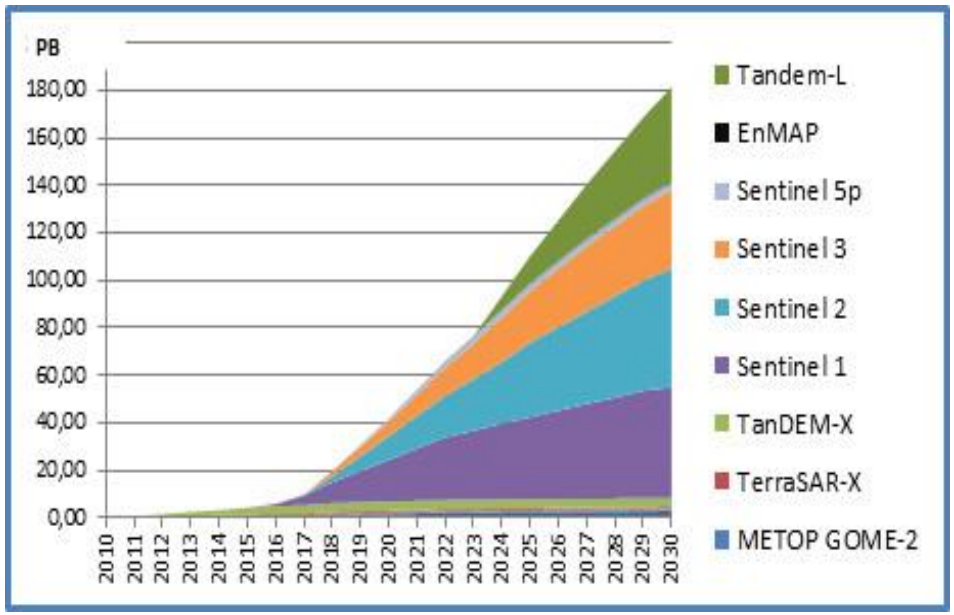
Selection of Reliable Machine Learning Algorithms for Geophysical Applications

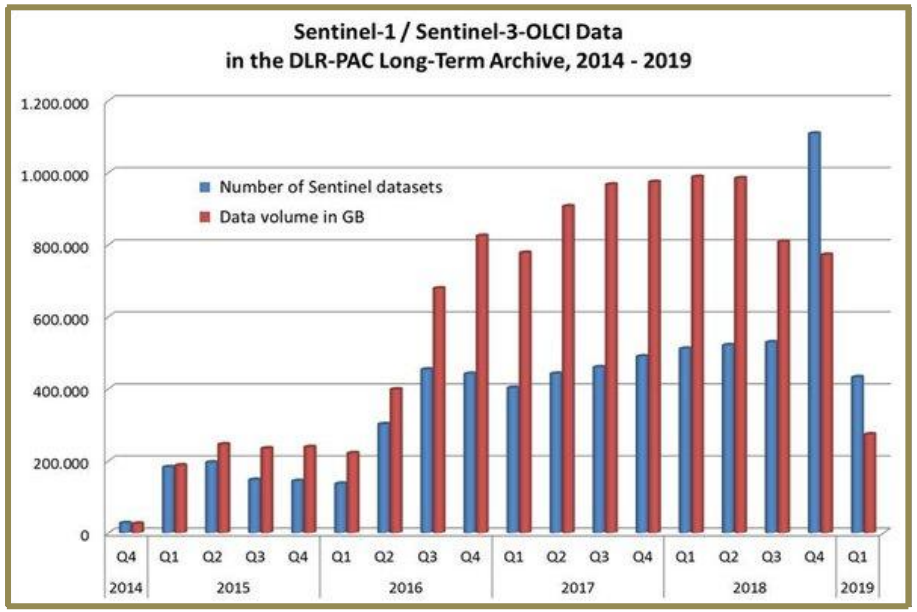
C.O. Dumitru, G. Schwarz, D. Ao, G. Dax, V. Andrei, C. Karmakar, and M. Datcu

DLR

Big Data & Sensors

The volume of data for different sensors over the years





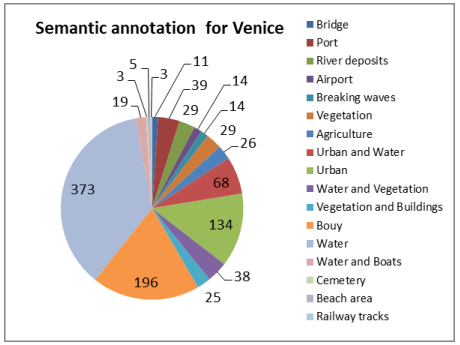
https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-12632/22039_read-51751

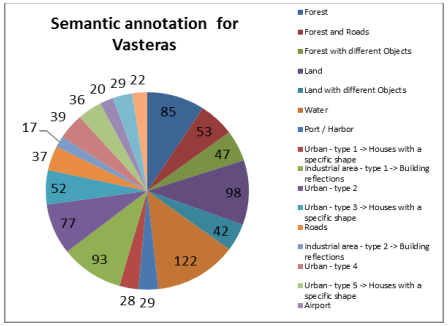
https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-13247/23165_read-54030

DLR

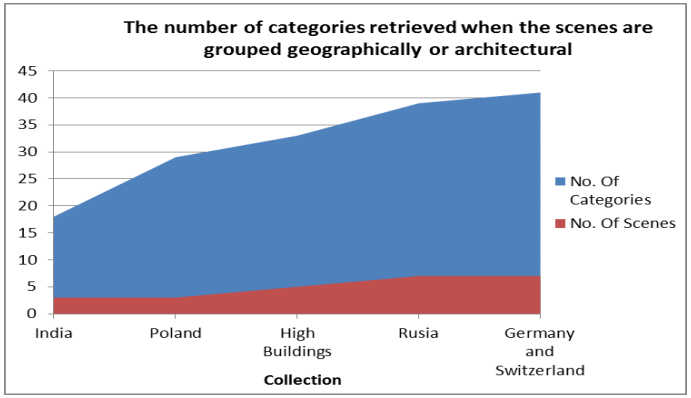
Applications: TerraSAR-X archive semantic catalogue

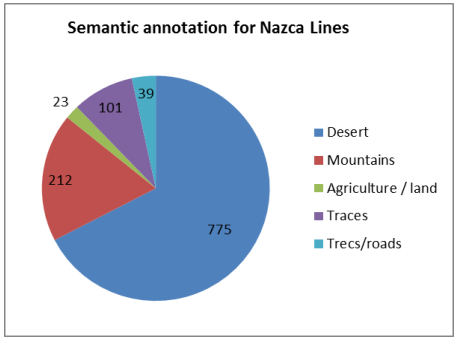
300 cities 1300 semantic labels





-Bangkok (Thailand);
-Shenyang (China);
-Nazca Lines (Peru);
-Havana (Cuba);
-Venice (Italy);
-Vasteras (Sweden);
-Oran (Algeria);
-Bogota (Columbia).

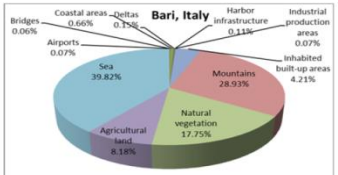


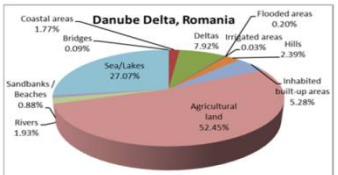


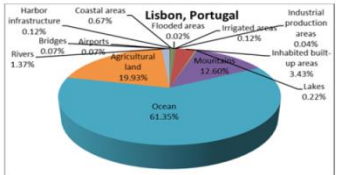
DLR

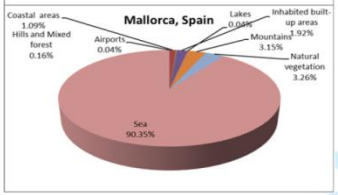
Applications: Protected areas in Europe - ECOPOTENTIAL

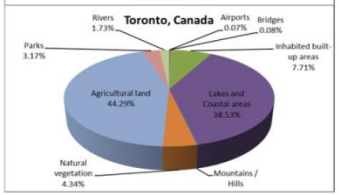
General description of Ecopotential protected areas

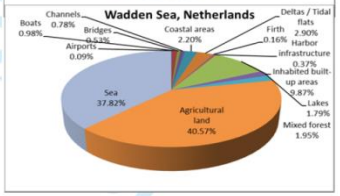












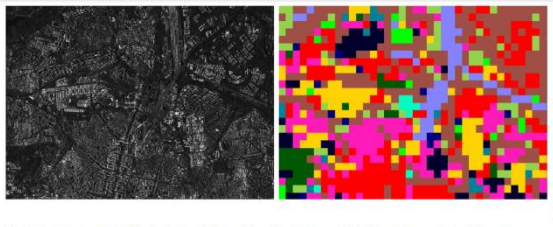


Fig. 21.b. TerraSAR-X quick-look view (left) and classification map (right) for an image of Madrid, Spain.

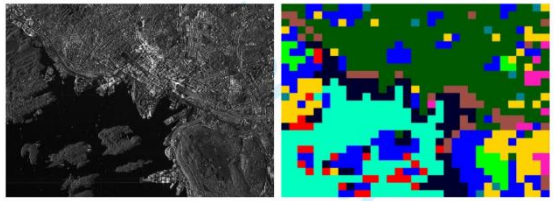
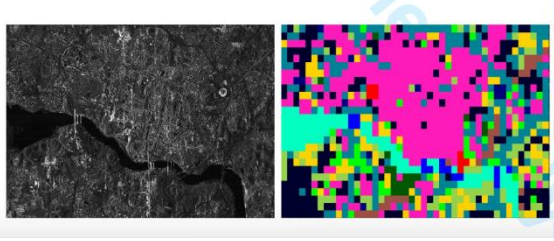
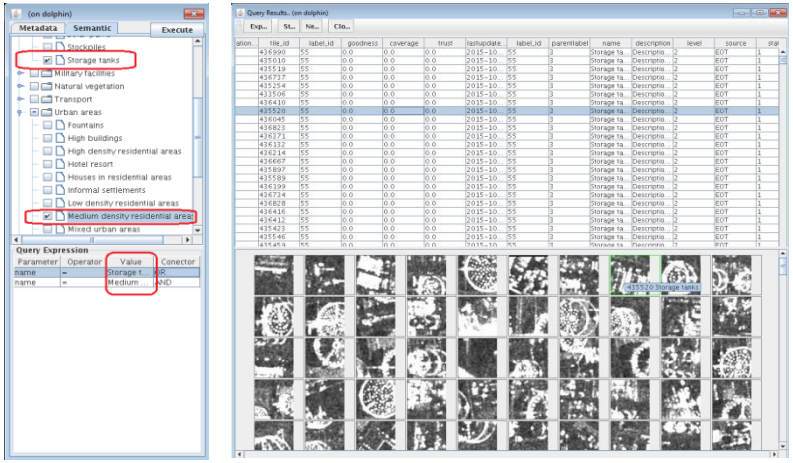


Fig. 21.c. TerraSAR-X quick-look view (left) and classification map (right) for an image of Oslo, Norway.

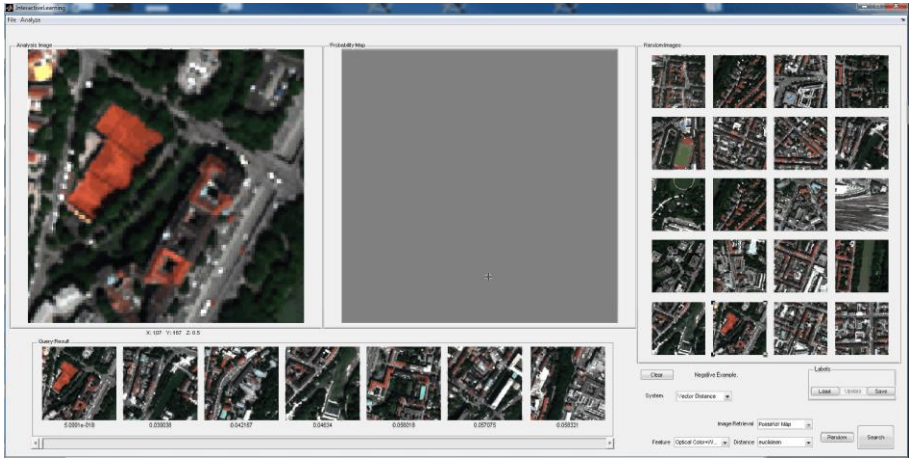


Data Mining Tools

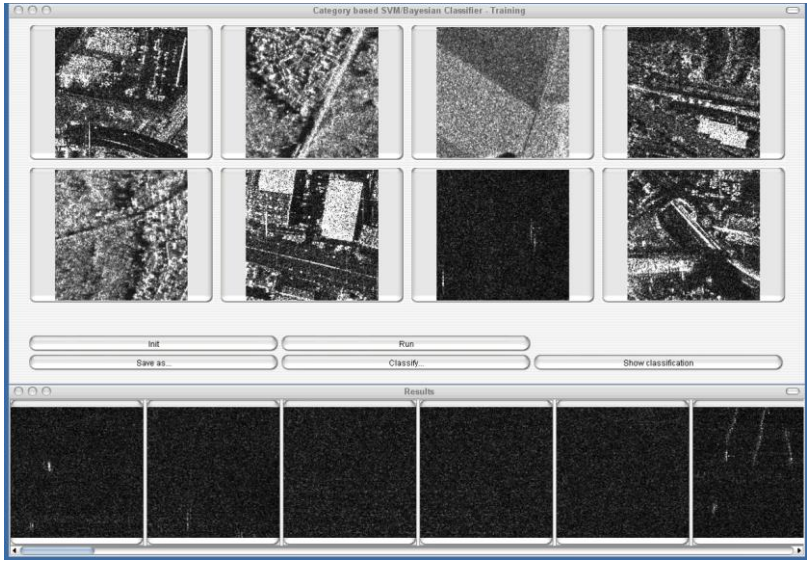
Multimodal queries



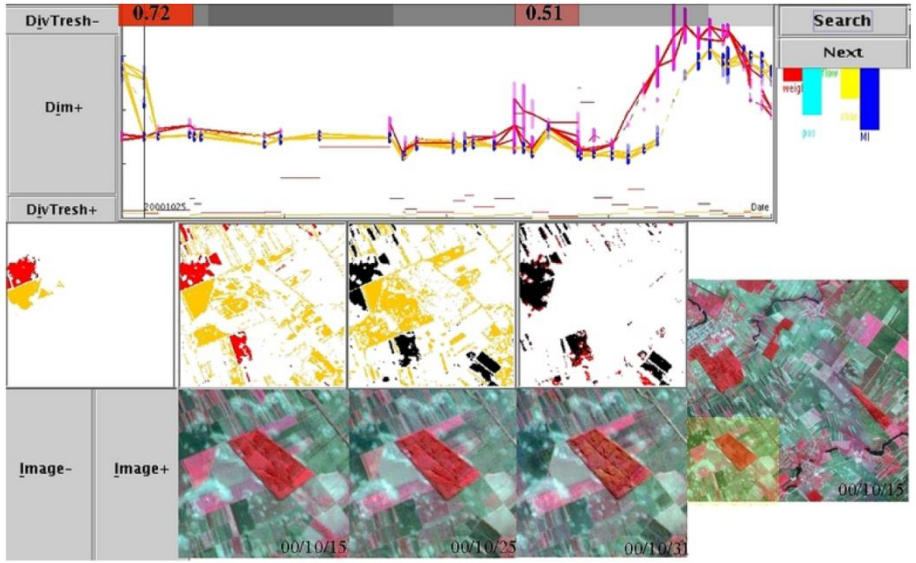
Query by example



KDD and semantic extraction

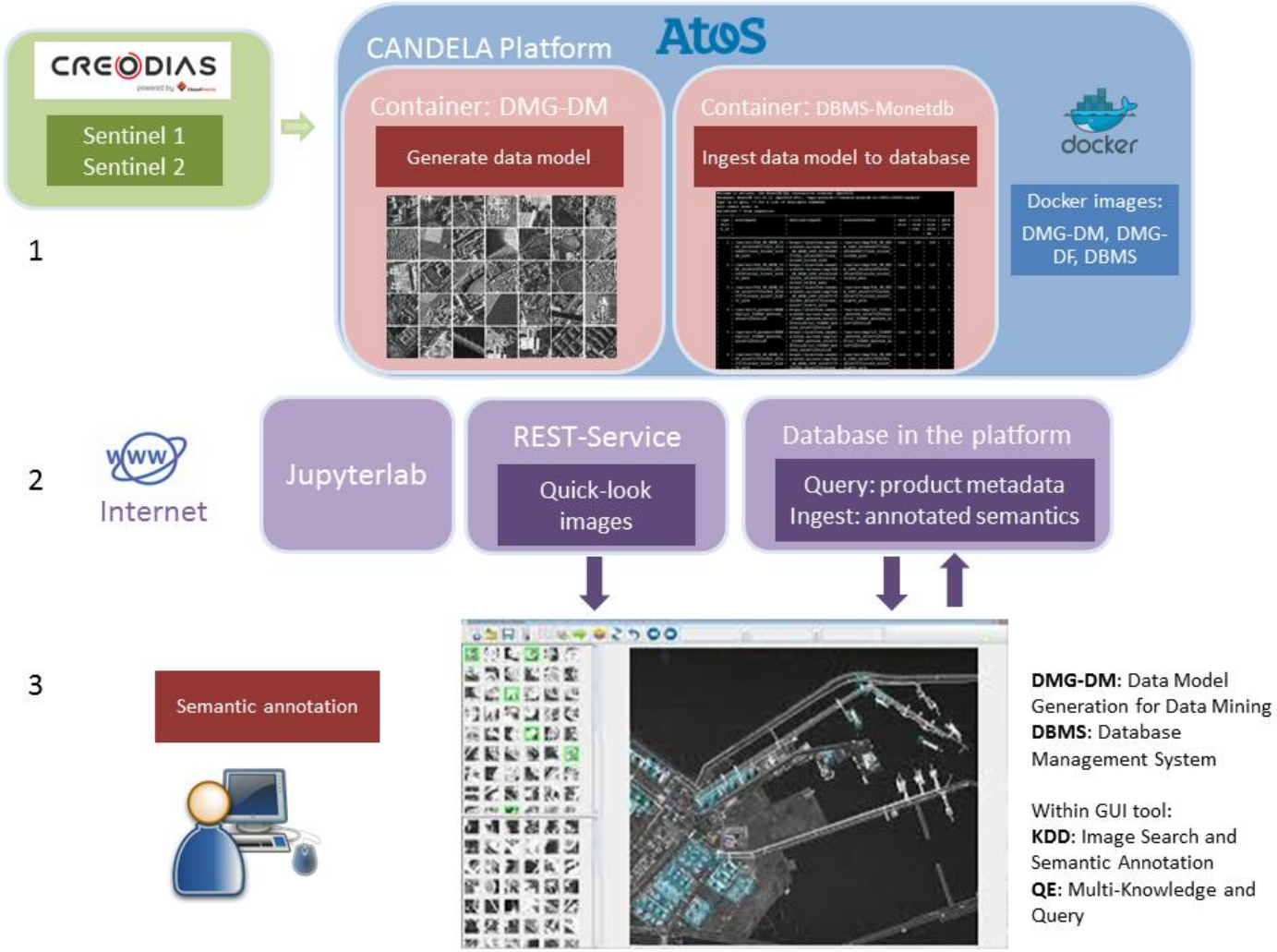


Spatio temporal patterns discovery



DLR

CANDELA



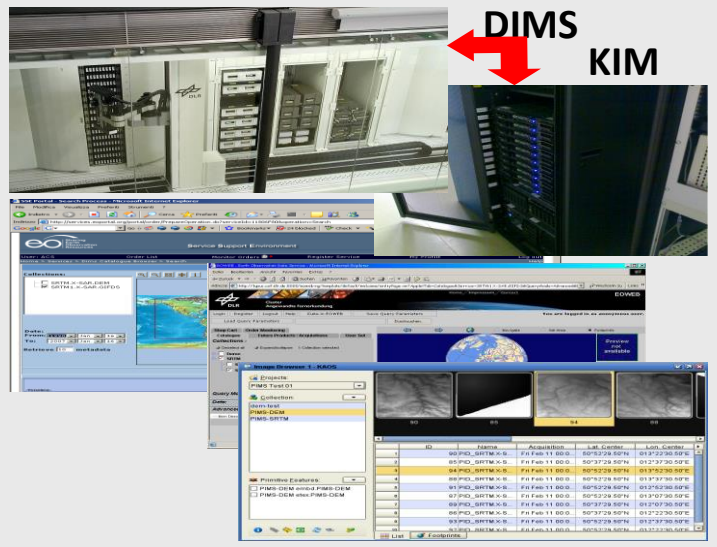
DMG-DM: Data Model Generation for Data Mining
DBMS: Database Management System

Within GUI tool:
KDD: Image Search and Semantic Annotation
QE: Multi-Knowledge and Query

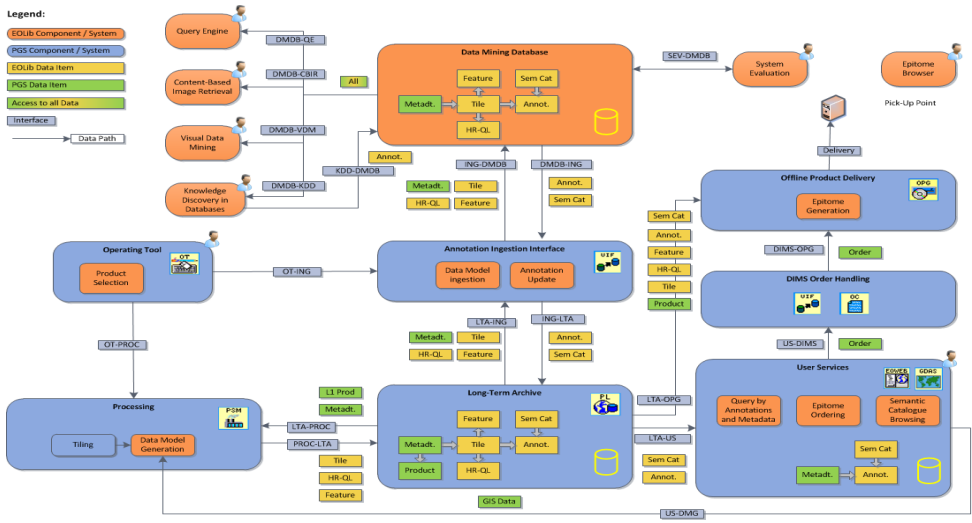
DLR

IA for AI: The Information Architecture

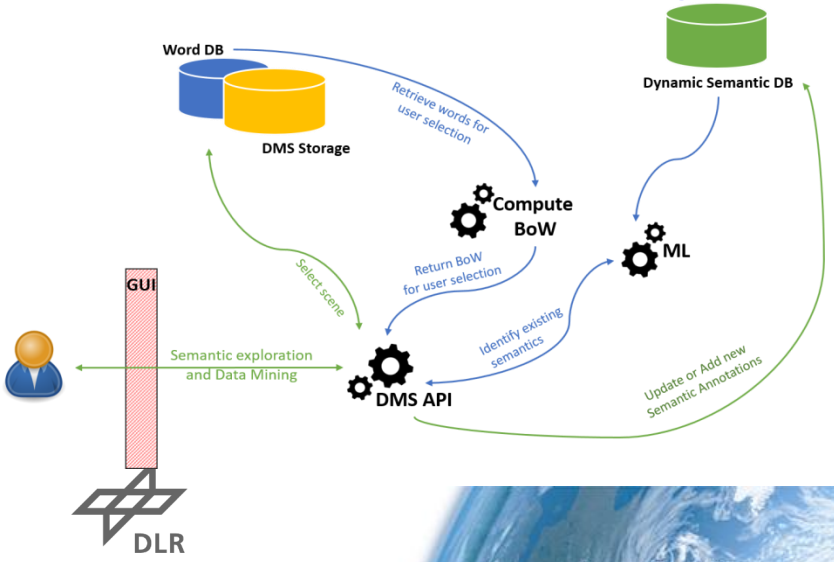
KIM: Bring Data to Algorithms



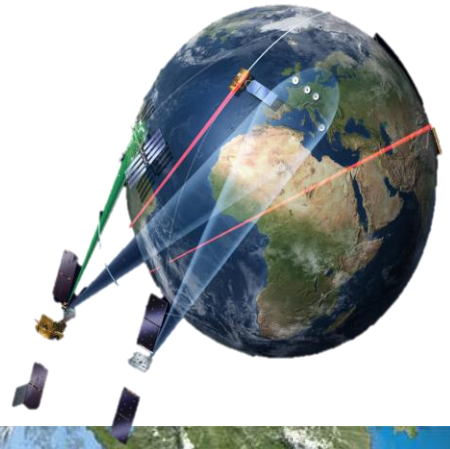
EOLib: Bring Algorithms to Data



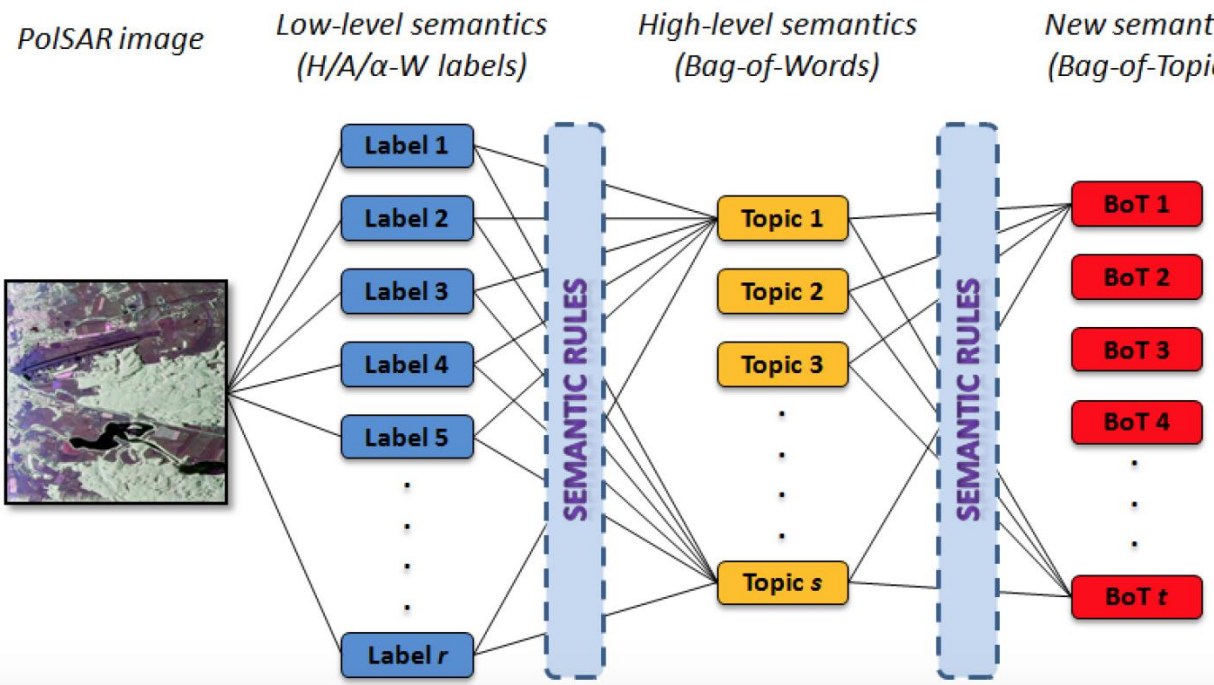
EO Block chain: federated systems



Laser Comm: EO Internet

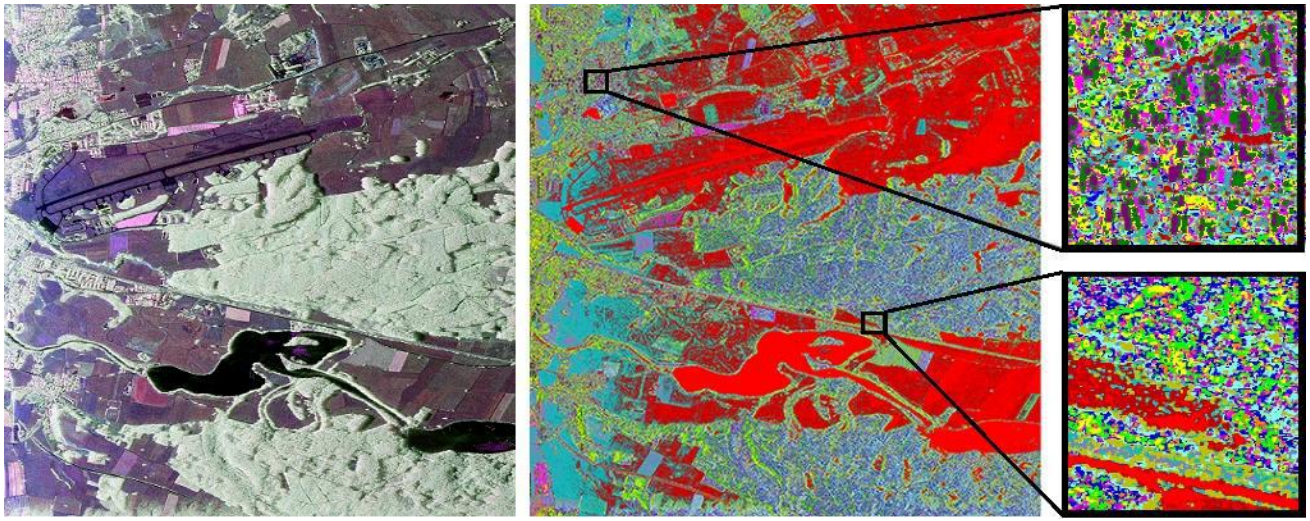


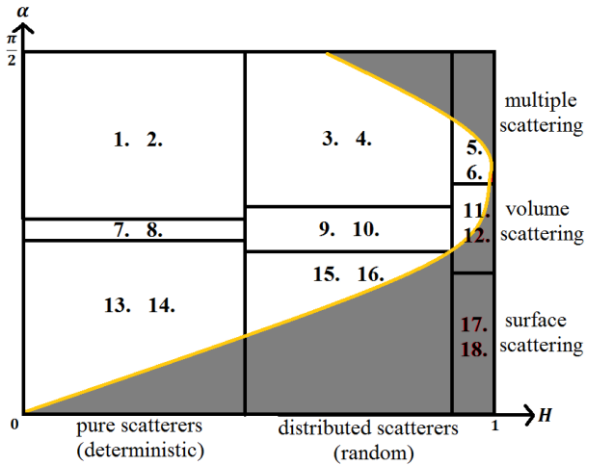
Discovery of semantic relationships: PolSAR L-Band



DLR

PolSAR L-Band Low-level semantics – H\A\α classes





Label	Semantics
L1	Low-entropy multiple scattering with low anisotropy
L2	Low-entropy multiple scattering with high anisotropy
L3	Medium-entropy multiple scattering with low anisotropy
L4	Medium-entropy multiple scattering with high anisotropy
L5	High-entropy multiple scattering with low anisotropy
L6	High-entropy multiple scattering with high anisotropy
L7	Low-entropy volume scattering with low anisotropy
L8	Low-entropy volume scattering with high anisotropy
L9	Medium-entropy volume scattering with low anisotropy
L10	Medium-entropy volume scattering with high anisotropy
L11	High-entropy volume scattering with low anisotropy
L12	High-entropy volume scattering with high anisotropy
L13	Low-entropy surface scattering with low anisotropy
L14	Low-entropy surface scattering with high anisotropy
L15	Medium-entropy surface scattering with low anisotropy
L16	Medium-entropy surface scattering with high anisotropy

R Tănase, R Bahmanyar, G Schwarz, M Datcu

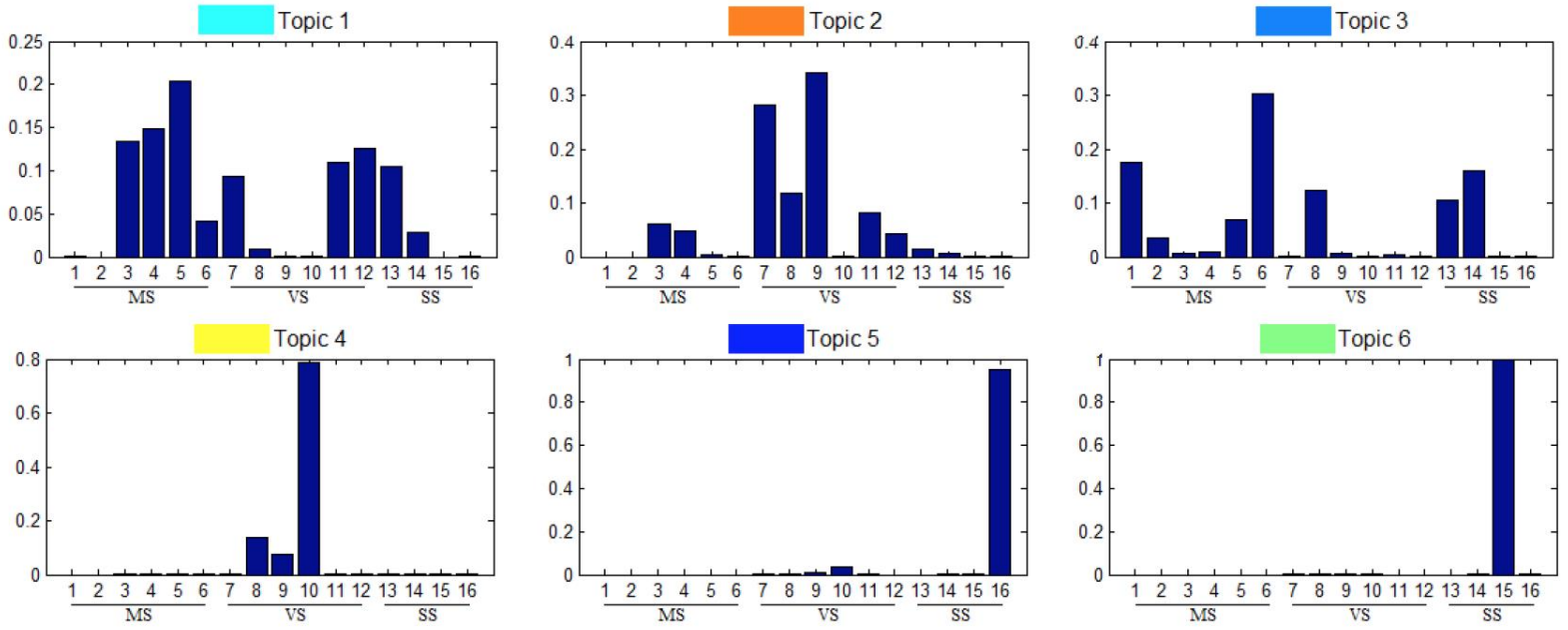
Discovery of Semantic Relationships in PolSAR Images Using Latent Dirichlet Allocation

IEEE Geoscience and Remote Sensing Letters 2017

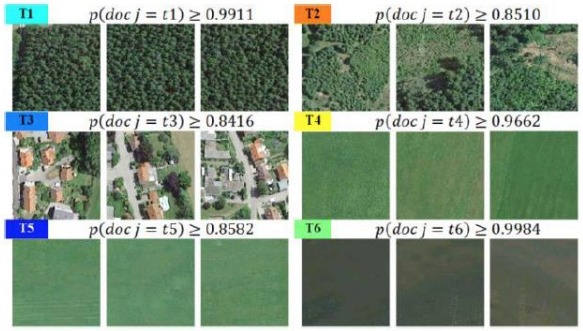
DLR

High-level semantics – LDA topics for PolSAR L-Band

$$p(w_{dn}|a, B) = \int p(\theta_d|a)\Big[\sum_{z_{dn}} p(z_{dn}|\theta_d)p(w_{dn}|z_{dn}, B)\Big]d\theta_d$$



Topic	Semantics	Semantic rule
T1	Woody vegetation	53% MS + 34% VS + 13% SS
T2	Mixed woody vegetation & shrubs	12% MS + 86% VS + 02% SS
T3	Artificial, man-made structures	60% MS + 14% VS + 26% SS
T4	Herbaceous vegetation	00% MS +100%VS + 00% SS
T5	Smooth surface	00% MS + 05% VS + 95% SS
T6	Specular surface	00% MS + 00% VS + 100%SS



DLR

Trends

- Enforce the market/business of AI and Data Analytics in EO
- Grow multi-disciplinary and disruptive thinking HR
- Apply AI where it is needed and works such to obtain more than 20% *better*
- Target to implement solutions for the big, required, but unsolved yet technologies for global and long term EO market:
  - Joint multi-mission and multi-source data valorization
  - Satellite Image Time Series valorization
  - Multisource model assimilation and physical parameters inversion
  - Predictive models and forecast Earth dynamic
  - Quantum technologies
- Elaborate an overall EO intelligence for the system of systems:
- DATA – SENSOR – MISSION – ARCHITCTURE

Threats

- DNN: in 2019 more than 500 papers/month
- Research is often wasted effort
- ML faces a deep reproducibility crisis
- Training data is as important as the learning algorithm
- ML finds any pattern in data, it may be irrelevant
- We need the actual patterns of the Earth processes
- Big EO Data accentuate the crisis

DLR

References

- Dumitru, C.O.; Schwarz, G.; Castel, F.; Lorenzo, J.; Datcu, M.; Artificial Intelligence Data Science Methodology for Earth Observation, in Advanced Analytics and Artificial Intelligence Applications, InTech Publishing, London, UK, September 2019, pp. 75-97, DOI: http://dx.doi.org/10.5772/intechopen.86886
- Tong, X.Y.; Xia, G.S.; Hu, F.; Zhong, Y.; Datcu, M.; Liangpei Zhang Exploiting Deep Features for Remote Sensing Image Retrieval: A Systematic Investigation, IEEE Transactions on Big Data 2019, Early Access Article.
- Bahmanyar, R.; Espinoza-Molina, D.; Datcu, M.; Multisensor Earth Observation Image Classification Based on a Multimodal Latent Dirichlet Allocation Model, IEEE Geoscience and Remote Sensing Letters, vol. 15, issue 3, pp. 459-63, 2018. URL: https://ieeexplore.ieee.org/abstract/document/8278834
- Dumitru, C.O; Schwarz, G.; Datcu, M.; SAR Image Land Cover Datasets for Classification Benchmarking of Temporal Changes, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 5, pp. 1571-92, 2018, IEEE ISSN 1939-1404, URL: https://ieeexplore.ieee.org/abstract/document/8303669
- Murillo Montes de Oca, A. ; Bahmanyar, R.; Nistor, N.; Datcu, M., Earth Observation Image Semantic Bias: A Collaborative User Annotation Approach, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 6, pp. 2462 - 2477, 2017, DOI: 10.1109/JSTARS.2017.2697003, URL: https://ieeexplore.ieee.org/abstract/document/7926346
- Tănase, R.; Bahmanyar, R.; Schwarz, G.; Datcu, M., Discovery of Semantic Relationships in PolSAR Images Using Latent Dirichlet Allocation, IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 2, pp. 237 - 241, 2017, DOI: 10.1109/LGRS.2016.2636663, URL: http://ieeexplore.ieee.org/document/7797191/
- Dumitru, C. O.; Schwarz, G.; Datcu, M., Land Cover Semantic Annotation Derived from High-Resolution SAR Images, in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 6, pp, 2215 - 2232, 2016, DOI: 10.1109/JSTARS.2016.2549557, URL: http://ieeexplore.ieee.org/document/7464254/
- Marmanis, D.; Datcu, M.; Esch, T.; Stilla, U., "Deep Learning Earth Observation Classification Using ImageNet Pretrained Networks," in Geoscience and Remote Sensing Letters, IEEE , vol.13, no.1, pp.105-109, Jan. 2016, doi: 10.1109/LGRS.2015.2499239
- EOLib project; URL: http://wiki.services.eoportal.org/tiki-index.php?page=EOLIB+Project
- ECOPOTENTIAL project; URL: https://www.ecopotential-project.eu/
- CANDELA project; URL: http://candela-h2020.eu/