Organic and Biological Analytical Chemistry Group

Machine learning, the separation scientist acolyte for innovative method development

Pierre-Hugues Stefanuto, Meriem Gaida, and friends

Molecular System, Organic & Biological Analytical Chemistry Group, Liège University, Belgium

Acknowledgements

OBIACHEM
Organic and Biological Analytical Chemistry Group

LIÈGE université
MolSys

Prof. Jef Focant

Anais Rodrigues

Thibaut Dejong

Thibault Massenet

Anika Lokker

Géraldine Dumont

Djulia Bensaada

Dr Kinjal Bhatt

Dr. Armélinda Agnello

Stéphanie Hendrick

Grégory Gridelet

Vicky Borguet

Roberto

TRAJAN

LECO

FDA

Fondation Léon Fredericq

fnrs
LA LIBERTÉ DE CHERCHER

NIST

Disclaimer: *This presentation has been generated using AI-Tools*

What is a separation scientist acolyte?

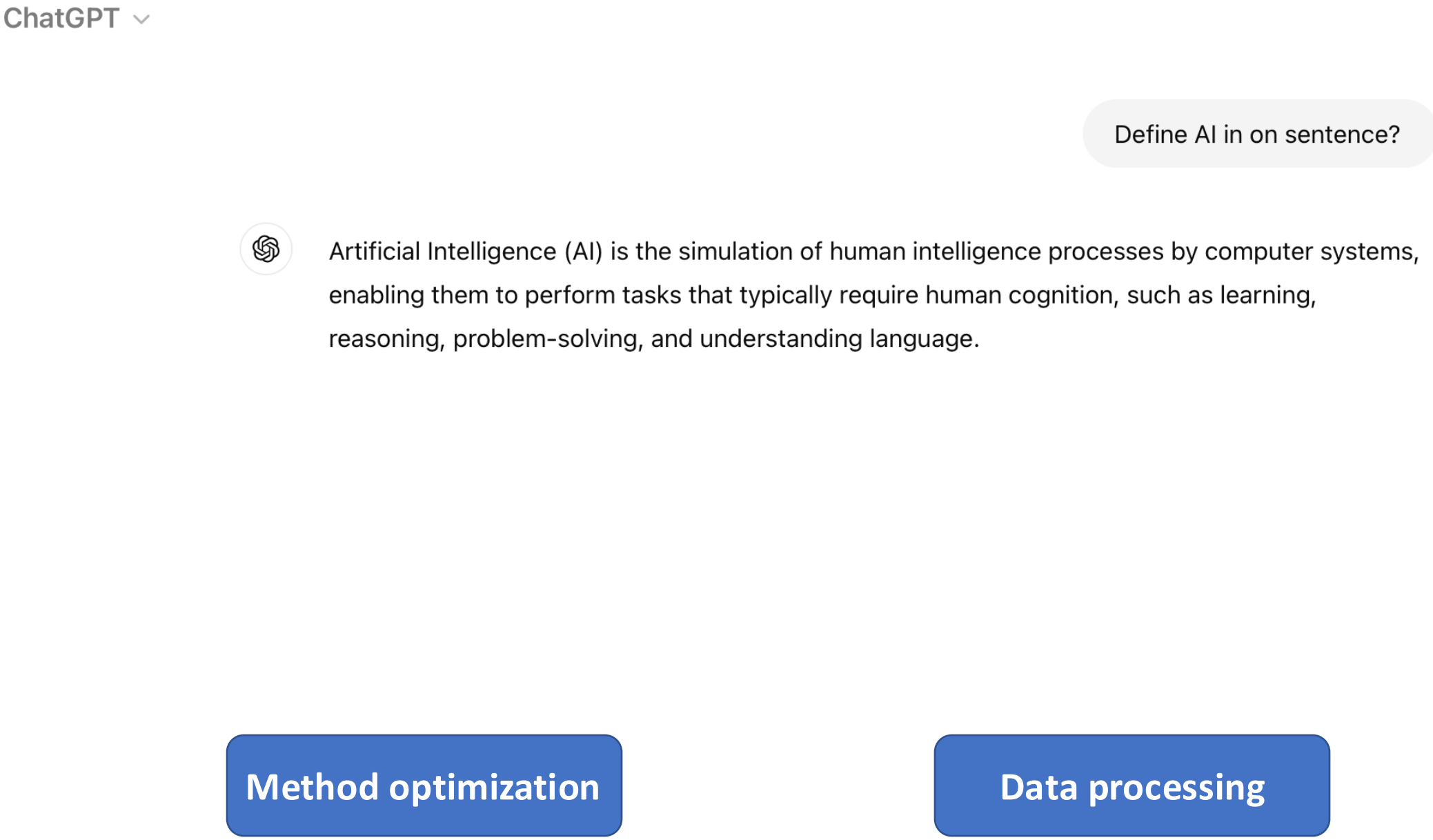
*An **acolyte** refer to a devoted follower or **assistant of a prominent figure**. An acolyte is someone who **supports and learns from a more experienced person**, sometimes in political, academic, or professional environments.*

*AI and Machine learning is a **support to separation scientist**, but it needs to learn from **high quality information.***

ChatGPT

Define AI in on sentence?

Text generated from ChatGPT



Text generated from ChatGPT

GPT concept

GPT: Generative **P**re-trained **T**ransformer

Generative: *The model is generating information* → What we ask

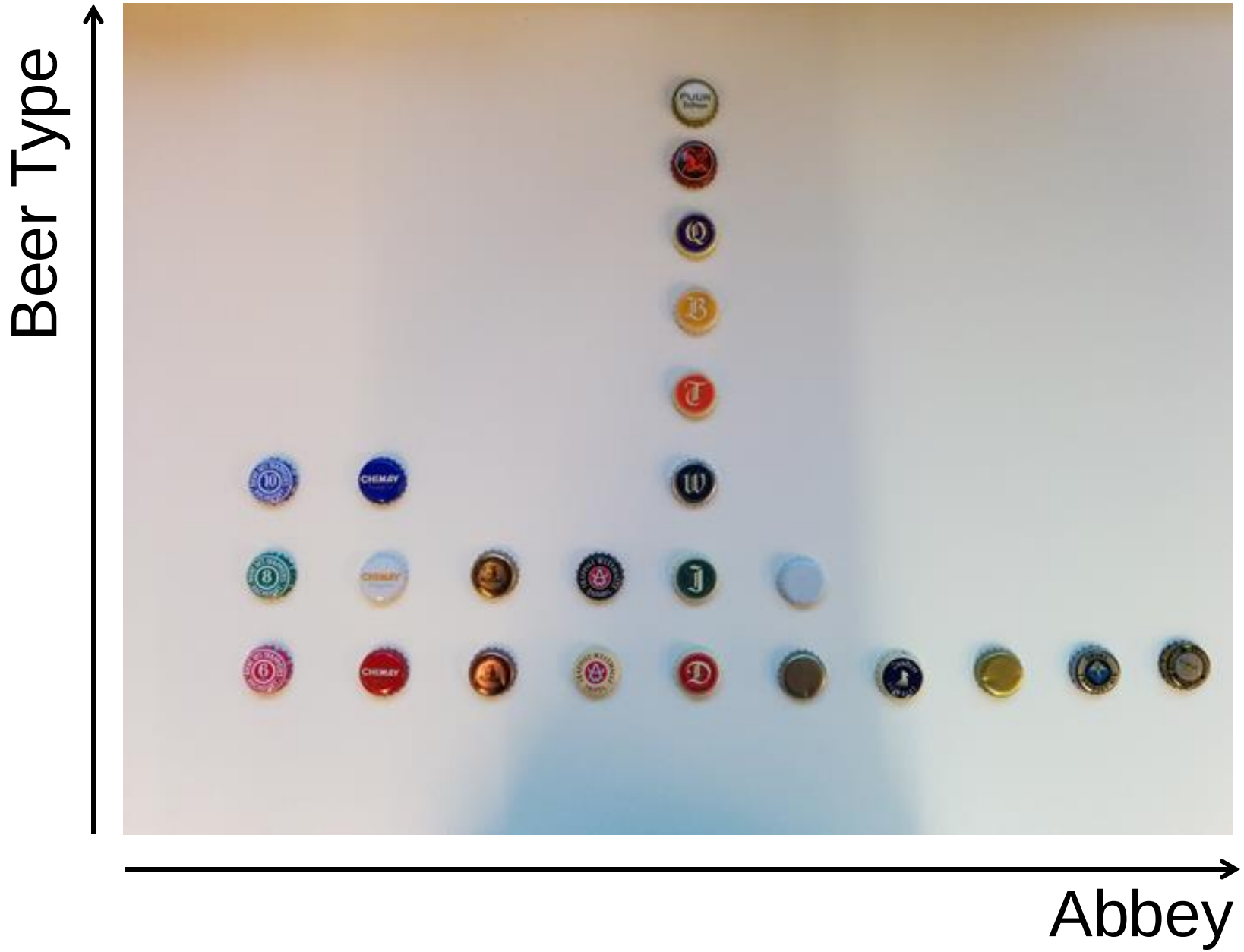
Pre-trained: *The model has been trained on a large amount of data* → The quality of the training

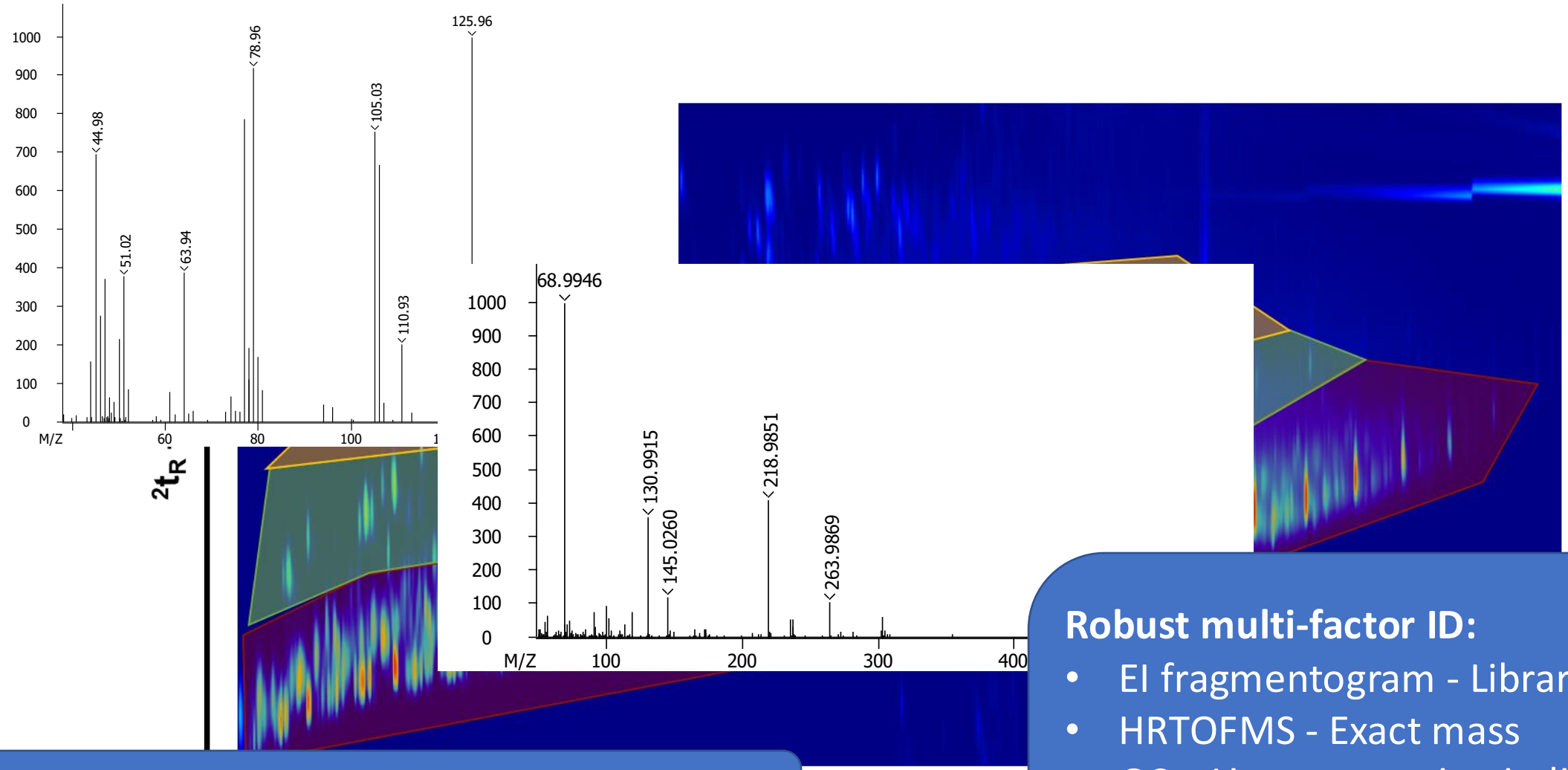
Transformer: *The underlying architecture of neural network model* → The way he manage the information

GC×GC-TOFMS: the principle

The sample

Abbey


Beer Type
Abbey



How to handle the method development?

How to manage the data processing?

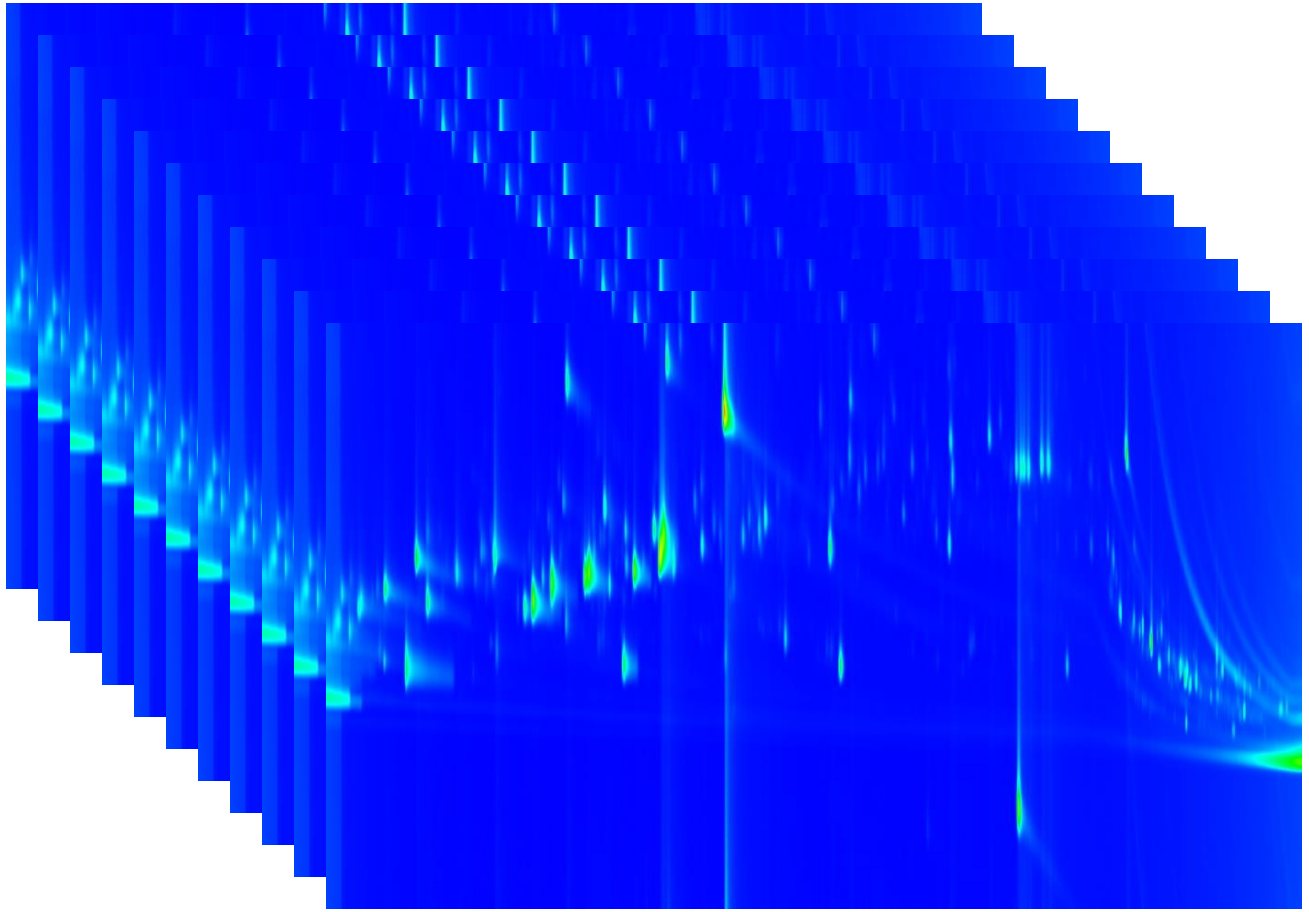
Robust multi-factor ID:

- EI fragmentogram - Library match
- HRTOFMS - Exact mass
- GC – Linear retention indices
- GCxGC – Elution pattern

→ **Level 2 MSI ID**

Generative I: Method optimization

LECO Simply GC×GC®



Generative I: Method optimization

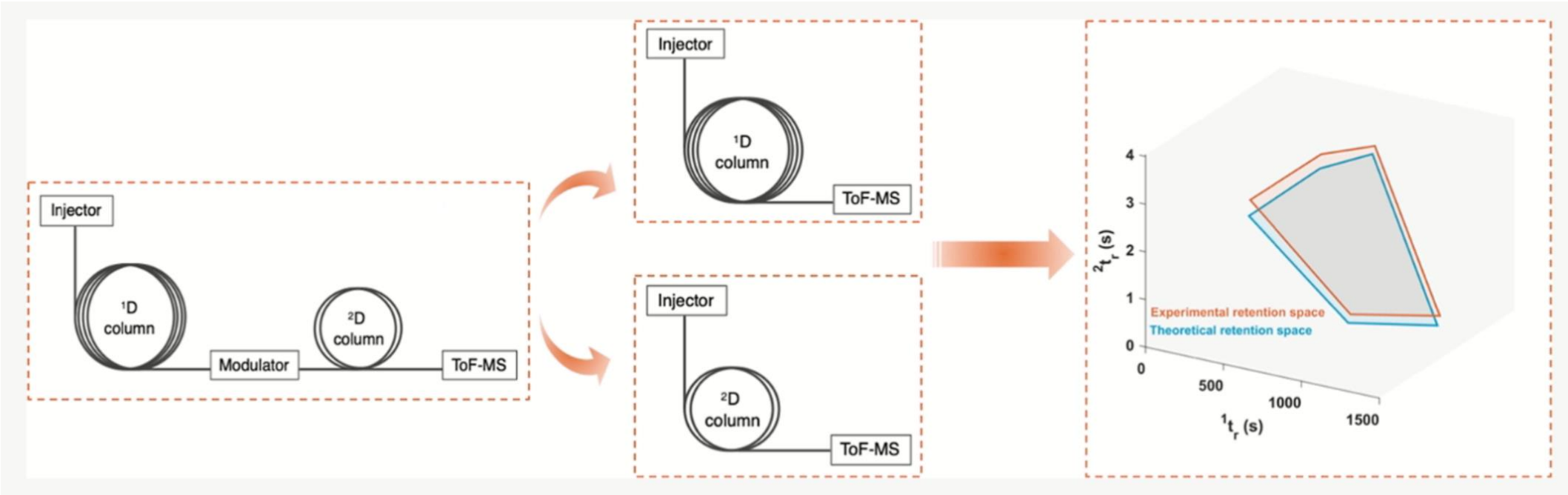
analytical chemistry

pubs.acs.org/ac Article

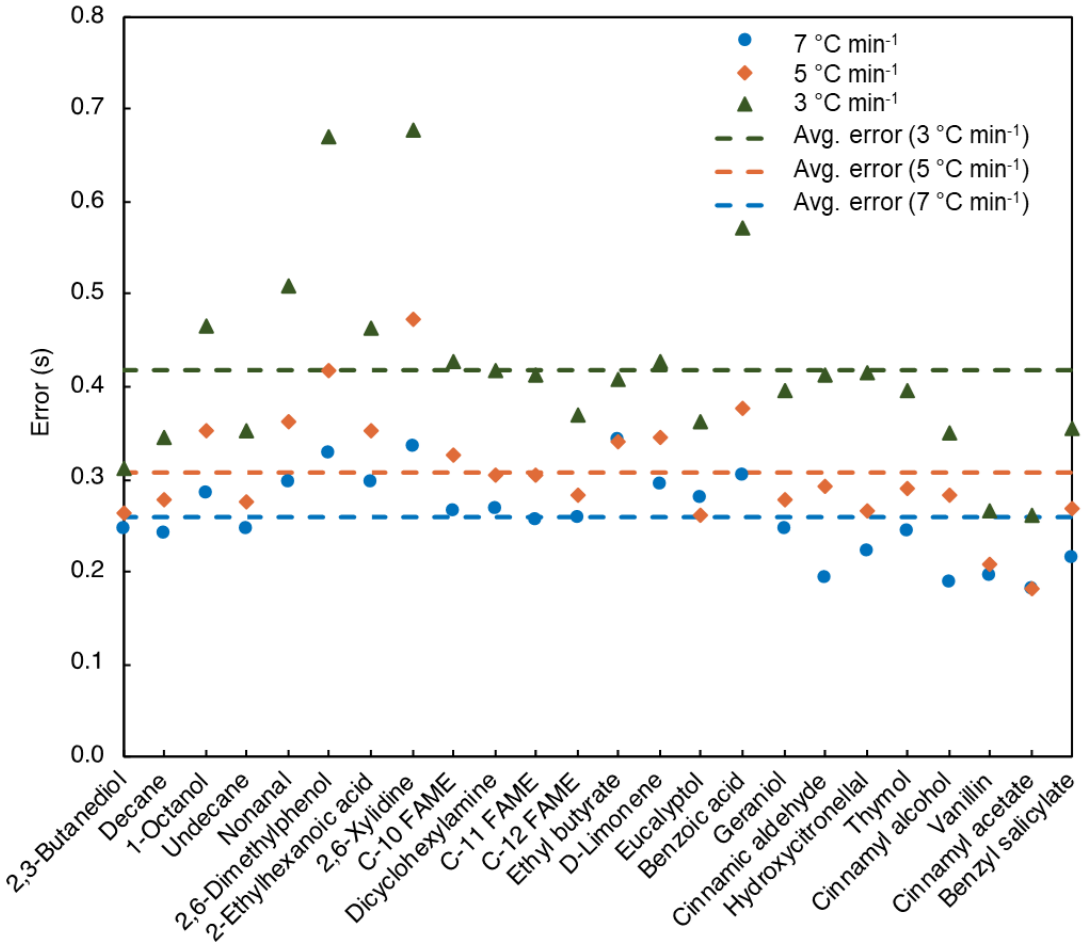
Top-Down Approach to Retention Time Prediction in Comprehensive Two-Dimensional Gas Chromatography–Mass Spectrometry

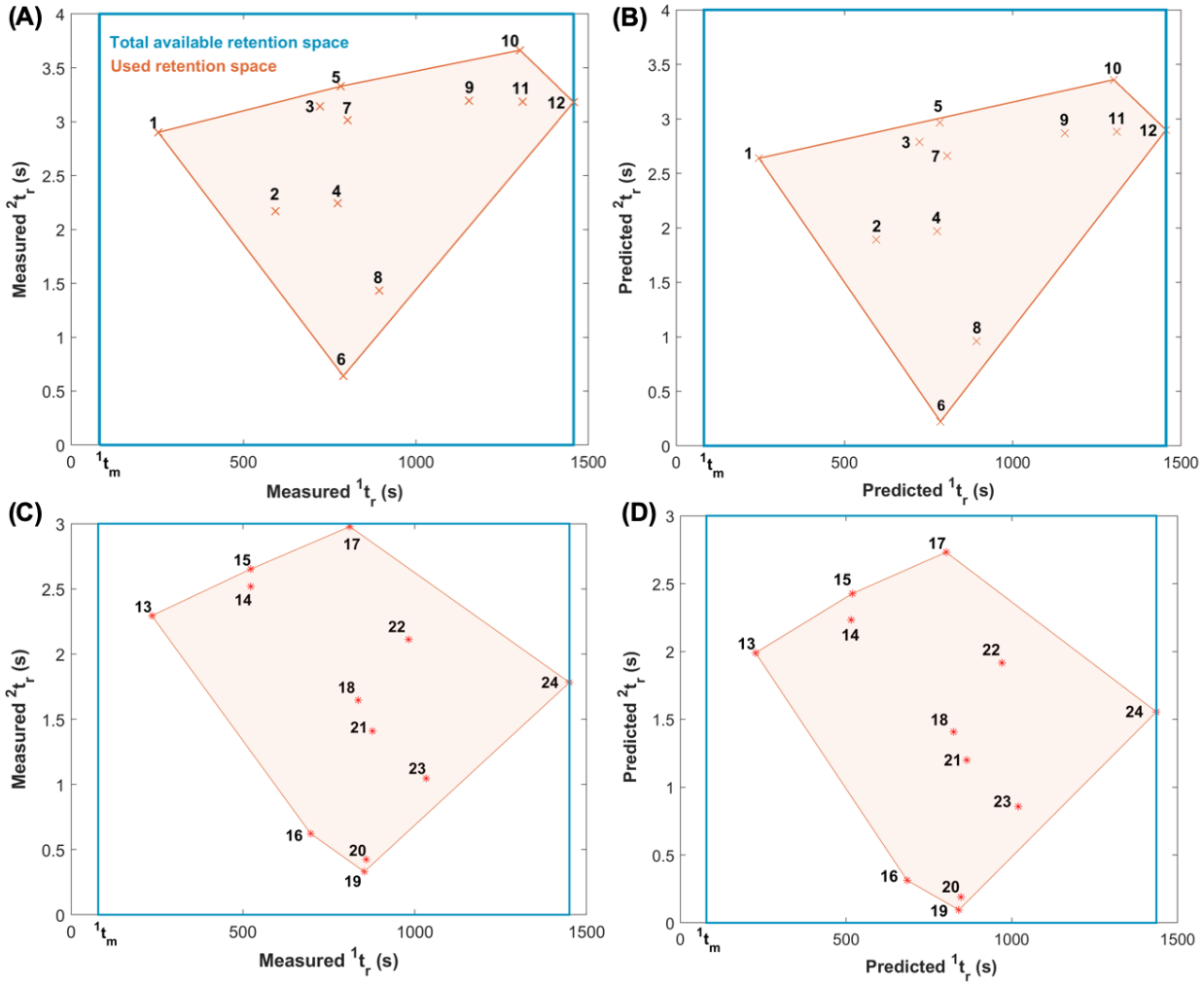
Meriem Gaida,* Flavio A. Franchina, Pierre-Hugues Stefanuto, and Jean-François Focant

Cite This: *Anal. Chem.* 2022, 94, 17081–17089 Read Online

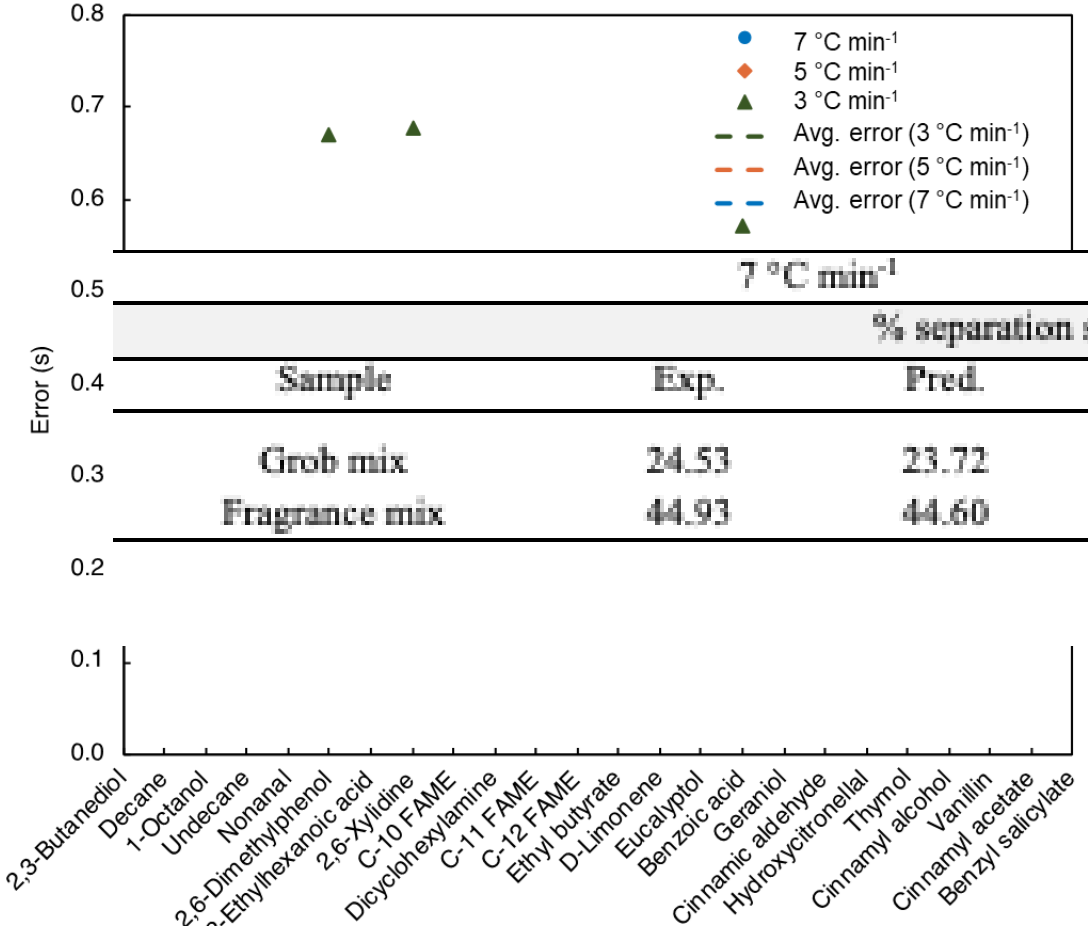


Generative I: Method optimization

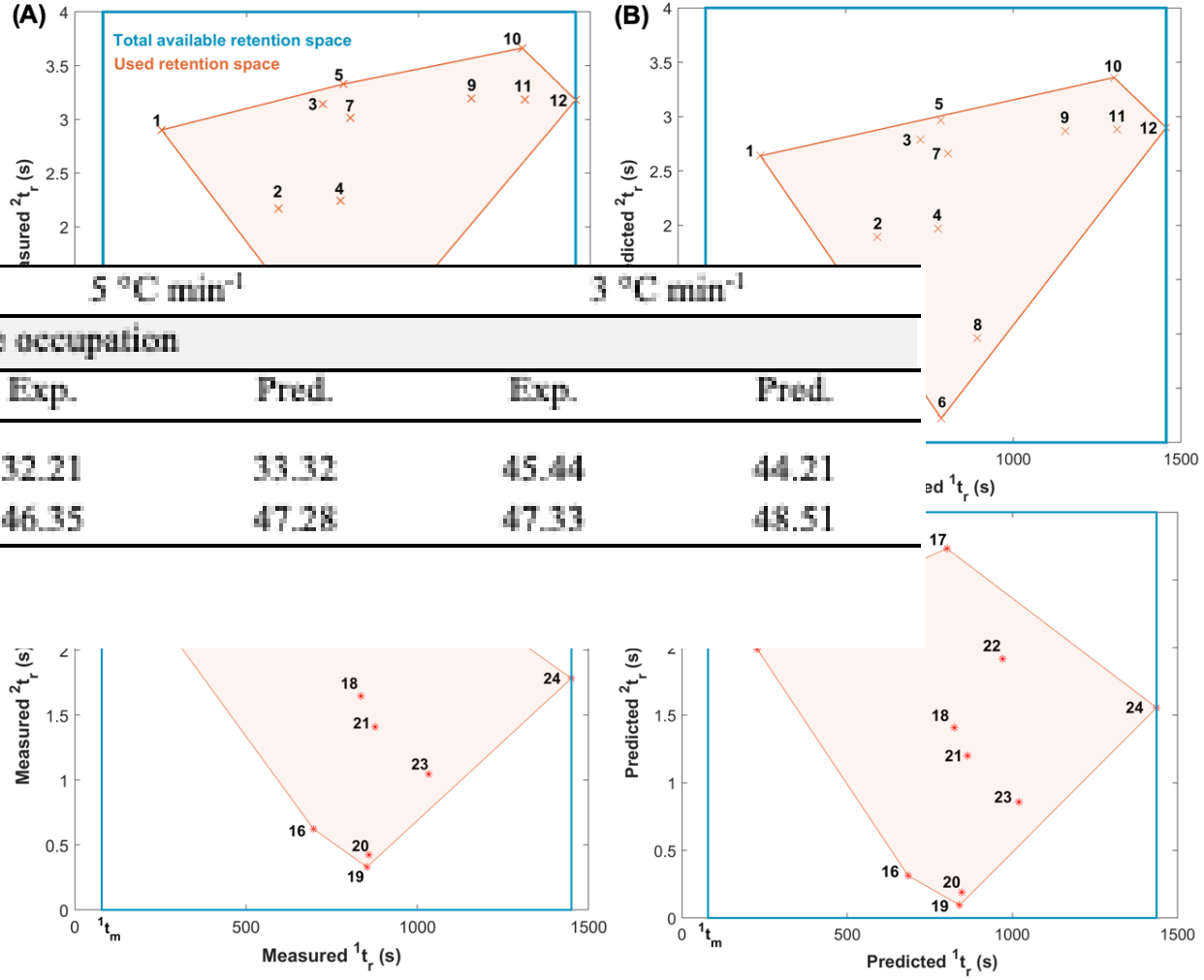




Generative I: Method optimization



% separation space occupation						
	7 °C min^{-1}		5 °C min^{-1}		3 °C min^{-1}	
Sample	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.
Grob mix	24.53	23.72	32.21	33.32	45.44	44.21
Fragrance mix	44.93	44.60	46.35	47.28	47.33	48.51



Generative I: Method optimization

RETURN TO ISSUE | < PREV ECOTOXICOLOGY AND PU... NEXT >

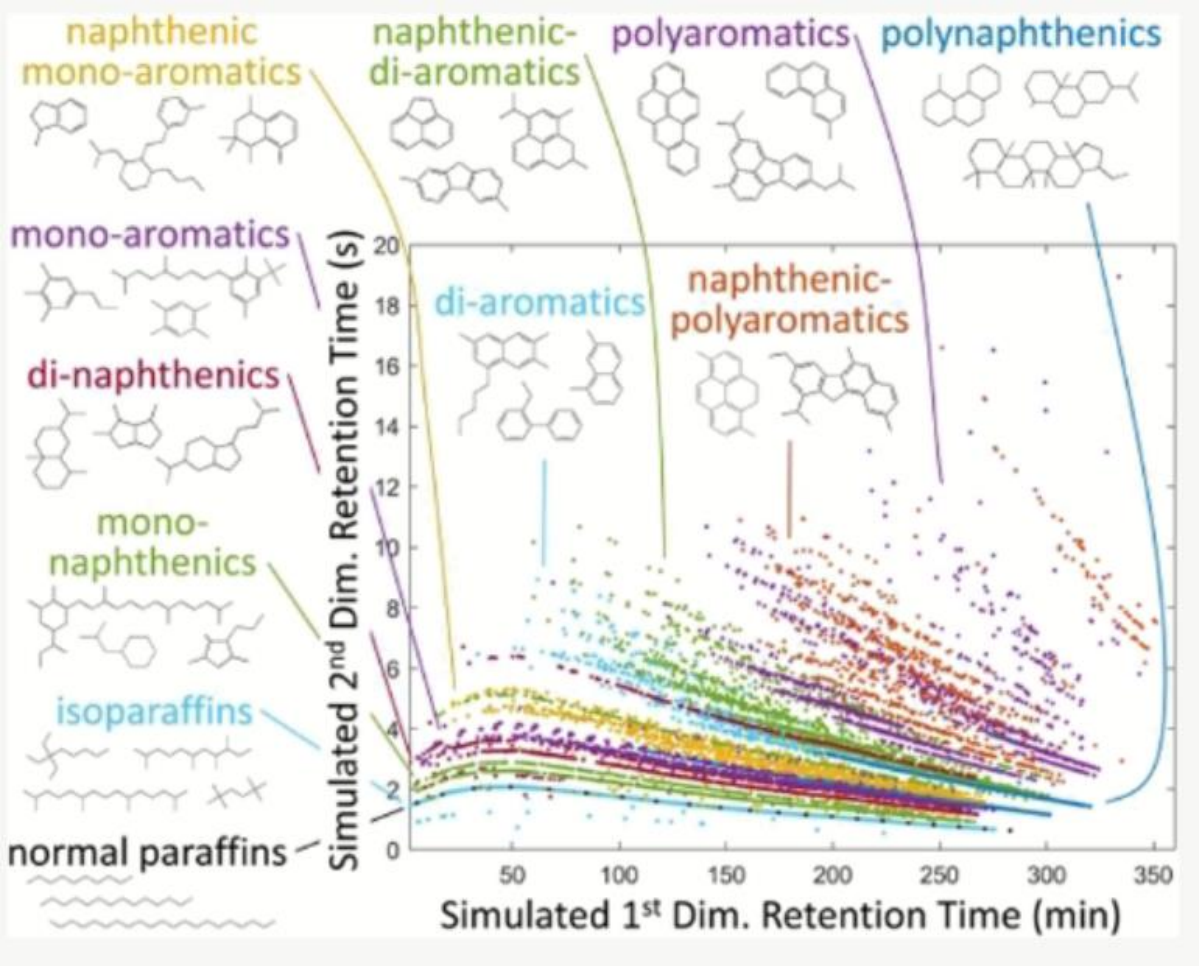
Modeling the GCxGC Elution Patterns of a Hydrocarbon Structure Library To Innovate Environmental Risk Assessments of Petroleum Substances

J. Samuel Arey, Alberto Martin Aparicio, Eleni Vaiopoulou, Stuart Forbes, and Delina Lyon*

Cite this: *Environ. Sci. Technol.* 2022, 56, 24, 17913−17923
Publication Date: December 7, 2022
https://doi.org/10.1021/acs.est.2c06922
Copyright © 2022 The Authors. Published by American Chemical Society. This publication is licensed under CC-BY-NC-ND 4.0.
Open Access

Article Views 1879 | Altmetric 7 | Citations 1
LEARN ABOUT THESE METRICS

PDF (5 MB) Get it@ULiège SI Supporting Info (2) »



Bringing AI in the game!

Journal of Chromatography A 1612 (2020) 460661

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Steroid identification via deep learning retention time predictions and two-dimensional gas chromatography-high resolution mass spectrometry

Giuseppe Marco Randazzo[a,*], Andrea Bileck[b], Andrea Danani[a], Bruno Vogt[b], Michael Groessl[b,**]

[a] Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA), Scuola Universitaria Professionale della Svizzera italiana (SUPSI), Università della Svizzera italiana (USI), CH-6928 Manno, Switzerland
[b] Department of Nephrology and Hypertension and Department of BioMedical Research, Inselspital, Bern University Hospital, University of Bern, Switzerland

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

DeepRel: Deep learning-based gas chromatographic retention index predictor

Tomáš Vrzal [a, *], Michaela Malečková [a, b], Jana Olšovská [a]

[a] Research Institute of Brewing and Malting, Plc., Lípová 511/15, 120 44, Prague 2, Czech Republic
[b] Charles University, Faculty of Science, Department of Analytical Chemistry, Albertov 6, 128 43, Prague 2, Czech Republic

Pre-trained: reuse previous data

analytical chemistry

pubs.acs.org/ac Article

Top-Down Approach to Retention Time Prediction in Comprehensive Two-Dimensional Gas Chromatography–Mass Spectrometry

Meriem Gaida,* Flavio A. Franchina, Pierre-Hugues Stefanuto, and Jean-François Focant

Cite This: *Anal. Chem.* 2022, 94, 17081–17089 Read Online

Journal of Chromatography A 1612 (2020) 460661

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Steroid identification via deep learning retention time predictions and two-dimensional gas chromatography-high resolution mass spectrometry

Giuseppe Marco Randazzo[a,*], Andrea Bileck[b], Andrea Danani[a], Bruno Vogt[b], Michael Groessl[b,**]

[a] Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA), Scuola Universitaria Professionale della Svizzera italiana (SUPSI), Università della Svizzera italiana (USI), CH-6928 Manno, Switzerland
[b] Department of Nephrology and Hypertension and Department of BioMedical Research, Inselspital, Bern University Hospital, University of Bern, Switzerland

RETURN TO ISSUE < PREV ECOTOXICOLOGY AND PU... NEXT >

Modeling the GCxGC Elution Patterns of a Hydrocarbon Structure Library To Innovate Environmental Risk Assessments of Petroleum Substances

J. Samuel Arey, Alberto Martin Aparicio, Eleni Vaiopoulou, Stuart Forbes, and Delina Lyon*

Cite this: *Environ. Sci. Technol.* 2022, 56, 24, 17913–17923
Publication Date: December 7, 2022
https://doi.org/10.1021/acs.est.2c06922

Copyright © 2022 The Authors. Published by American Chemical Society. This publication is licensed under CC-BY-NC-ND 4.0.
Open Access

Article Views 1879 | Altmetric 7 | Citations 1

LEARN ABOUT THESE METRICS

PDF (5 MB) Get it@ULiège SI Supporting Info (2) »

Contents lists available at ScienceDirect

Analytica Chimica Acta

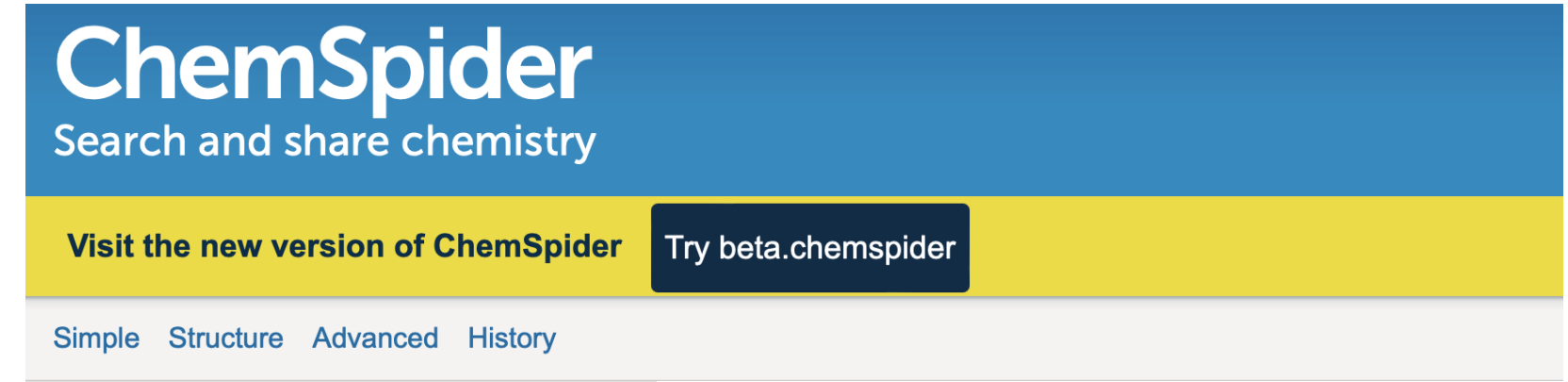
journal homepage: www.elsevier.com/locate/aca

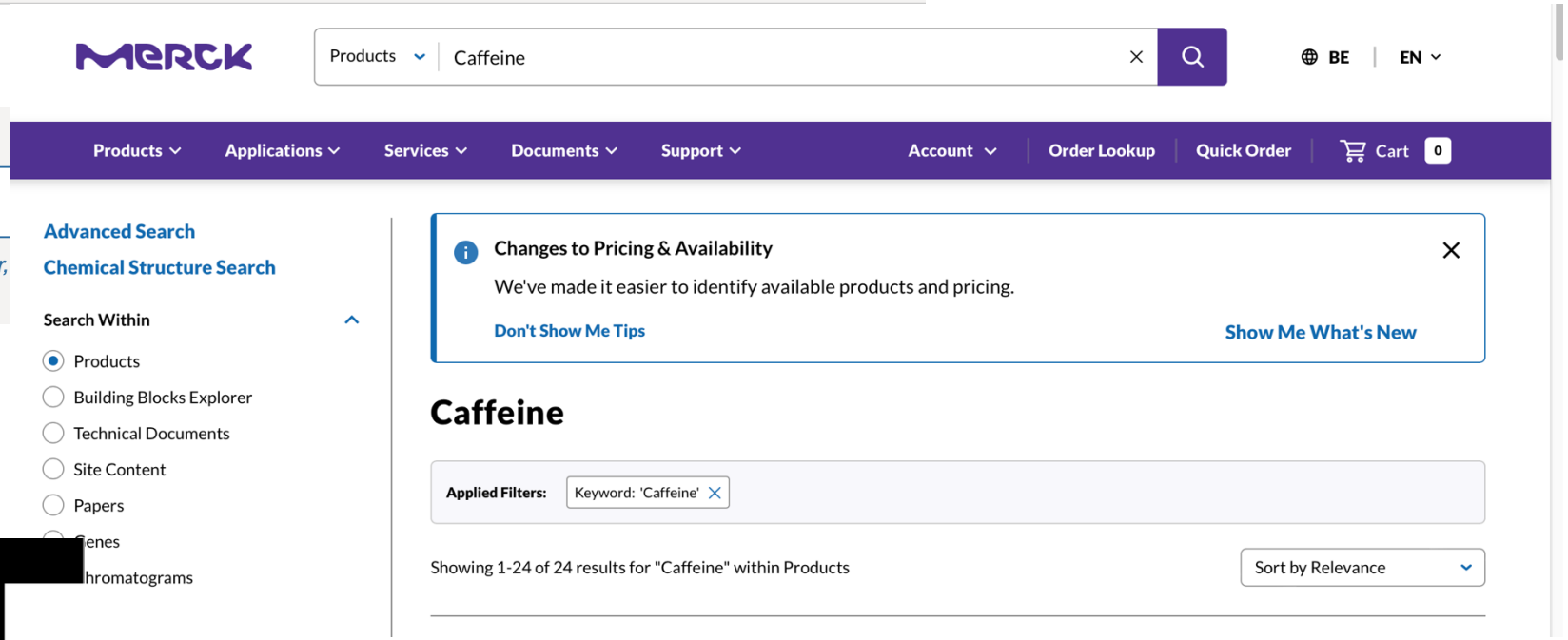
DeepRel: Deep learning-based gas chromatographic retention index predictor

Tomáš Vrzal [a, *], Michaela Malečková [a, b], Jana Olšovská [a]

[a] Research Institute of Brewing and Malting, Plc., Lípová 511/15, 120 44, Prague 2, Czech Republic
[b] Charles University, Faculty of Science, Department of Analytical Chemistry, Albertov 6, 128 43, Prague 2, Czech Republic

Pre-trained: we need large data!





Pre-trained: **generating data as a community**

ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Chromatography A, 1019 (2003) 261–272

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Standardized test mixture for the characterization of comprehensive two-dimensional gas chromatography columns: the Phillips mix

Jean-Marie D. Dimandja [a,*], Garrick C. Clouden [b], Ivelisse Colón [c], Jean-François Focant [d], Whitney V. Cabey [a], Ritchard C. Parry [e]

[a] *Department of Chemistry, Spelman College, 350 Spelman Lane, SW Box 279 Atlanta, GA 30314, USA*
[b] *Meharry Medical College, 1005 Dr. D. B. Todd Jr. Blvd., Nashville, TN 37208, USA*
[c] *Pfizer Global Research and Development, Eastern Point Road, Groton, CT 06340, USA*
[d] *Centers for Disease Control and Prevention, 4770 Buford Highway NE, Atlanta, GA 30041, USA*
[e] *LECO Corporation, 3000 Lakeview Avenue, St. Joseph, MI 49085, USA*

NIST FDA

Pre-trained: **generating data as a community**

1 – Column classification

- Century mix analysis across **20 Column combinations**

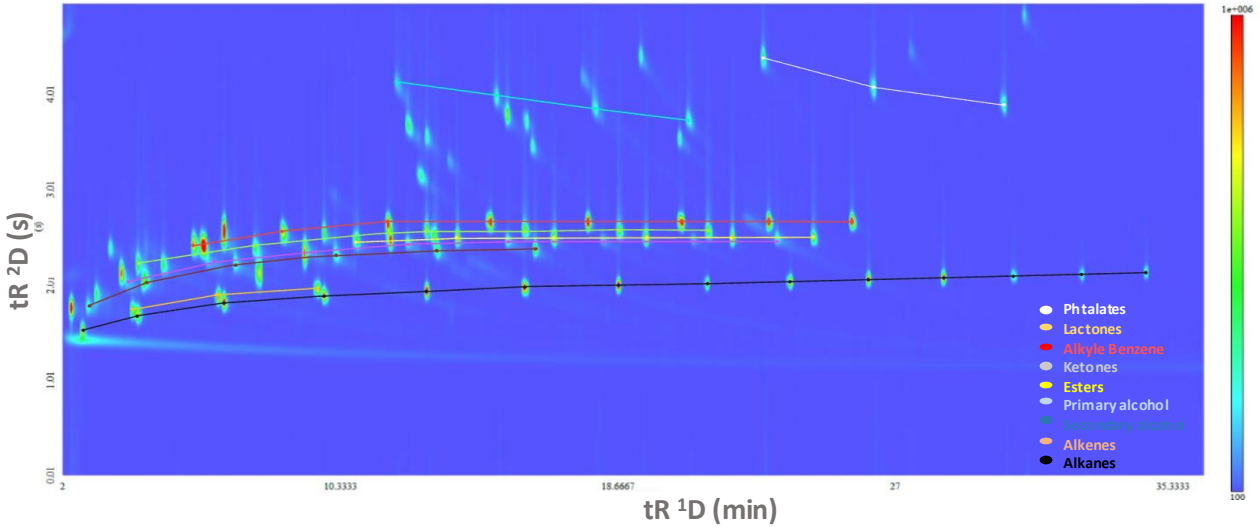
Polarity	Apolar	Mid-Polar	Polar
Apolar			
Mid-Polar			
Polar			

Normal orthogonality

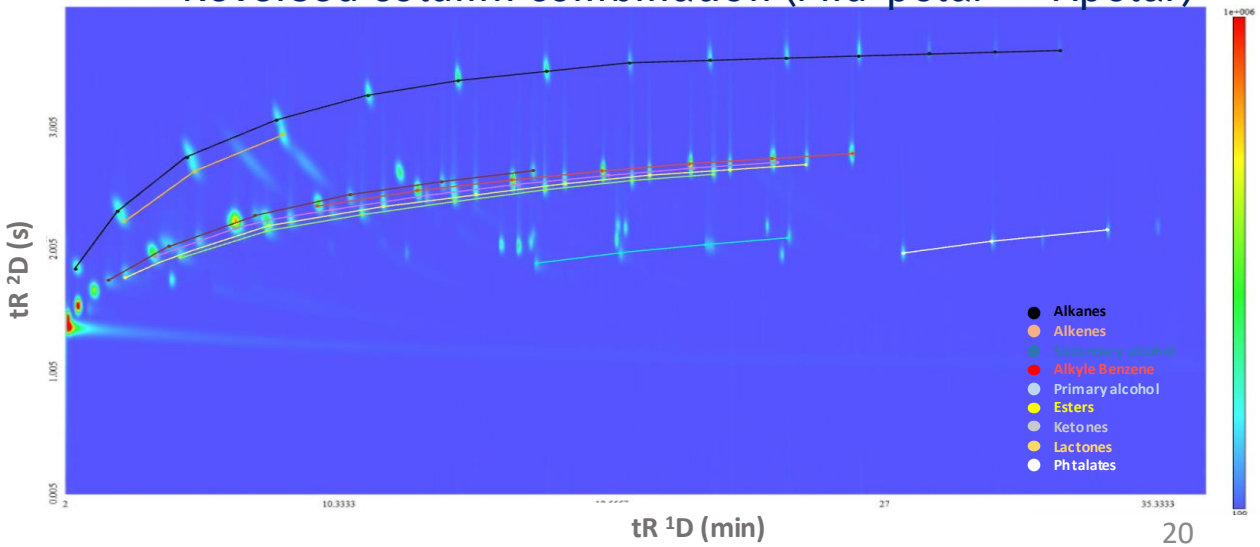
Non-orthogonal

Reversed orthogonality

Normal column combination (Apolar × Mid-polar)



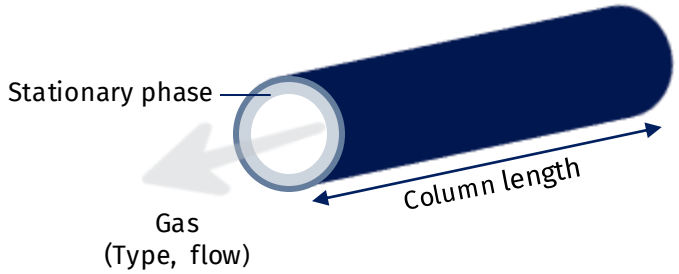
Reversed column combination (Mid-polar × Apolar)



Pre-trained: **generating data as a community**

1 – System evaluation

- Stationary phase film thickness (0.1 à 1.4 µm)
- Temperature ramp (2 à 15 °C/min)
- Gas type (He, H_2, N_2)
- Gas flow (0.8 à 1.5 mL/min)

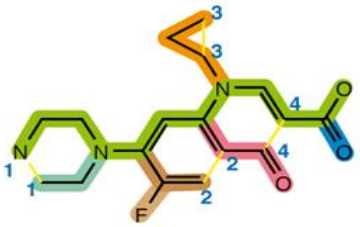


2 – Data Collection

Retention time , elution order , geometric peak repartition Etc

+

Boiling points, polarity moment, **SMILES Annotation**



Ciprofloxacin SMILES Annotation

N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O

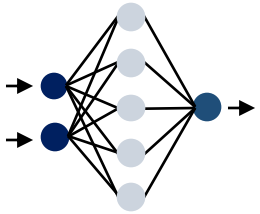
3 – Modelisation-Prediction

Retention indices prediction

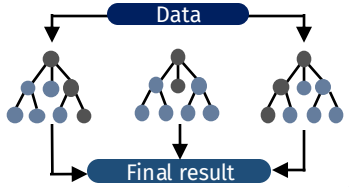
(System-independent constants)

Machine Learning Algorithm

Artificial neural network



Random Forest



Transformer: mostly neural network

Define neural network in one sentence

Text generated from ChatGPT

Deep Learning, a step further for transformers

Machine Learning	Deep Learning
Possible to train with fewer data	Large datasets for training
Statistical algorithms	Artificial Neural Networks (ANN)
Structured data	Unstructured data
Limited tuning capabilities	Can be tuned in multiple ways
Simpler applications	More complex applications

Automation

New methods

Greener methods

Images generated on ideogram.ai

16TH Multidimensional Chromatography Workshop

LIÈGE, BELGIUM

FEBRUARY 3-5, 2025

FREE registration

PRESENTATIONS by key speakers

DISCUSSION on hot topics

POSTER SESSION with awards

www.multidimensionalchromatography.com

Liège University

Think tank

Networking hub