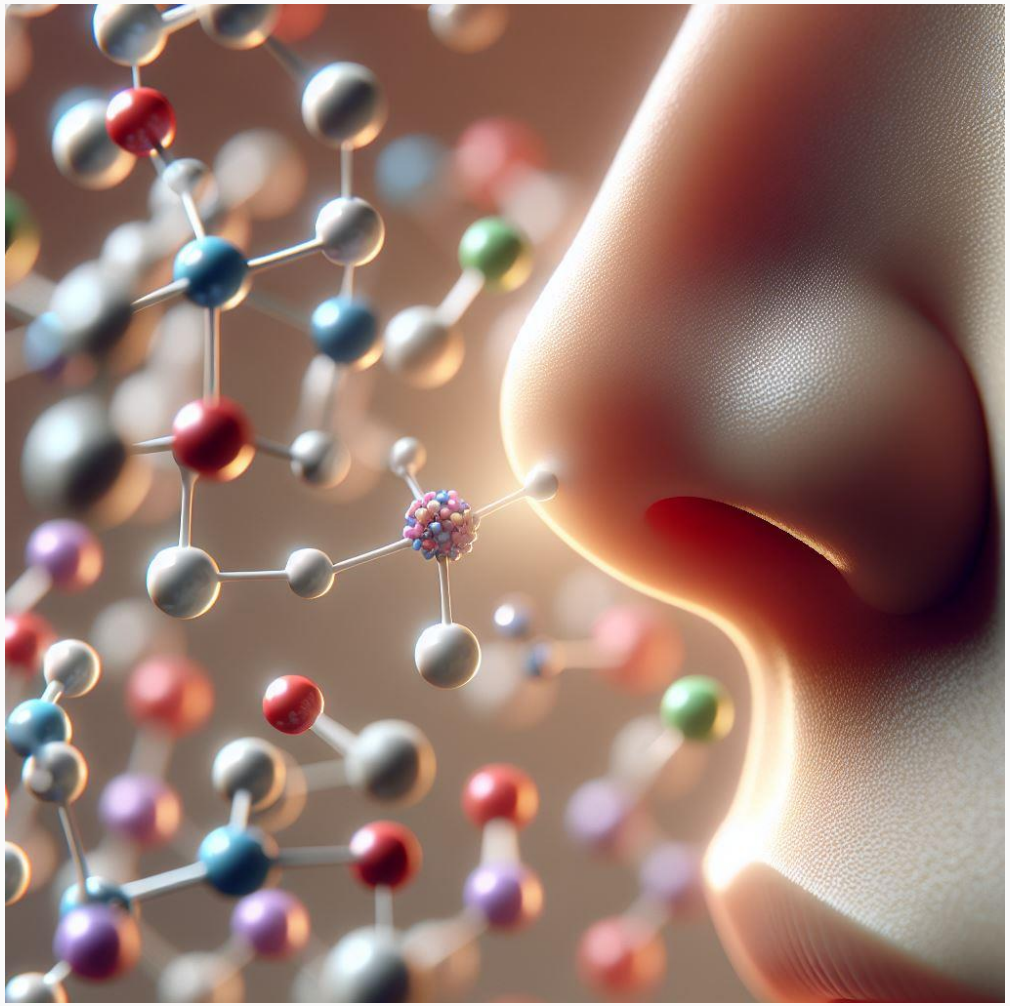
Flavour

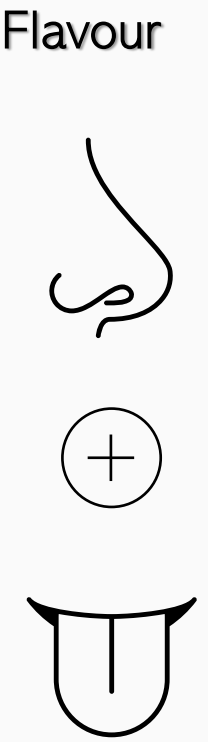


FOOD FLAVOUR

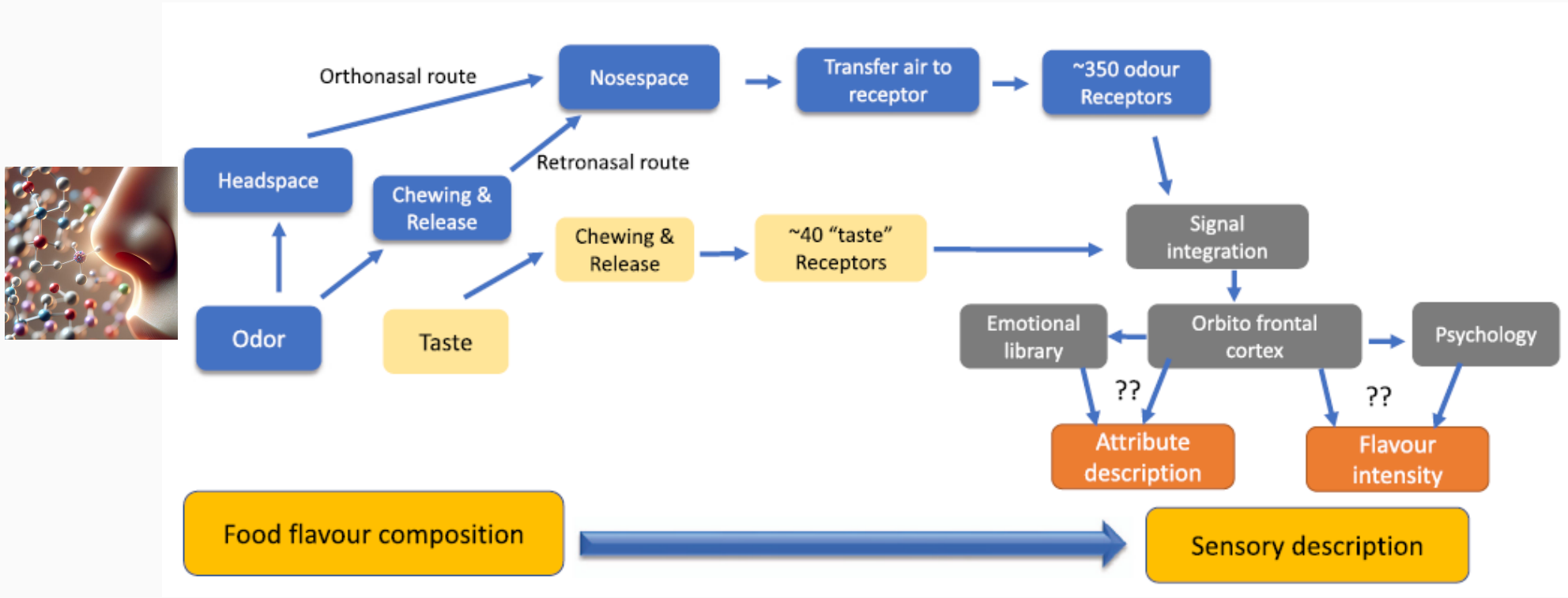


Is it that simple?



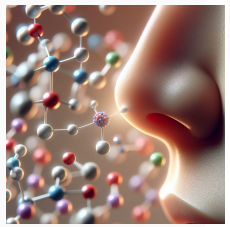


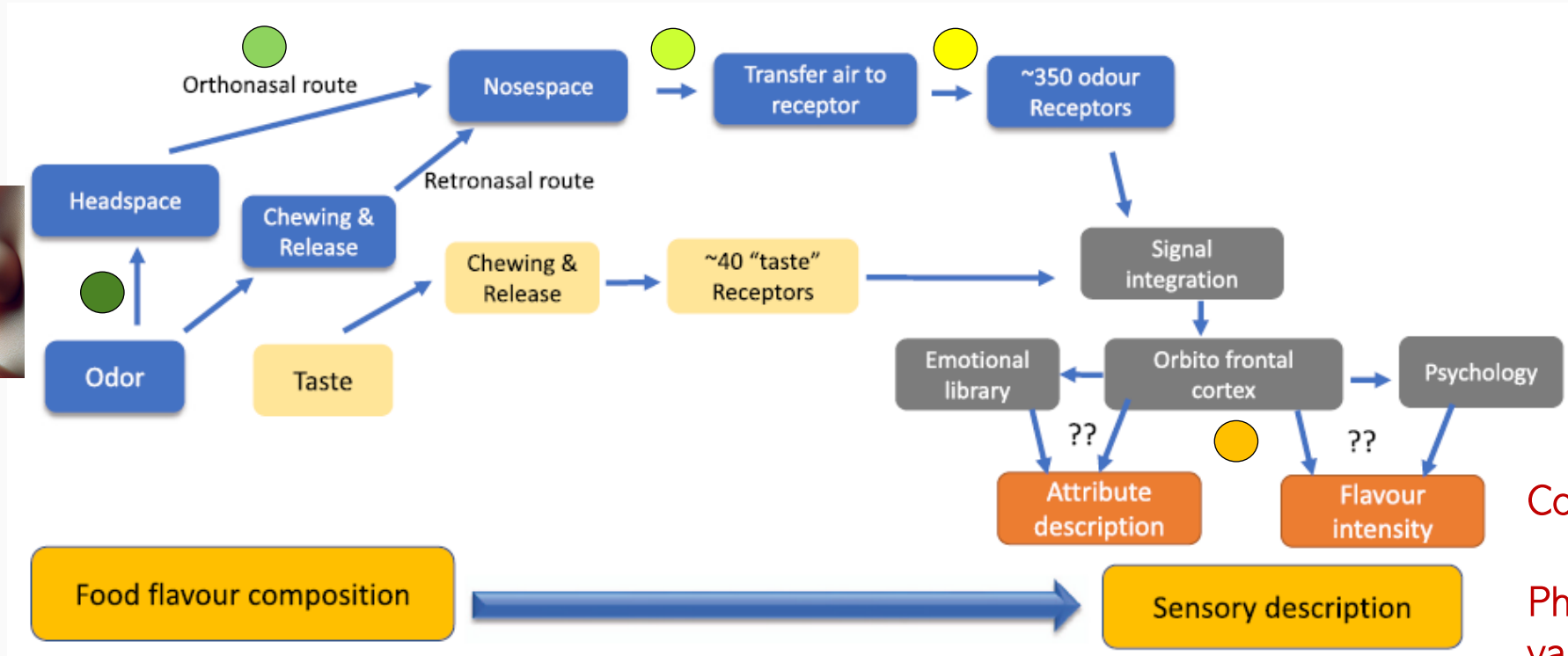
FROM FOOD FLAVOUR TO SENSORY PERCEPTION



Taylor, A.J. (2023) Chapter 9: Predicting sensory properties from chemical profiles, the ultimate flavour puzzle. In: Digital Sensory Science, Woodhead Publishing.

FROM FOOD FLAVOUR TO SENSORY PERCEPTION





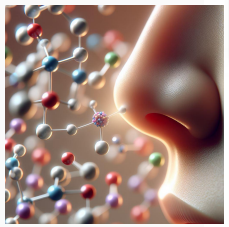
Complexity

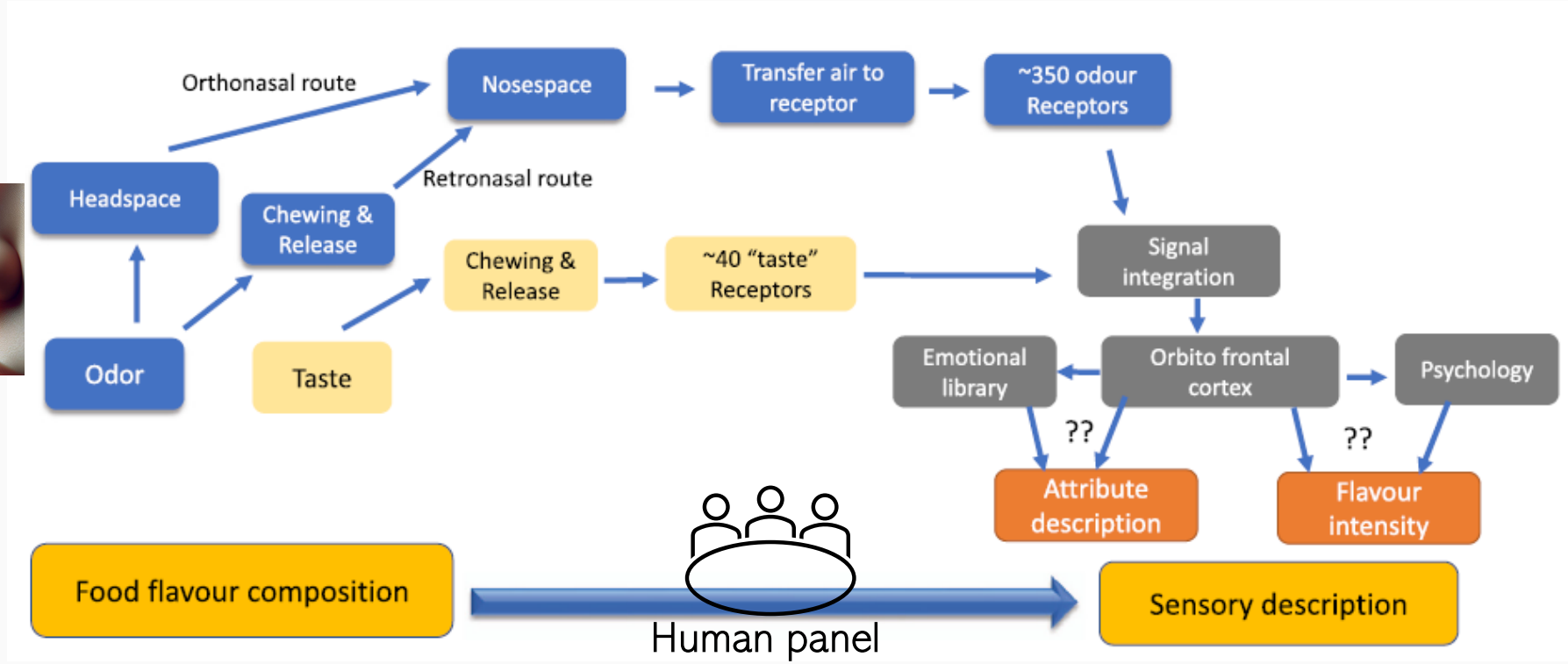
Physiological variability

Differences in the state of knowledge

Taylor, A.J. (2023) Chapter 9: Predicting sensory properties from chemical profiles, the ultimate flavour puzzle. In: Digital Sensory Science, Woodhead Publishing.

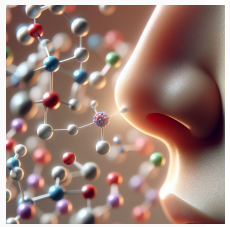
FROM FOOD FLAVOUR TO SENSORY PERCEPTION

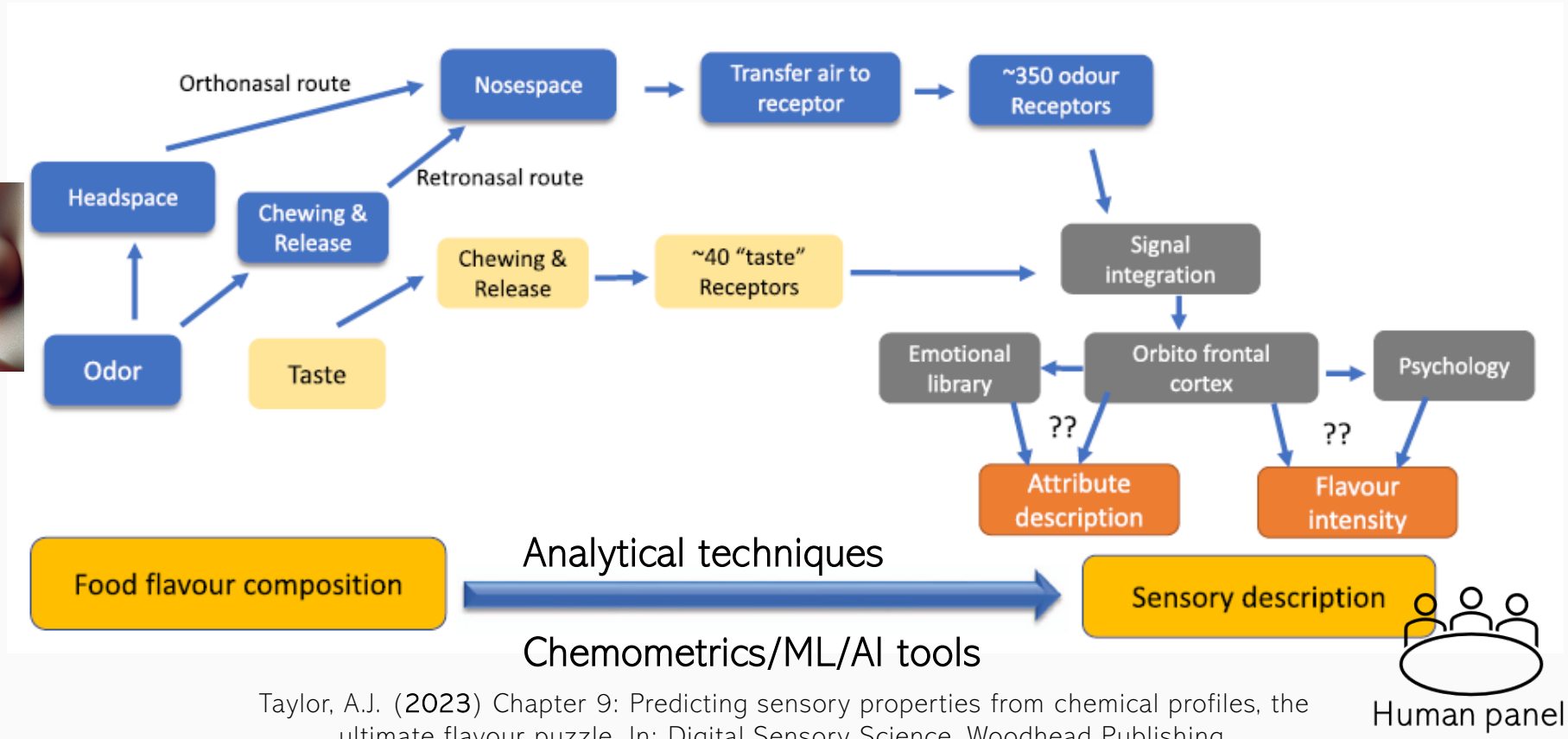




Taylor, A.J. (2023) Chapter 9: Predicting sensory properties from chemical profiles, the ultimate flavour puzzle. In: Digital Sensory Science, Woodhead Publishing.

Digitalizing flavour

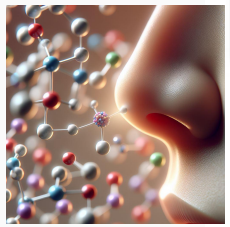


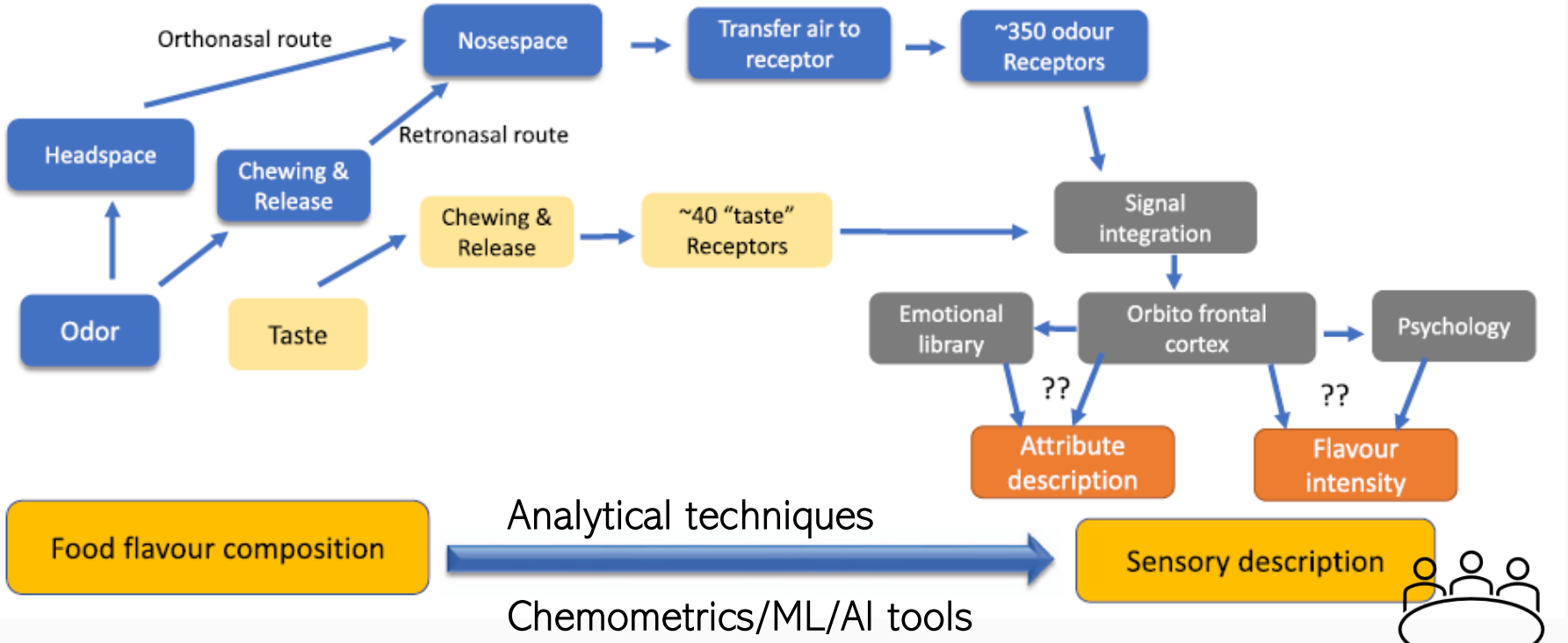


Taylor, A.J. (2023) Chapter 9: Predicting sensory properties from chemical profiles, the ultimate flavour puzzle. In: Digital Sensory Science, Woodhead Publishing.

Digitalizing flavour: why is this relevant?

To optimize food products design and processing
To better understand and control off-flavour formation





Taylor, A.J. (2023) Chapter 9: Predicting sensory properties from chemical profiles, the ultimate flavour puzzle. In: Digital Sensory Science, Woodhead Publishing.

DIGITALIZING FLAVOUR

Chemical subspace

Sensory perception subspace

Volatilomics (*untargeted*)

- GC-MS (+ PARAFAC2; Rasmus Bro's talk)
- GCxGC-MS

DIGITALIZING FLAVOUR

Food flavour composition → Sensory description

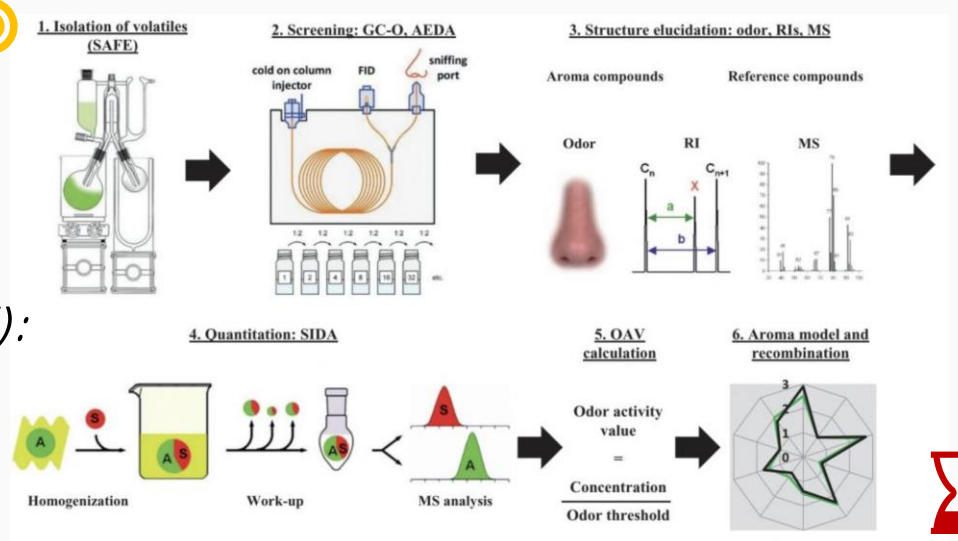
Chemical subspace — Sensory perception subspace

Volatilomics (*untargeted*)

- GC-MS (+ PARAFAC2; Rasmus Bro's talk)
- GCxGC-MS

Sensomics method *(targeted)*: identification and quantification of odour-active compounds

Unified method *(targeted)*:
Odorants and tastants
By UHPLC-MS/MS



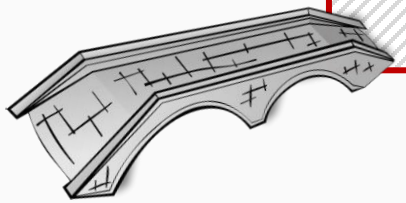
Granvogl, M., Schieberle, P. (2022) Chapter 2: The sensomics approach (...). In: Comprehensive Analytical Chemistry, volume 96, ElSevier B.V.

DIGITALIZING FLAVOUR

Food flavour composition → Sensory description

Chemical subspace

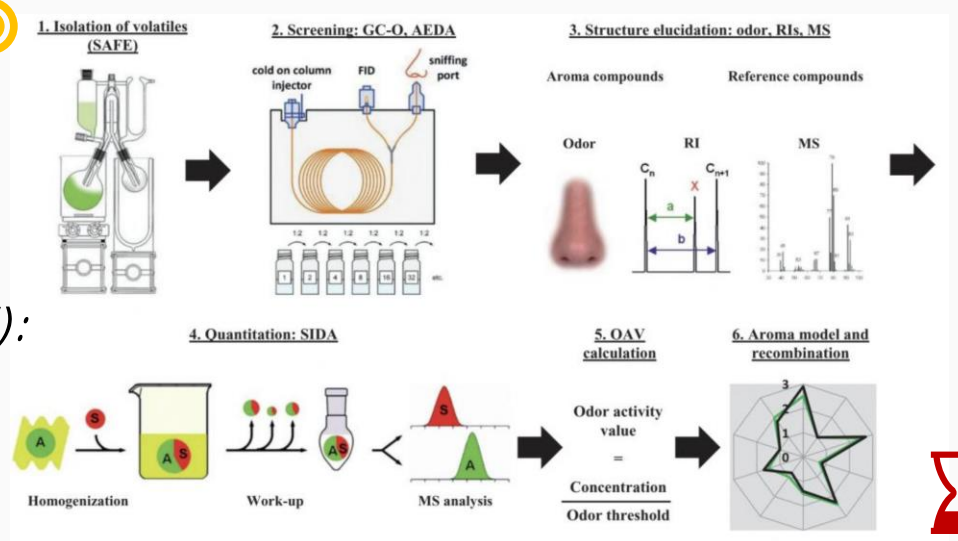
Volatilomics (*untargeted*)
- GC-MS (+ PARAFAC2; Rasmus Bro's talk)
- GCxGC-MS



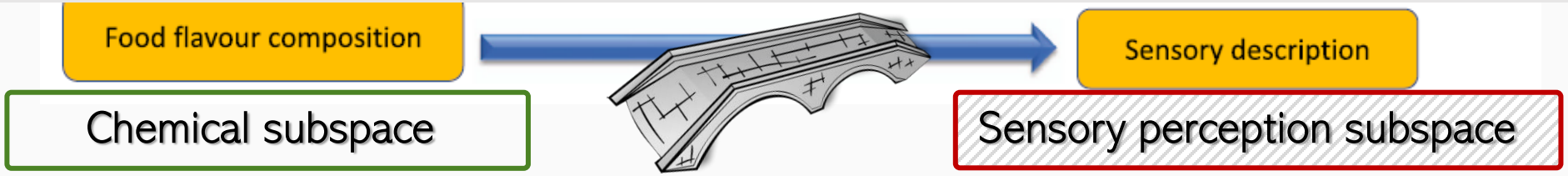
Sensory perception subspace

Sensomics method *(targeted)*: identification and quantification of odour-active compounds

Unified method *(targeted)*:
Odorants and tastants
By UHPLC-MS/MS

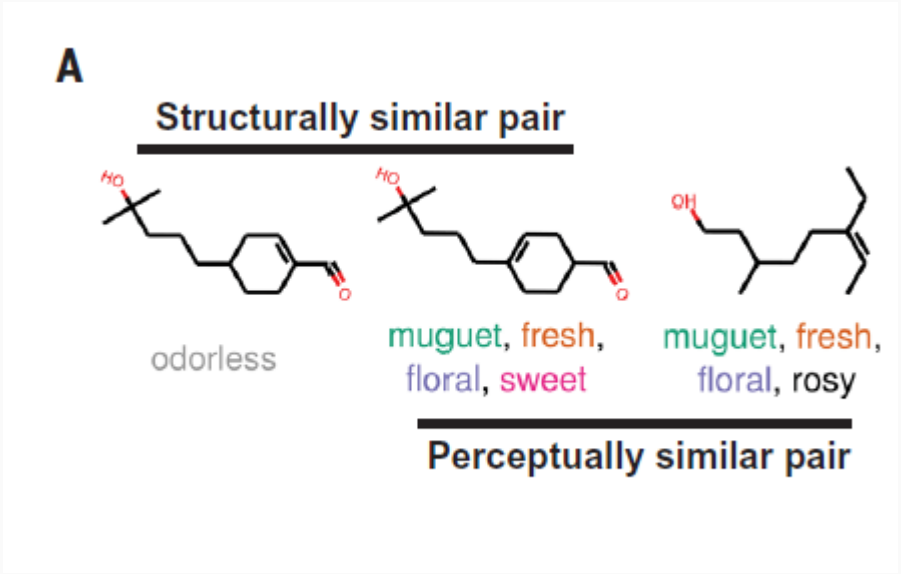


Granvogl, M., Schieberle, P. (2022) Chapter 2: The sensomics approach (...). In: Comprehensive Analytical Chemistry, volume 96, ElSevier B.V.

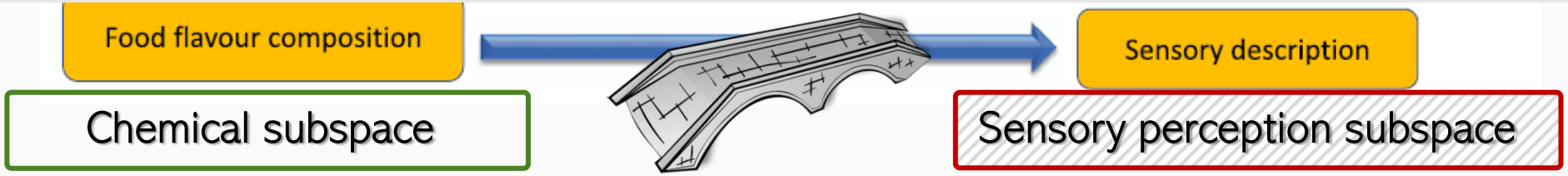


Molecule-based approaches:

- **QSPR:** chemical-physical parameters (vapour pressure, hydrophobicity, solubility, partition coefficient).



DIGITALIZING FLAVOUR



Molecule-based approaches:

- **QSPR:** chemical-physical parameters (vapour pressure, hydrophobicity, solubility, partition coefficient.

- Relevant studies/projects:

ODORIFY: DNN to identify what molecules bind to human odour receptors. Potential prediction of odorants interaction.

JBC RESEARCH ARTICLE

OdoriFy: A conglomerate of artificial intelligence–driven prediction engines for olfactory decoding

Received for publication, March 22, 2021, and in revised form, June 24, 2021 Published, Papers in Press, July 12, 2021,
https://doi.org/10.1016/j.jbc.2021.100956

Ria Gupta[1,‡], Aayushi Mittal[1,‡], Vishesh Agrawal[1,‡], Sushant Gupta[1], Krishan Gupta[2], Rishi Raj Jain[3], Prakriti Garg[1], Sanjay Kumar Mohanty[1], Riya Sogani[1], Harshit Singh Chhabra[1], Vishakha Gautam[1], Tripti Mishra[4], Debarka Sengupta[1,2,5,6], and Gaurav Ahuja[1,*]

From the [1]Department of Computational Biology, [2]Department of Computer Science and Engineering, [3]Department of Computer Science and Design, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), New Delhi, India; [4]Pathfinder Research and Training Foundation, Greater Noida, Uttar Pradesh, India; [5]Centre for Artificial Intelligence, Indraprastha Institute of Information Technology, New Delhi, India; [6]Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia

OSMO's AI scent platform

Graph Neural Network ~5000 molecules with odour labels from flavour DB

NEUROSCIENCE

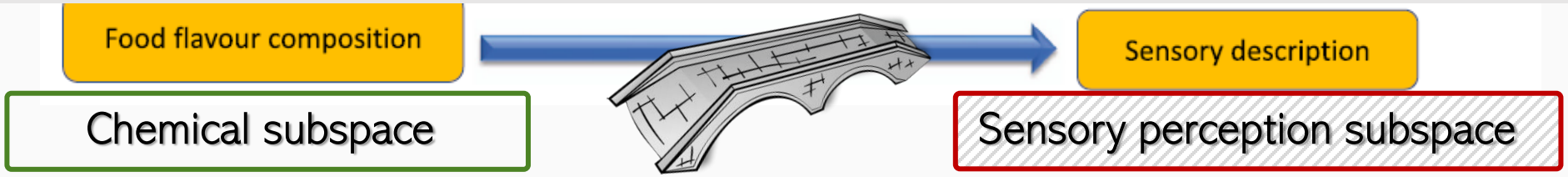
A principal odor map unifies diverse tasks in olfactory perception

Brian K. Lee[1]†, Emily J. Mayhew[2,3]†, Benjamin Sanchez-Lengeling[1], Jennifer N. Wei[1], Wesley W. Qian[4,1,5], Kelsie A. Little[2], Matthew Andres[2], Britney B. Nguyen[2], Theresa Moloy[2], Jacob Yasonik[4,1], Jane K. Parker[6], Richard C. Gerkin[4,1,7], Joel D. Mainland[2,8]*, Alexander B. Wiltschko[4,1]*

Mapping molecular structure to odor perception is a key challenge in olfaction. We used graph neural networks to generate a principal odor map (POM) that preserves perceptual relationships and enables odor quality prediction for previously uncharacterized odorants. The model was as reliable as a

Single molecules
Ratios and mixtures play a role

DIGITALIZING FLAVOUR

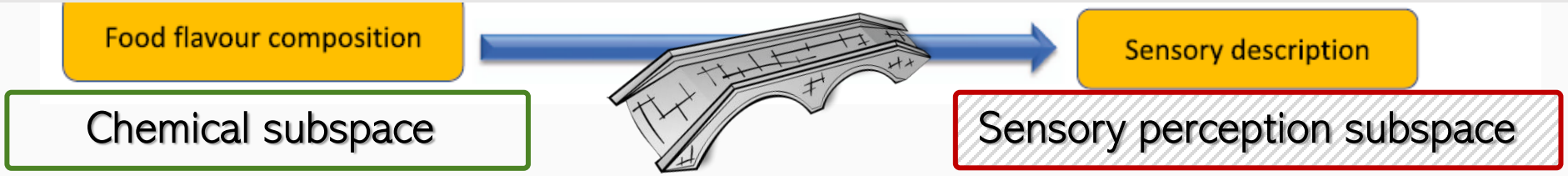


Volatilomics Untargeted → High-throughput data: training of flavour prediction ML models

- GC-MS (+ PARAFAC2; Rasmus Bro's talk)
- GCxGC-MS
  (Cozzolino, 2022)

+ Other measurements non-volatile composition
(Spectroscopy: IR, Fluorescence, Raman, NMR)

DIGITALIZING FLAVOUR: A TEXT ANALYSIS APPROACH



Volatilomics Untargeted → High-throughput data: training of flavour prediction ML models

- GC-MS (+ PARAFAC2; Rasmus Bro's talk)
- GCxGC-MS
  (Cozzolino, 2022)

+ Other measurements non-volatile composition
(Spectroscopy: IR, Fluorescence, Raman, NMR)

↑

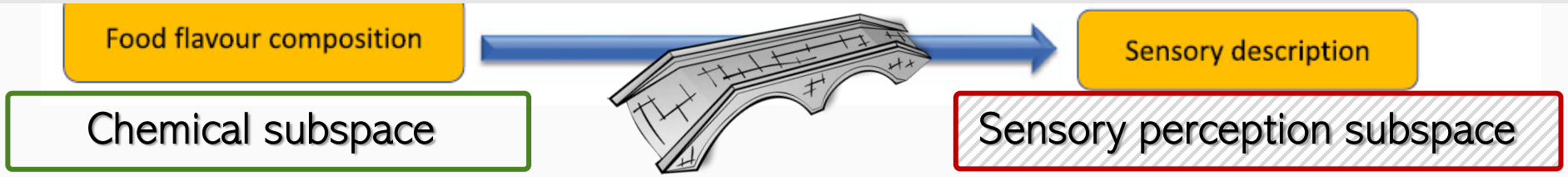
Text data

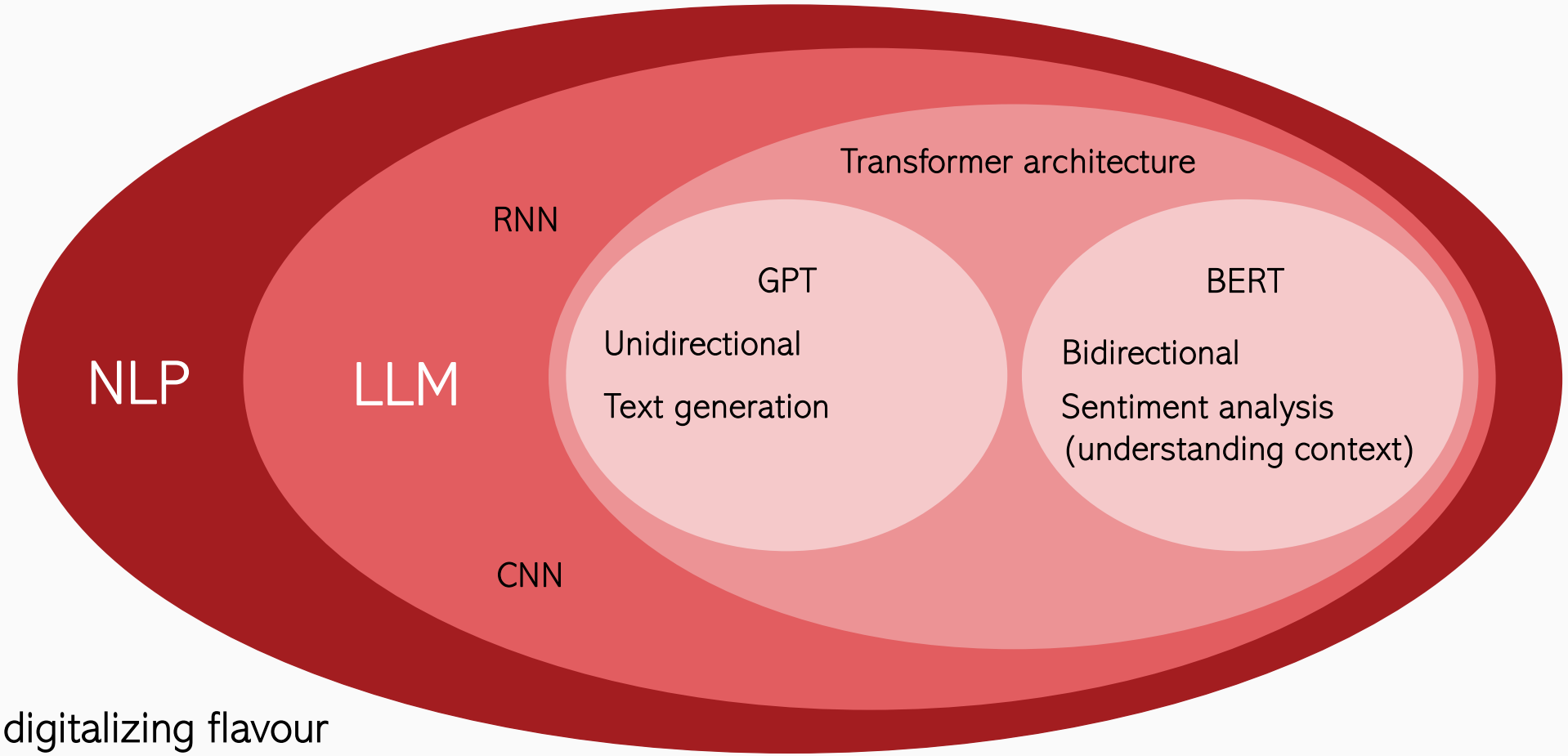
**Key role in harmonization process for*
The vocabulary used in sensory assessment

Background knowledge not actively used in both chemical composition and sensory assessment, available at:

- Scientific papers
- Books
- Databases: flavour and physico-chemical meta-data (e.g., flavorDB and Pubchem)

DIGITALIZING FLAVOUR: A TEXT ANALYSIS APPROACH





NLP for digitalizing flavour

DIGITALIZING FLAVOUR: A TEXT ANALYSIS APPROACH

Food flavour composition

Chemical subspace

Sensory description

Sensory perception subspace

Pilot structure

Interface

Large Language Model

Data querying

Internal Data

Public Data

amass

Amass' Scientific assistant

GPT

Unidirectional

Text generation

BERT

Bidirectional

Sentiment analysis (understanding context)

bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

PubMed

KU FOOD pilot

Internal PDF's, accessed by referring to "internal files" as part of questions.

Review and research papers from the main scientific journals in FOOD SCIENCE

~ 1 milon vectors

DIGITALIZING FLAVOUR: A TEXT ANALYSIS APPROACH

Food flavour composition → Sensory description

Chemical subspace → Sensory perception subspace

Pilot structure

Interface

Large Language Model

Data querying

Internal Data | Public Data

amass

Amass' Scientific assistant

GPT

Unidirectional

Text generation

BERT

Bidirectional

Sentiment analysis (understanding context)

KU FOOD pilot

Internal PDF's, accessed by referring to "internal files" as part of questions.

Review and research papers from the main scientific journals in FOOD SCIENCE

~ 1 milon vectors

bioRxiv THE PREPRINT SERVER FOR BIOLOGY

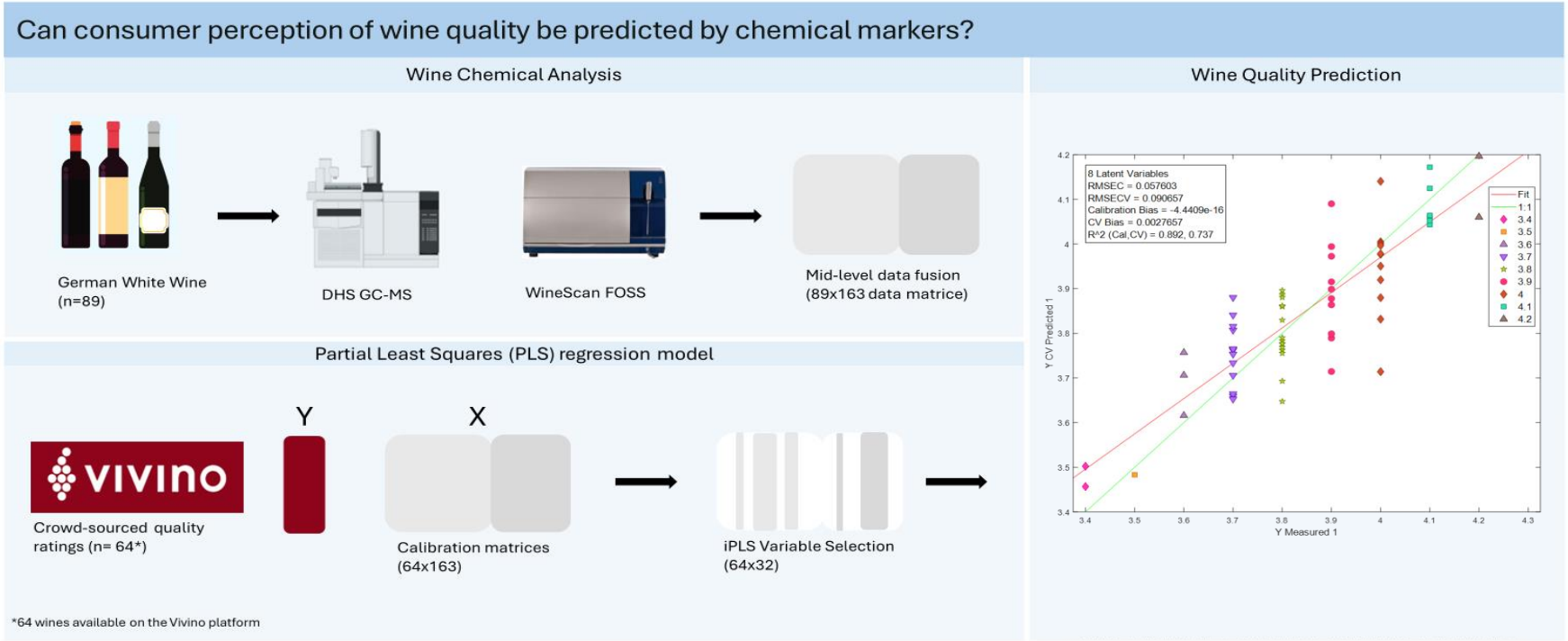
PubMed

Challenges:

- Unstructured data: mainly as text, but also in tables and figures.
- Symbols and Greek letters are often present in chemical names.
- Compound names are not unique

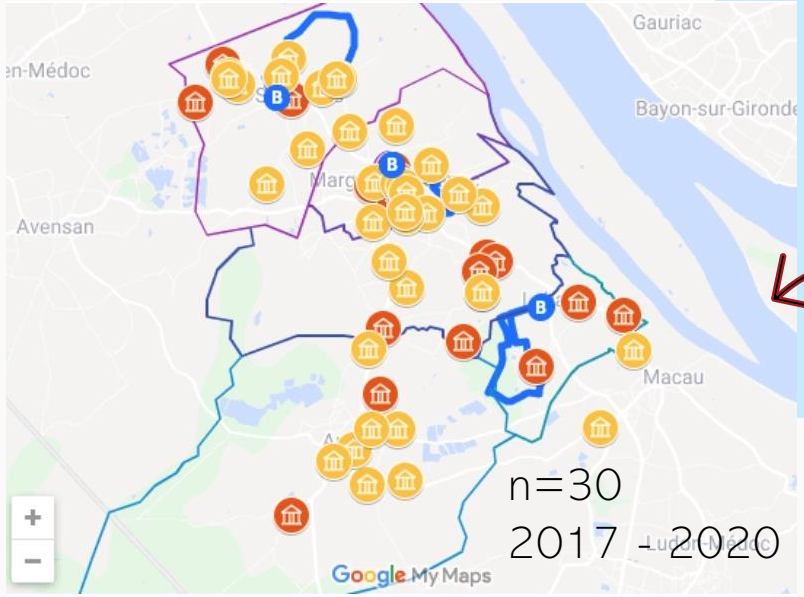
Preliminary study (Hjertholm/Götz et al., submitted for publication)

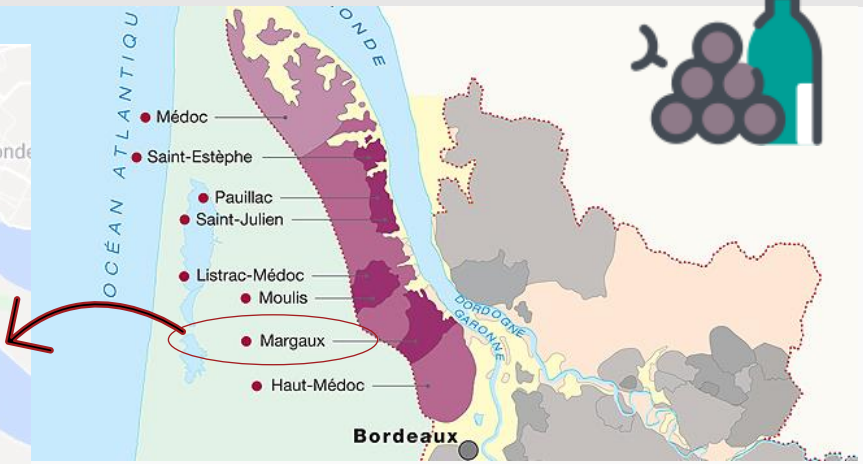
Frederikke Hjertholm's flash presentation



CASE-STUDY: WINE PROJECT

The Margaux wine project



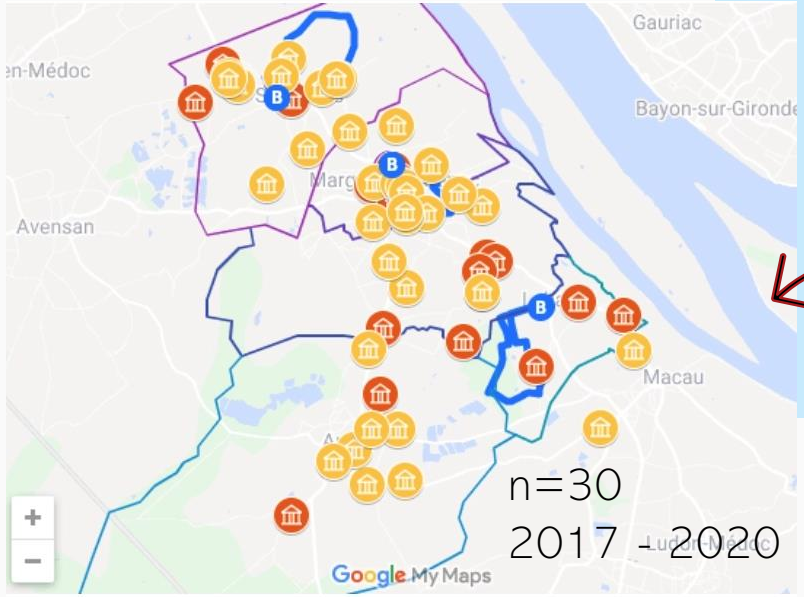


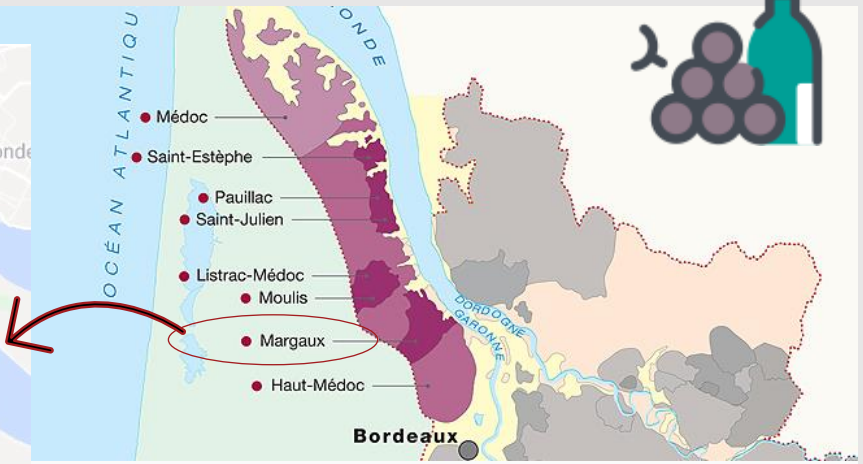
- 1855 classification
- Cru bourgeoise classifications

+ Non-alcoholic wines (*Frederikke's PhD*)

CASE-STUDY: WINE PROJECT

The Margaux wine project





- 1855 classification
- Cru bourgeoise classifications

+ Non-alcoholic wines (*Frederikke's PhD*)

Analytical measurements:

Volatile profile: GC-MS and PTR-TOF

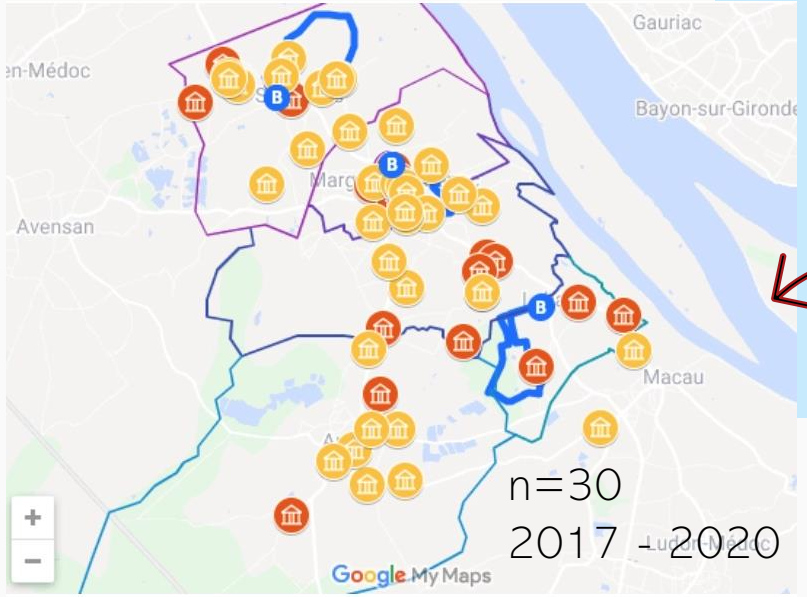
Phenolic composition: LC-HRMS and EEM

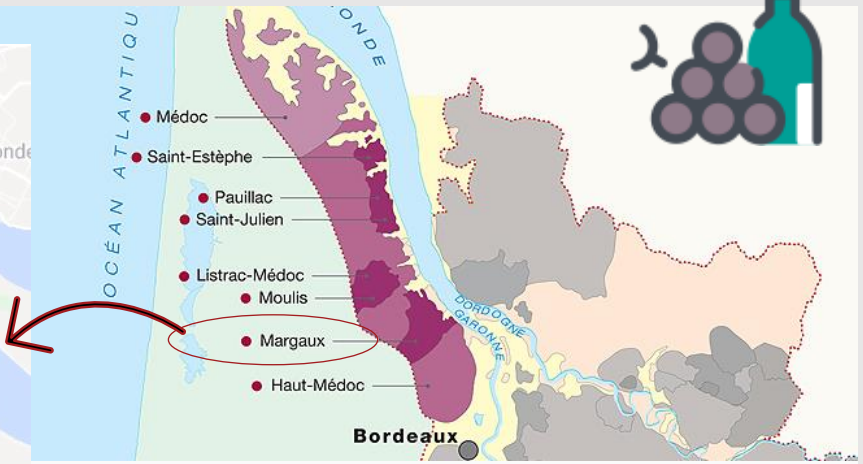
Metabolites composition: NMR

General quality control physicochemical parameters: FTIR (Winescan®, Foss, Hillerød DK

CASE-STUDY: WINE PROJECT

The Margaux wine project





- 1855 classification
- Cru bourgeoise classifications

+ Non-alcoholic wines (*Frederikke's PhD*)

Analytical measurements:

Volatile profile: GC-MS and PTR-TOF

Phenolic composition: LC-HRMS and EEM

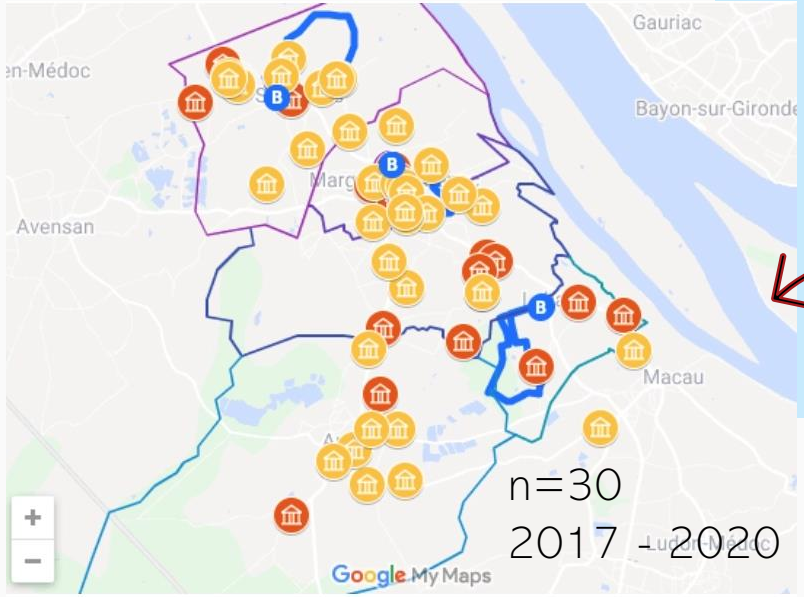
Metabolites composition: NMR

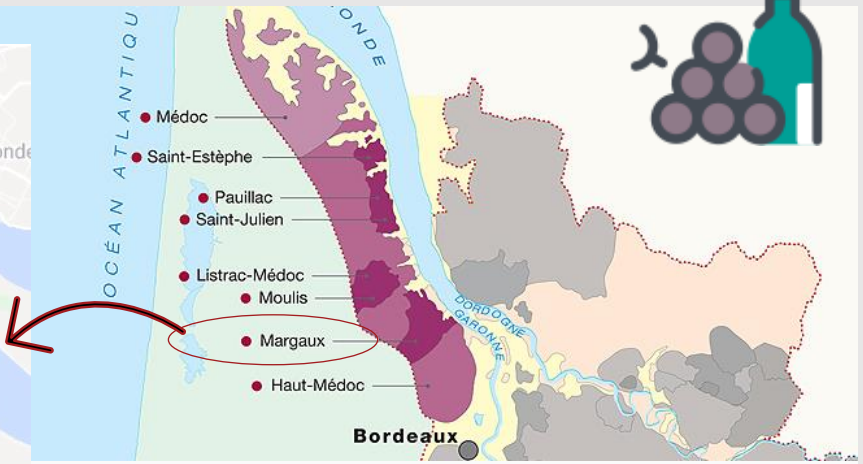
Chemometric methods development *(Julius J. Terp's PhD)*

General quality control physicochemical parameters: FTIR (Winescan®, Foss, Hillerød DK

CASE-STUDY: WINE PROJECT

The Margaux wine project





- 1855 classification
- Cru bourgeoise classifications

+ Non-alcoholic wines (*Frederikke's PhD*)

Analytical measurements:

Volatile profile: GC-MS and PTR-TOF

Phenolic composition: LC-HRMS and EEM

Metabolites composition: NMR

Chemometric methods development *(Julius J. Terp's PhD)*

Sensory analysis:

Panel of wine experts

RATA method – ratings of detected descriptors

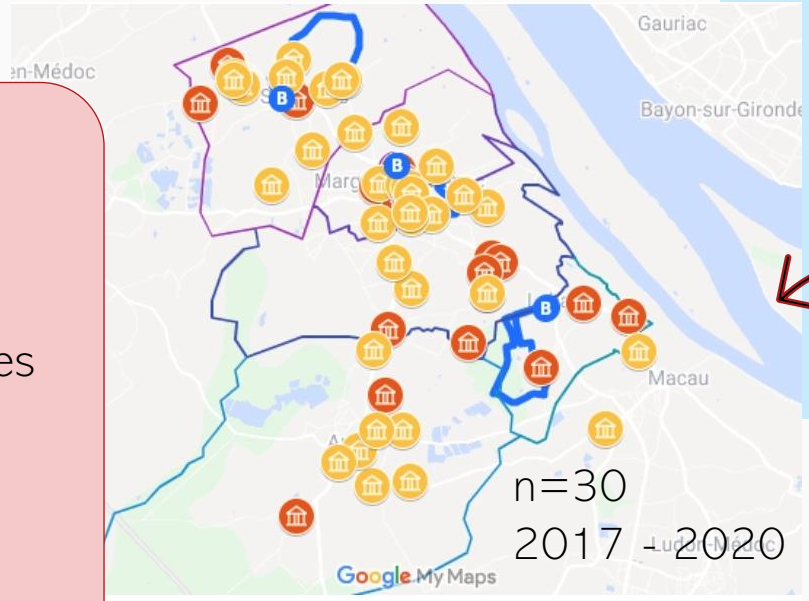
General quality control physicochemical parameters: FTIR (Winescan®, Foss, Hillerød DK

CASE-STUDY: WINE PROJECT

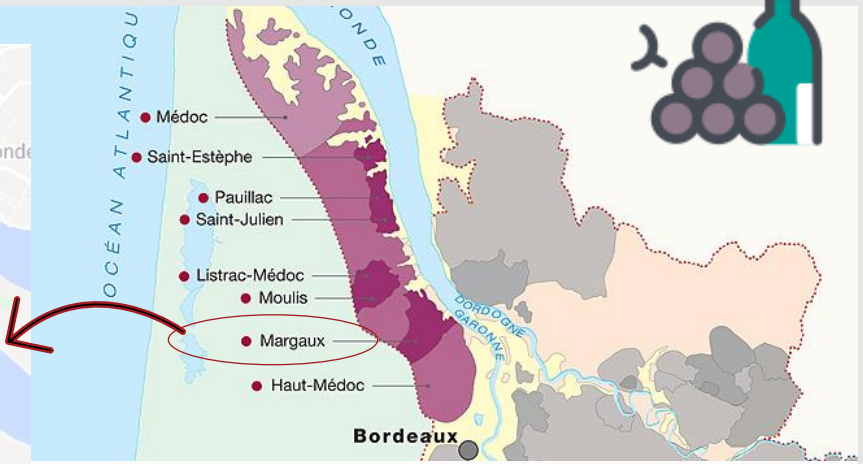
The Margaux wine project

Text analysis (BERT models)

- ✓ Meta-data from wine samples
- ✓ Expert reviews (numerical scores and tasting notes) from major sources (Decanter, Jancis Robinson, Wine Spectator)
- ✓ Background knowledge from scientific papers, DB, books



n=30
2017 - 2020



- 1855 classification
- Cru bourgeoise classifications

Non-alcoholic wines (*Frederikke's PhD*)

Analytical measurements:

Volatile profile: GC-MS and PTR-TOF

Phenolic composition: LC-HRMS and EEM

Metabolites composition: NMR

Chemometric methods development *(Julius J. Terp's PhD)*

Sensory analysis:

Panel of wine experts

RATA method – ratings of detected descriptors

General quality control physicochemical parameters: FTIR (Winescan®, Foss, Hillerød DK

Tusind tak!

Bridging aroma chemistry and sensory perception: a text analysis approach

Beatriz Quintanilla Casas

Postdoctoral researcher
Chemometrics and Machine Learning research group
Department of Food Science, University of Copenhagen

KØBENHAVNS UNIVERSITET

This work was supported by **Danish Data Science Academy** by means of the Postdoctoral fellowship awarded to B. Quintanilla-Casas, funded by the **Novo Nordisk Foundation** (**NNF21SA0069429**).

Rasmus Bro

Frederikke H. Nielsen

Julius J. Terp

& Other colleagues from KU working with wine/aroma/chemometrics

November 8th, 2024
DSK.2024 (Middelfart, Denmark)