HS-GC-IMS: analytical innovation applied to the recent experiences of the Coop laboratory

Pietro Morozzi
Analista Dati Qualità e Monitoraggio Performance del Sistema | AI Specialist
Direzione Qualità e Sicurezza Prodotto – Quality Assurance Team
e-mail: pietro.morozzi@coopitalia.coop.it

Coop Italia

- Italian Large-scale Organised Distribution (**GDO**)
- **72 Associated Cooperatives**, **18 Italian Regions**
- Annual revenue of around **€16.6 billion**
- More than **6 million members**
- Around **2260 shops** and **58000 employers**

Values of COOP products:

Safety *Ethical* *Cheap*

Tasty *Ecological* *Trasparent*

Biology

Chemistry

Sensory

Food Integrity Lab, Casalecchio di Reno (BO)

Food Integrity Lab – Chemistry Area

Gradual evolution of the Laboratory since 80’s and today towards **rapid**, **sustainable**, **semi-automated**, and **untargeted analyses** for quality control (not only food)

Analysis of volatile compounds

FlavourSpec®**25**
(HS-GC – IMS)

Heracles II
(HS-GC – FID)

Spectroscopic analyses

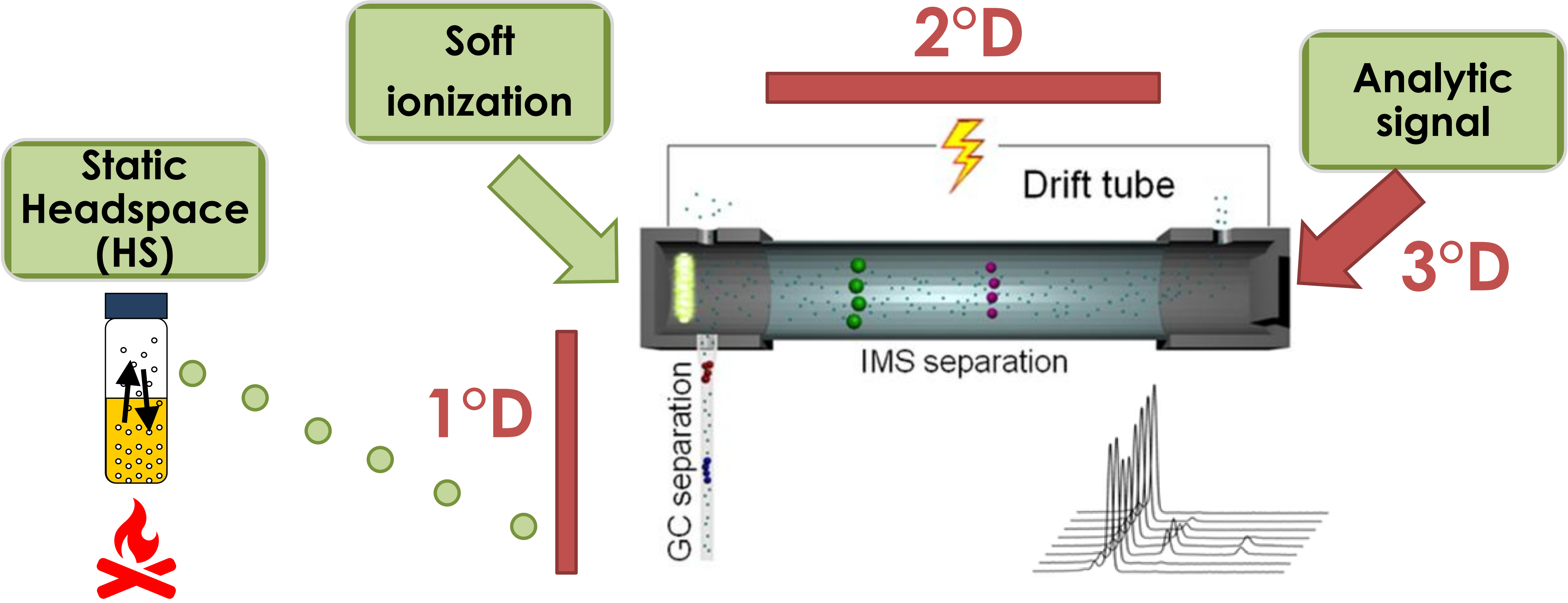
Multiskan GO
(Transmittance spectrophotometer)

DS-710
(Diffuse reflectance colorimeter)

MicroNIR OnSite-W
(NIR spectrometer)

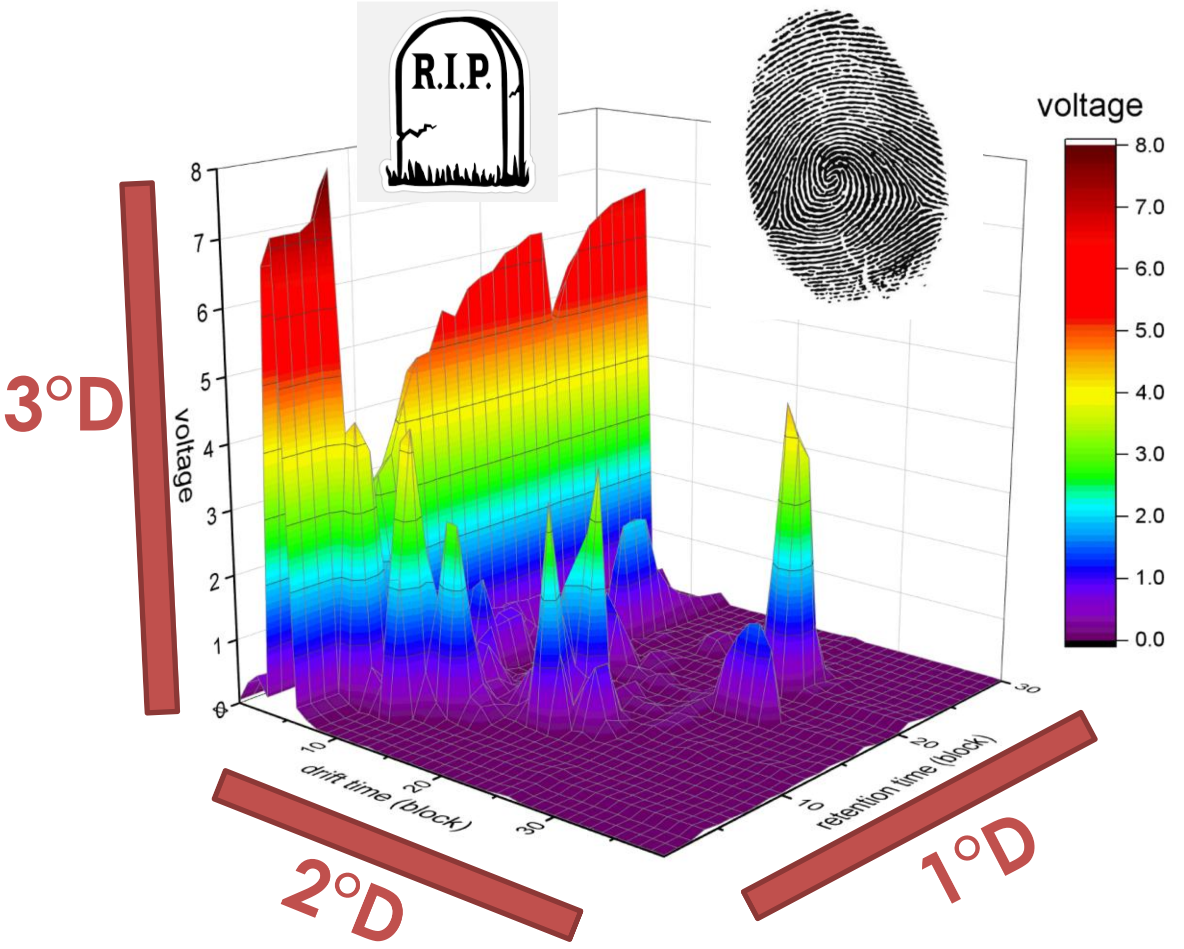
FlavourSpec®: the technology in 3 steps

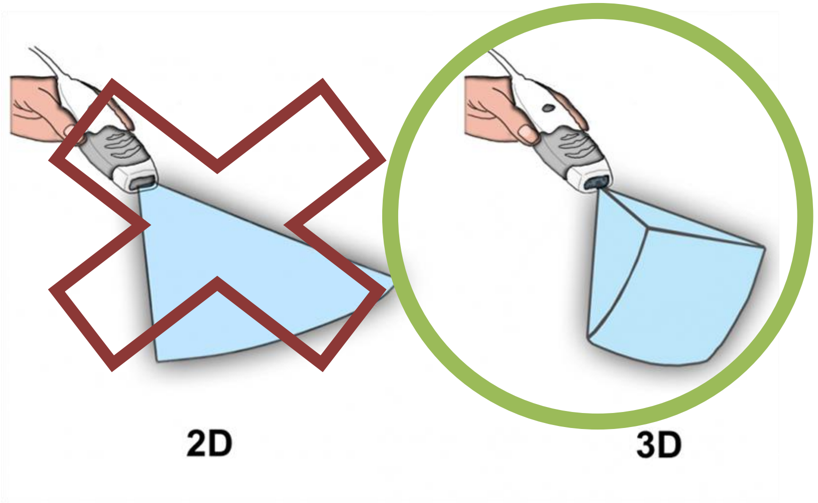
- Gas Chromatography (**GC**) coupled with ion mobility spectrometry (**IMS**)
- **<u>Hyphenated</u>** technique, that generates **3D data**



FlavourSpec®: raster data

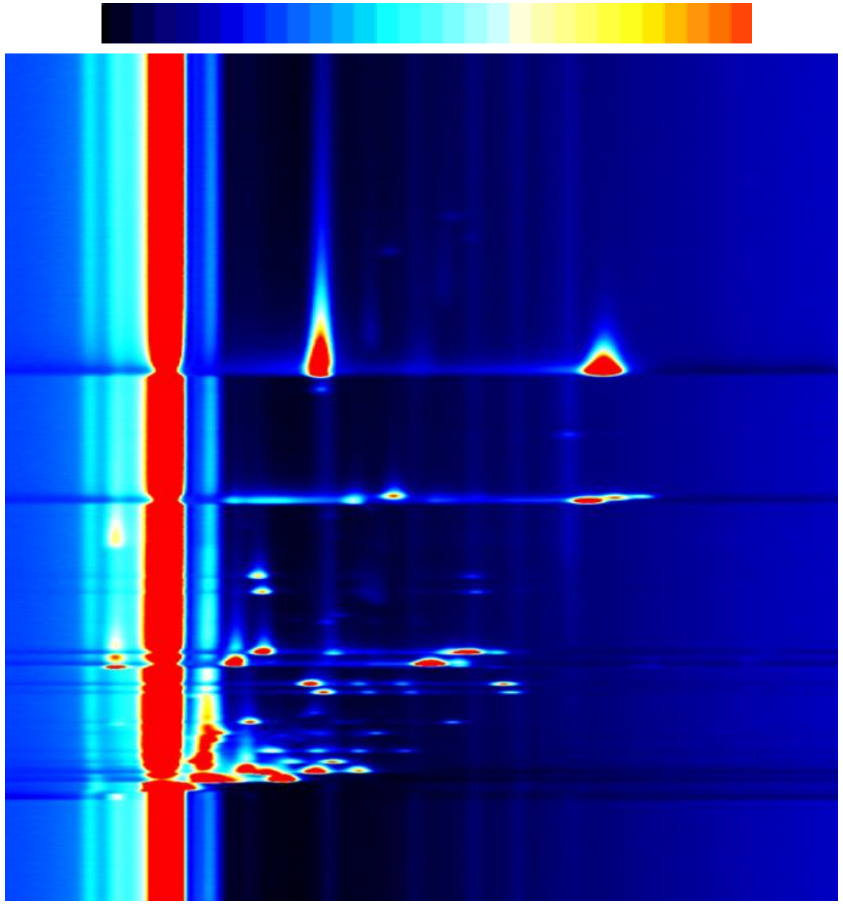
Data matrix with **topological information**:
3D topographic map



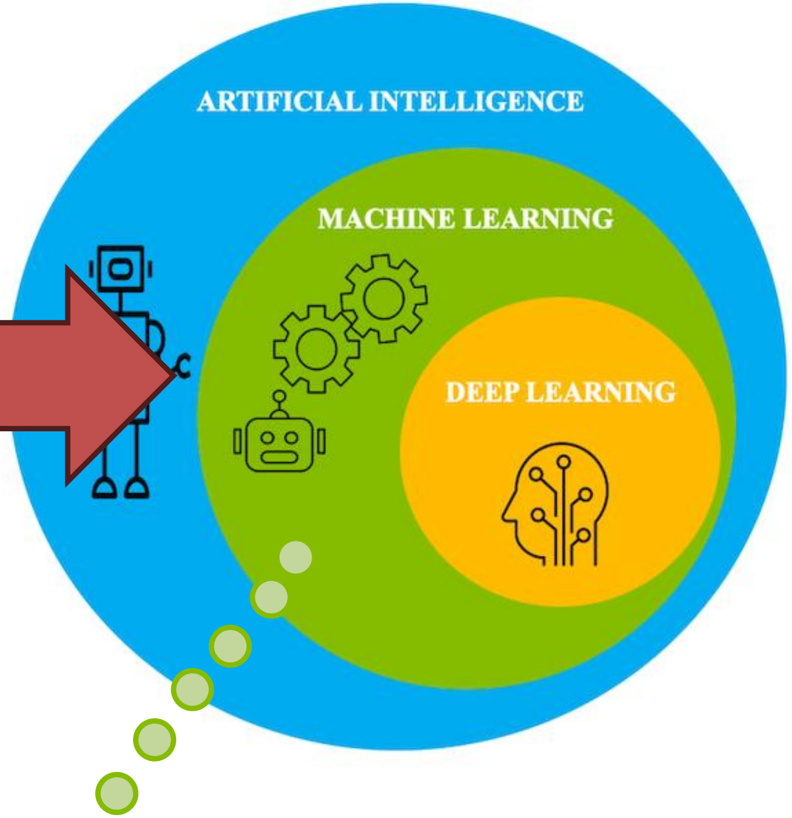


Mathematically equal to **geospatial data**:

FlavourSpec®: from "raster" to Machine Learning



2D visual representations of 3D information (retention time × drift time × voltage)



Subset of AI techniques that automatically enables a machine or system to learn and improve from experience

FlavourSpec®: Data approach

Macro: **general principle**

Meso: **learning**

Micro: **specific goal**

Top-down and **untargeted approach**

Secondary technique

Decision-making value only when **calibrated/trained on a primary technique or pre-defined classes**.

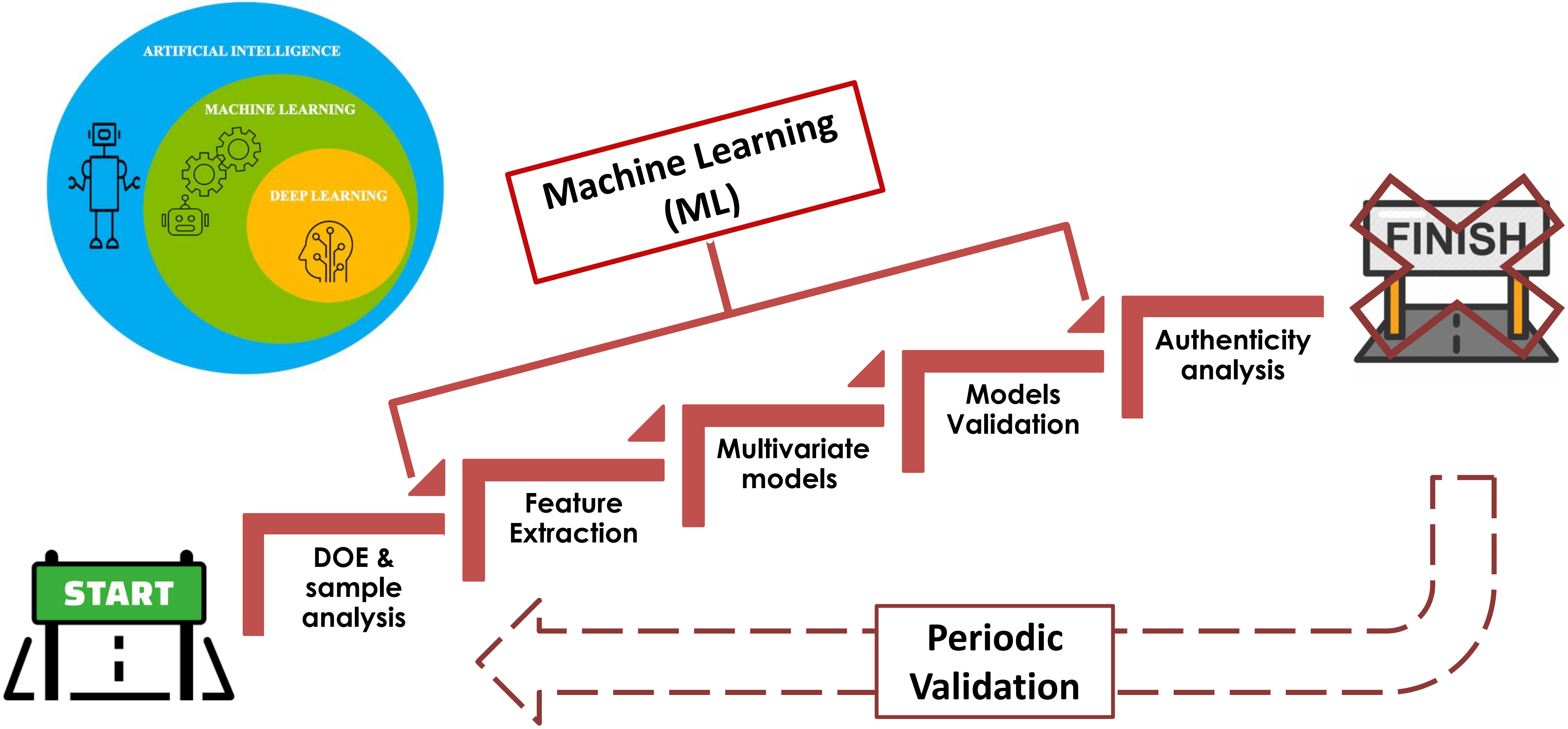
Supervised models

Output obtained from samples labeled by a primary reference → enables **operational screening from VOCs fingerprinting**.

Anomaly detection

Detects fingerprint alterations → **investigate with primary/targeted techniques**.

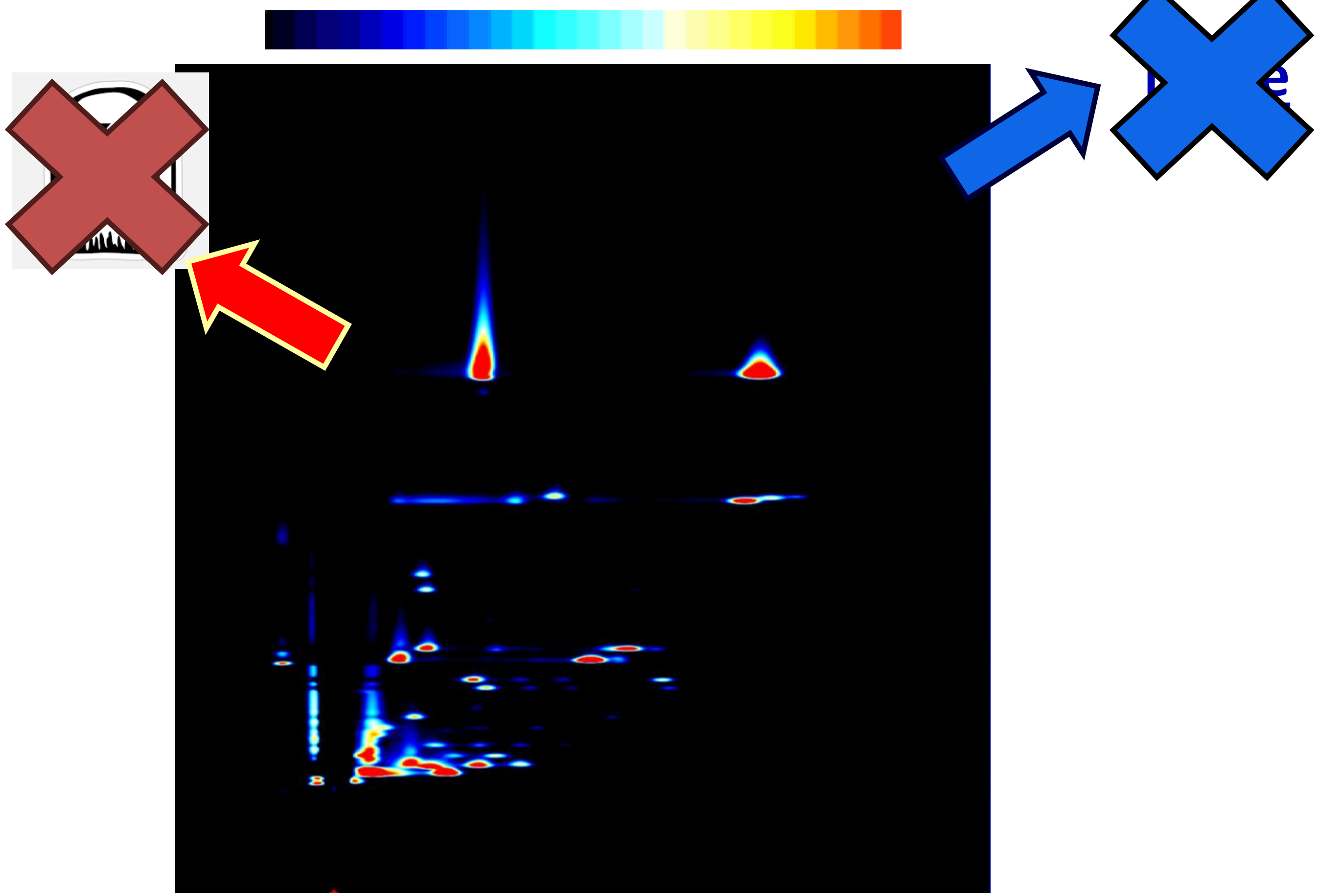
FlavourSpec®: Data approach



FlavourSpec®: Data Preparation for ML

- **<u>Data Cleaning</u>** and **<u>Feature Selection</u>** by G.A.S. Dortmund software (VOCal) + deNoise …

VOCal 0.4.33

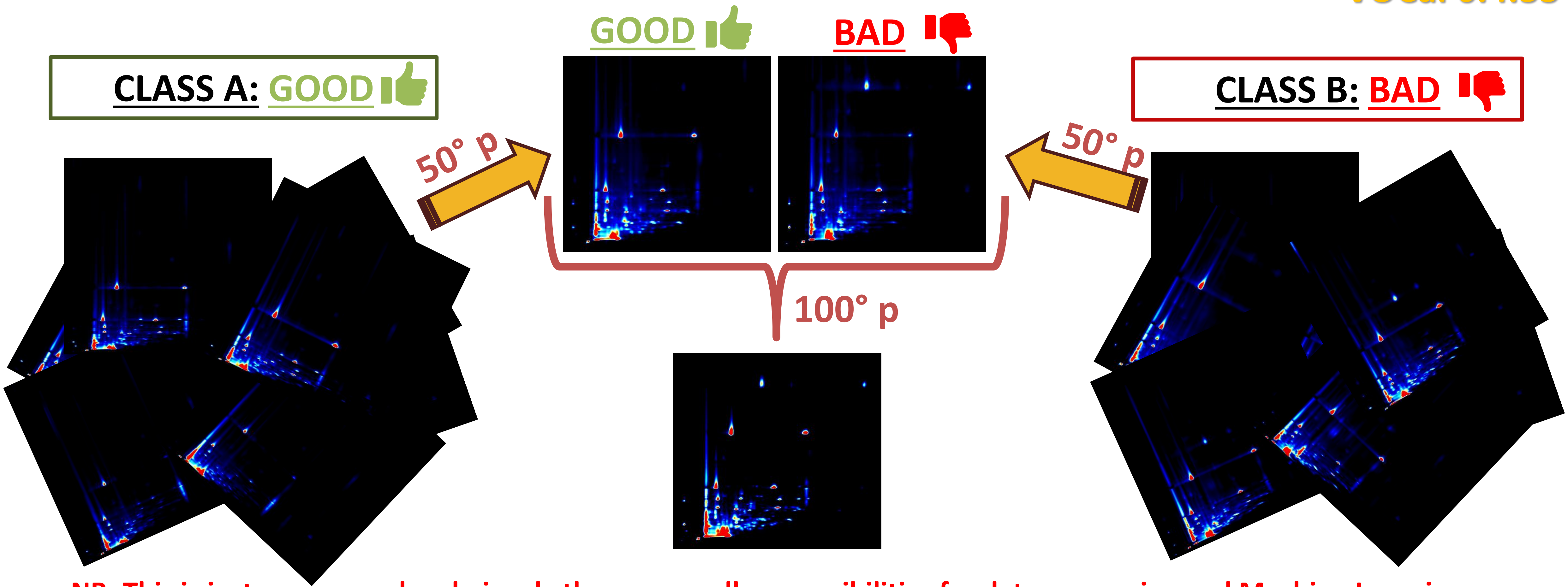


NB: This is just one example, obviously there are endless possibilities for data processing and Machine Learning …

FlavourSpec®: Data Preparation for ML

- **Data Cleaning** and **Feature Selection** by G.A.S. Dortmund software (VOCal) + deNoise & Class Percentile Modules

VOCal 0.4.33

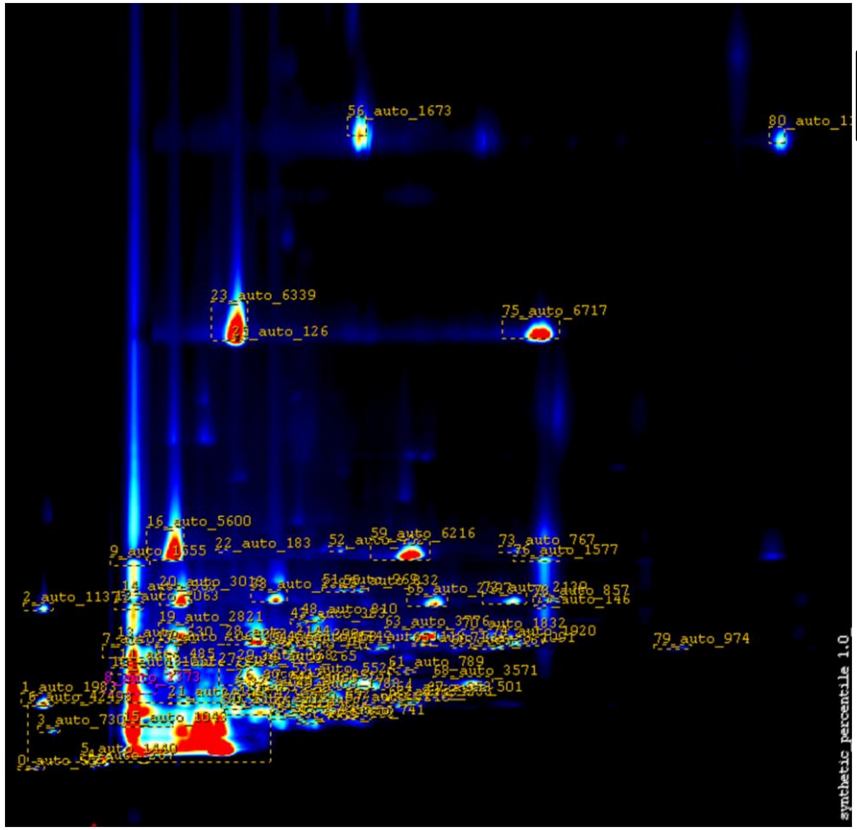


NB: This is just one example, obviously there are endless possibilities for data processing and Machine Learning …

FlavourSpec®: Data Preparation for ML

- **Data Cleaning** and **Feature Selection** by G.A.S. Dortmund software (VOCal) + AutoArea tool

VOCal 0.4.33



N samples x M areas

Sample	A priori Class	Area_1	Area_2	Area_...	Area_M
Sample_1	A	0.321	10.08	3.485	0.004
Sample_2	A	0.400	11.30	3.595	0.005
Sample_...	B	1.795	1.985	3.176	0.011
Sample_N	B	1.832	1.735	3.202	0.009

NB: This is just one example, obviously there are endless possibilities for data processing and Machine Learning …

Machine Learning: algorithms

Supervised Learning

Classification

- PLS-DA
- DTs
- Logit
- KNN
- SVM
- Naive Bayes
- RF
- AdaBoost

Regression

- Linear
- Polynomial
- SVR
- DTs
- Lasso
- Ridge

Unsupervised Learning

Generalization

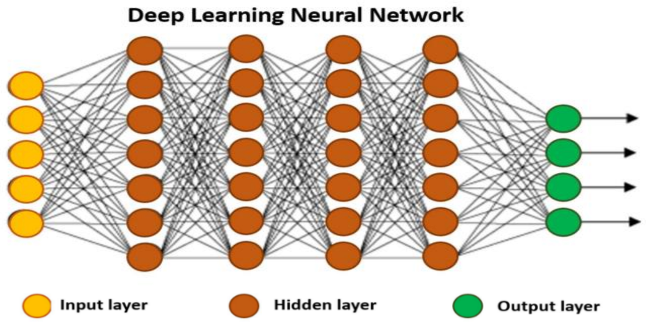
- PCA
- kPCA
- MDS
- TSNE
- LDA
- PMF

Clustering

- K-means
- DBScan
- Hierarchical
- Mean shift
- FCM
- GMM

Deep Learning

- ANN
- CNN
- RNN



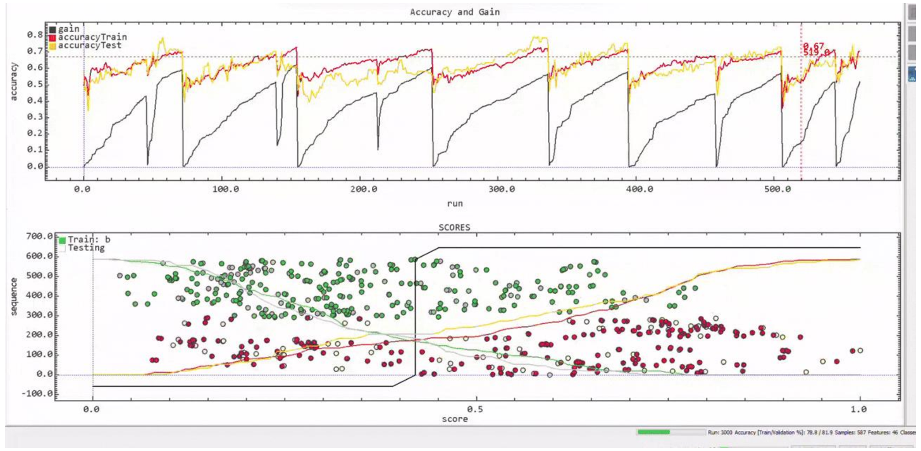
There are many other algorithms...

FlavourSpec®: Machine Learning

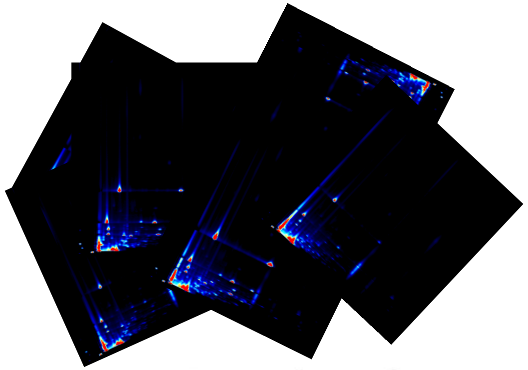
- **Multivariate Models** by G.A.S. Dortmund software (VOCal) + Dynamic PCA, Solver Module, MEA2Images & more…

VOCal 0.4.33

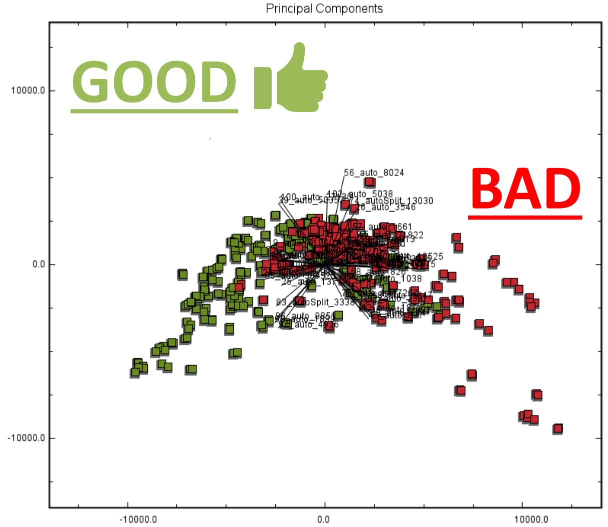
GitHub - dsandersGit/GIT_Solver

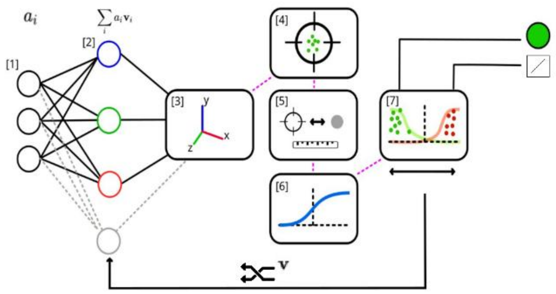


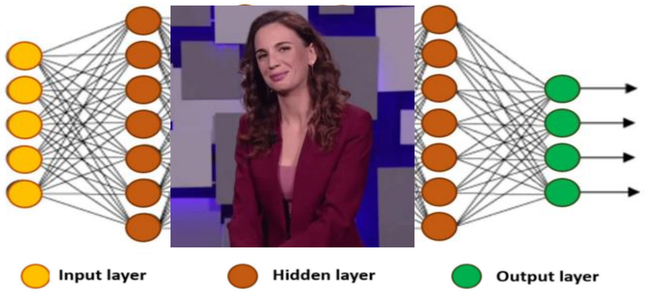
CNN – images (.png)



PCA



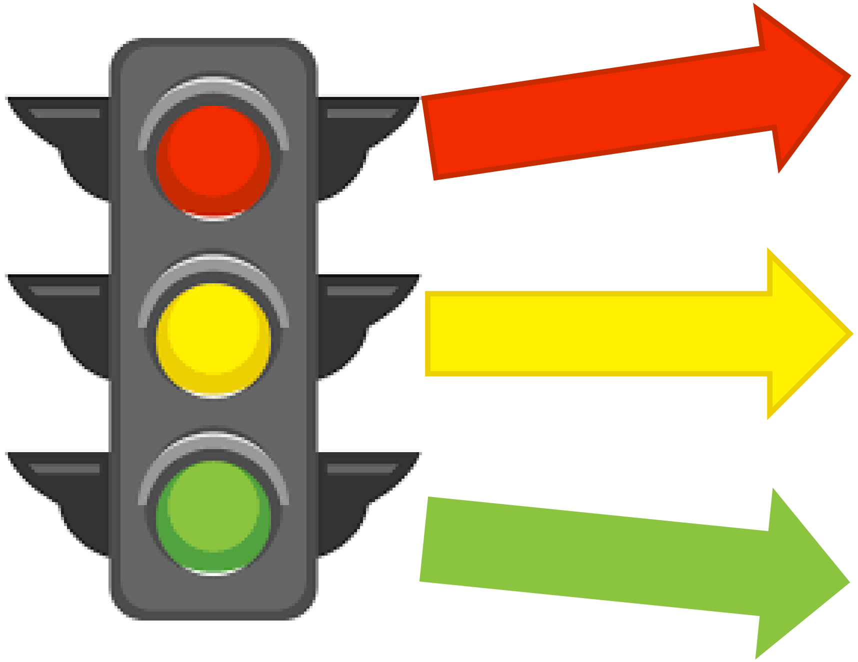




NB: This is just one example, obviously there are endless possibilities for data processing and Machine Learning …

COOP Italia: results

- **<u>Currently</u>**, these methods are not recognized by certification agencies.
  - **Internal** procedures, which return **alerts** to verify with other methods.



Not-compliant

Suspected

Compliant

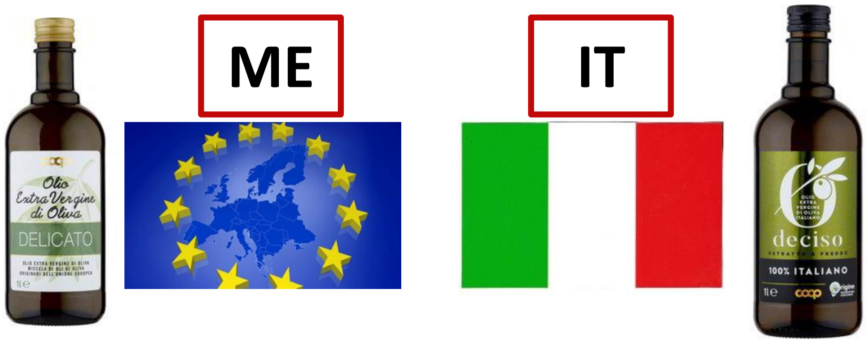
We know nothing very certainly,
but everything only probably,
and the probability has degrees
that are widely different.
Christiaan Huygens (1629–1695)

COOP case studies

Both short projects and routine activities that have consolidated over the years and are based on **<u>risk analysis</u>**, mainly in the food sector.

Olive oil

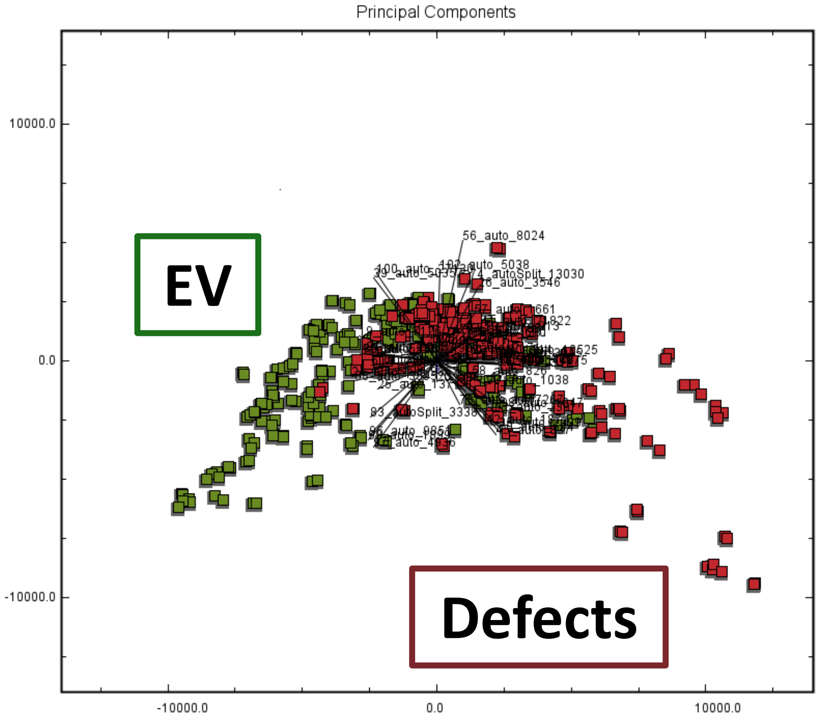
- Designation of origin



- Merceological classification (**extra-virgin** vs. **sensory defects**)

- Total and/or partial replacement with seed oil

Olive oil



Pre-2025 ➔ Post-2025

FlavourSpec®**25**
(HS-GC – IMS)

Train set (80%)

	EV	defects
EV	7%	93%
defects	98%	2%

Sensibility	Specificity	Precision	Accuracy	F1
0.93	0.98	0.97	0.96	0.95

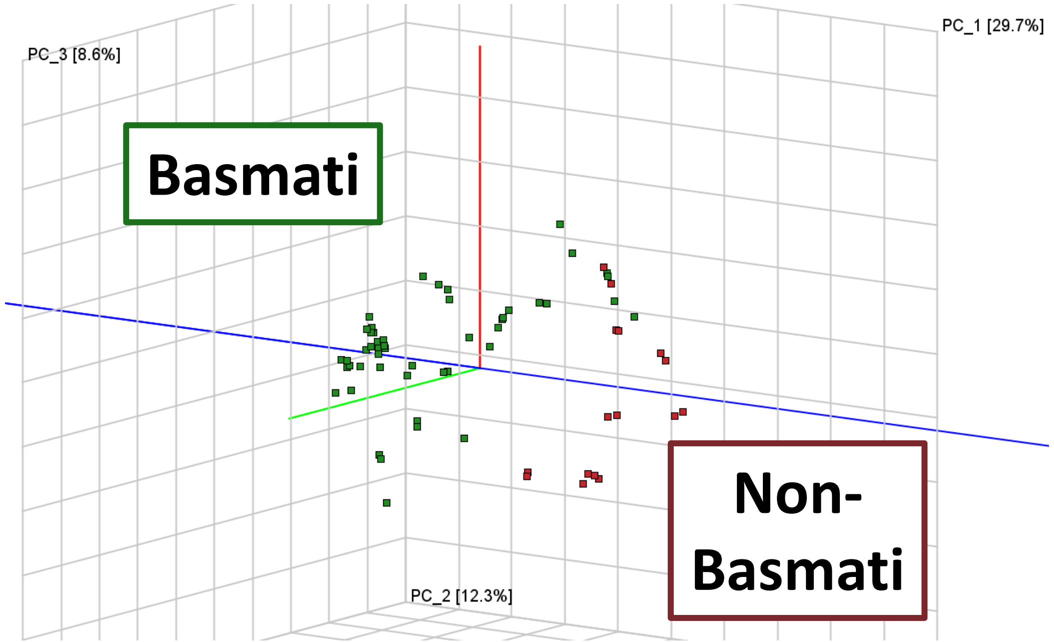
Test set (20%)

	EV	defects
EV	7%	93%
defects	94%	6%

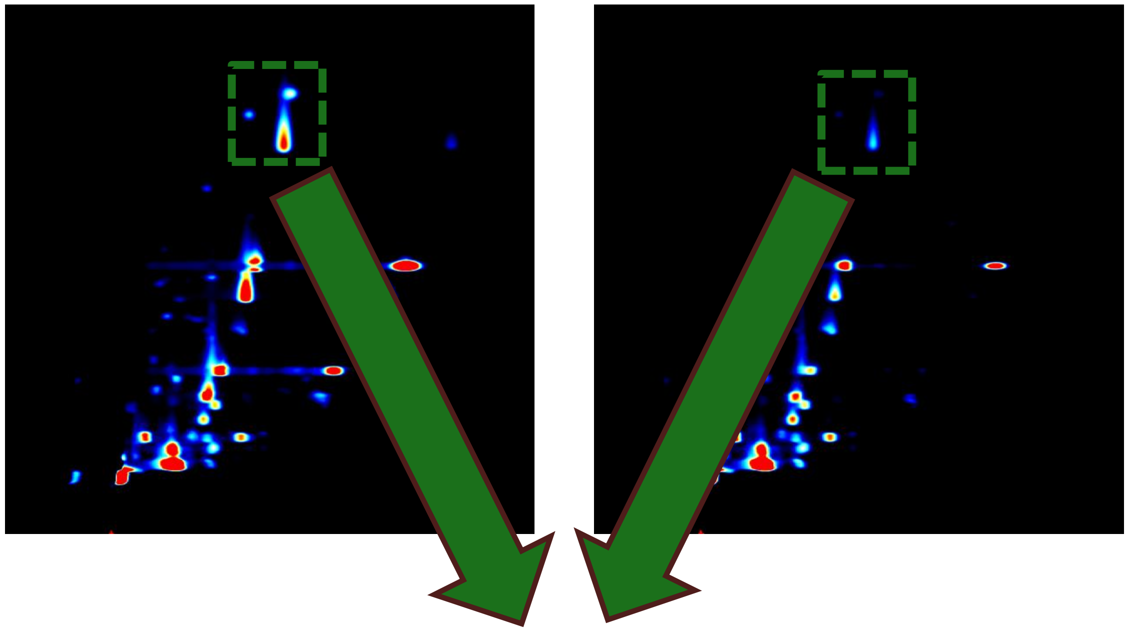
Sensibility	Specificity	Precision	Accuracy	F1
0.93	0.94	0.93	0.93	0.93

Rice: aromatic vs non-aromatic cultivars

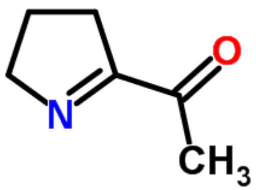
Principal Components Analysis (PCA)

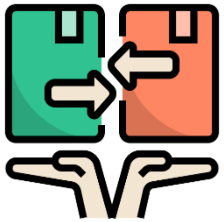


50° p Basmati | **50° p Non-Basmati**



2-Acetyl-1-pyrroline (CAS-85213-22-5)





Digital PCR (dPCR) enables absolute quantification of non-Basmati rice (%) by comparing copy numbers of a **BADH2-specific marker (non-Basmati)** and a rice reference gene, using Poisson statistics.

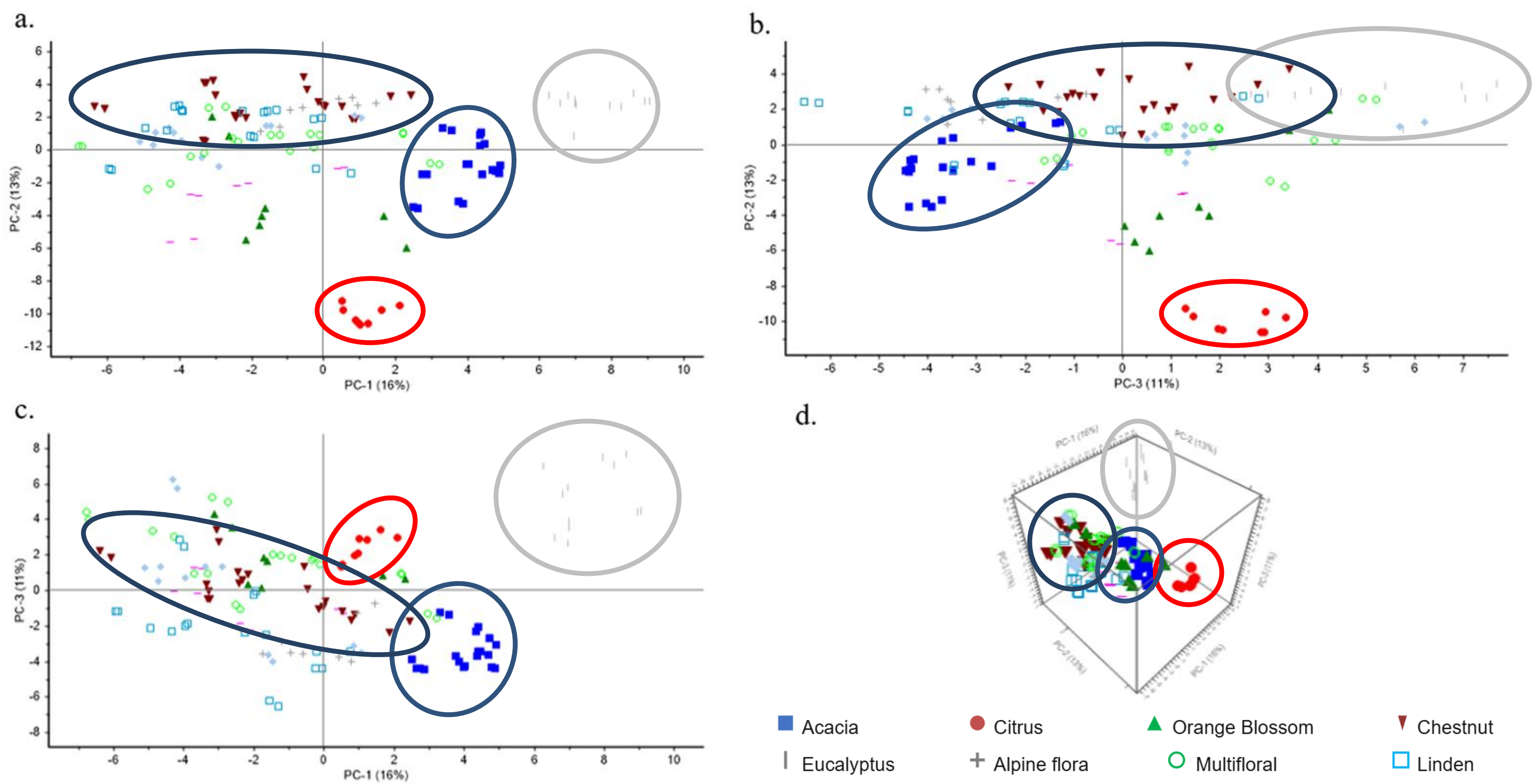
Honey: authenticity

- Botanical origin authentication

multiflorar

monoflorar

Honey: authenticity



POSTER (2025): [COOP_UNIBO: Honey Authenticity]

Food Integrity Lab: Why FlavourSpec®

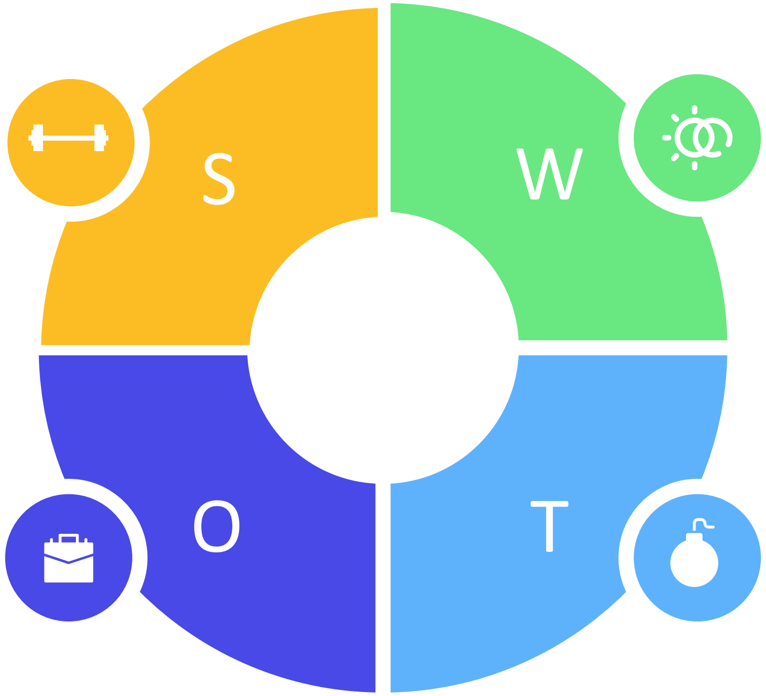
Minimal or no sample preparation: reduced preparative steps up to direct analysis.

Shortens decision time: pre-filters suspects before confirmatory tests (panel and/or targeted).

Scalability: throughput and sustainability suited to widespread controls.

Integrates, doesn't replace existing instruments ➔ **Multiblock approach**

SWOT Analysis: 2022➔2026



Strengths

- Quick
- Applicability
- Sustainable
- Sensibility

Weaknesses

- Expertise
- Sensibility (noise)
- Need of improvements

Opportunities

- Progress
- Willingness to improve

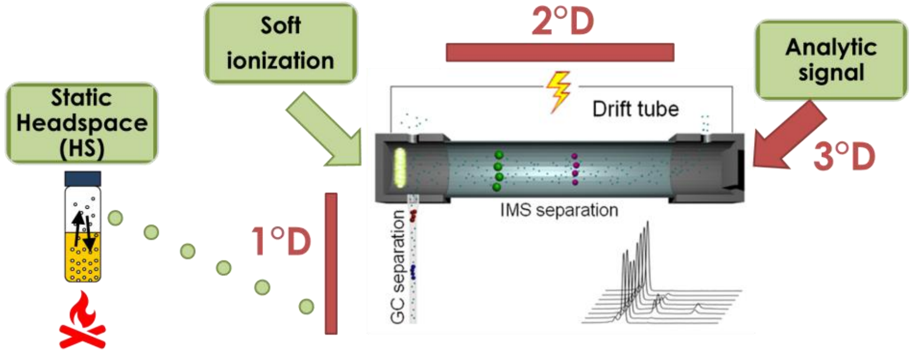
Threats

- Competition
- No accreditation

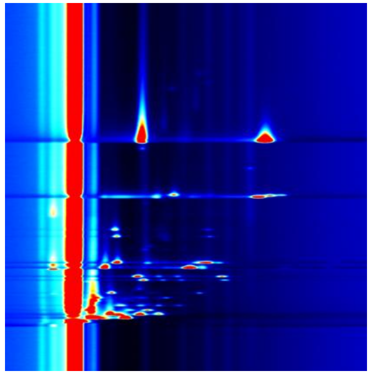
<u>Internal procedures, not accredited</u>, but validated and used by Universities, Authorities, Private companies, Distributor, etc…

Keyword: harmonization
(for each application and/or food matrix)

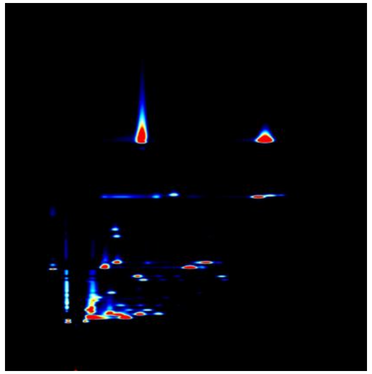
- HS preparation & GC-IMS parameters



- Raw data quality validation & Proficiency Testing (PT)



- Data Preparation for ML



- Shareable and accessible databases

- Shareable and accessible ML models

Thanks for your attention

Pietro Morozzi
Analista Dati Qualità e Monitoraggio Performance del Sistema | AI Specialist
Direzione Qualità e Sicurezza Prodotto – Quality Assurance Team
e-mail: pietro.morozzi@coopitalia.coop.it