

Towards a European AI for Earth Observation Research & Innovation Agenda

Proceedings of a workshop at ESA Φ -lab

Reference Φ -lab/mc/2018-23
Issue 1
Revision 3
Date of Issue 24/9/2018
Document Type Workshop proceedings

APPROVAL

Title Towards a European AI4EO R&I Agenda	
Issue 1	Revision 2
Author Pierre-Philippe Mathieu, Sveinung Loekken et al (p17)	Date 16/7/2018
Approved by Pierluigi Silvestrin	Date 7/9/2018

CHANGE LOG

Reason for change	Issue	Revision	Date
Formatting, sub-editing and minor editorial changes	1	3	24/9/2018

Table of contents:

1	Executive Summary	4
2	Scope	5
3	Background	5
4	Challenges and Opportunities	6
5	Timely need for a European initiative on AI4EO	8
6	Getting ready for AI - Community Recommendations	9
6.1	Research and Innovation Activities	10
6.1.1	Exploratory Activities.....	10
6.1.2	Demonstration Activities	12
6.1.3	Challenges Activities	12
6.2	Capacity Building Activities	12
6.2.1	Enabling Open Data sets	12
6.2.2	Expertise and Training.....	13
6.2.3	Tools and Infrastructure	13
6.3	Ecosystem Building Activities.....	14
6.3.1	Research community	14
6.3.2	Private sector	15
6.3.3	Startup and Investor ecosystem	15
7	References	15
	ANNEX 1: Workshop Participants	17
	Annex 2: Glossary	18

1 EXECUTIVE SUMMARY

Developments in digital technologies and our capability to monitor our home planet with Earth Observation (EO) satellites have led to new opportunities for science and business. There is an increasing need to mine the large amounts of data generated by the new generation of satellites coming online, including Copernicus and New Space satellites. Artificial Intelligence (AI) is an important part of the solution, enabling scalable exploration of big data and bringing new insight and predictive capabilities.

AI for Earth Observation (AI4EO) has great potential that remains largely untapped. This is a summary of the recommendations of an expert workshop convened by ESA to better understand how AI and earth observation can support each other. While today, AI4EO is mainly computer vision applied to very high-resolution satellite imagery, there are many areas of Earth science that could benefit from AI. The EO research and business communities are awakening to these opportunities, calling for European collaborative effort to:

Foster AI4EO capacity and the ecosystem:

- Ensure Europe supports the people, skills, and resources to develop AI4EO;
- Increase the supply of AI talent, by training students and data entrepreneurs to develop the new generation of AI-powered EO-based businesses;
- Provide a digital environment for researchers and innovators to rapidly prototype new AI4EO applications, including tools, clean quality-controlled training data sets, computing power and easy access to EO data, training material and expertise;
- Combine strengths across the innovation ecosystem by bringing together universities, start-ups, and investors.

Build AI4EO knowledge and demonstrate scientific and socio-economic value:

- Develop solutions in partnership with industry addressing real world problems,
- Explore key research questions, by setting up scientific and sectorial grand challenges,
- Promote the value through a series of use cases in partnership with users.

Resolve key research questions on:

- Classification and recognition, detection, indexing, understanding and prediction,
- Data fusion, inversion problems, automation and autonomous smart systems.

Addressing some of these challenges at European level, including by leveraging ESA assets, would help realise the full potential of EO and deliver socio-economic benefits. Not doing so would miss huge opportunities for Europe to shape the future of a rapidly changing, global AI landscape now dominated by the US and China. With AI becoming a transformative technology and European EO capability delivering a unique and comprehensive dataset on the planet that could be analysed by AI, we are at a cross roads. Making the most of this window of opportunity is an urgent challenge.

2 SCOPE

This document addresses elements of what needs to be done at European level, some specifically by ESA, to harness the full potential of artificial intelligence to exploit earth observation data and the data supply chain and vice versa.

It captures the recommendations of a community-led AI4EO workshop held at ESA/ESRIN on 27 Mar 2018 (c.f. participant list in Annex 1). The aim of the workshop was to informally assess progress in the development and application of AI techniques to the world of EO and to explore the potential value of a concerted **Research and Innovation (R&I)** effort on this topic at **European level**. This document is meant to be a *dynamic* report capturing the evolving needs of the community. It will be reviewed and updated in a follow-on workshop at ESA/ESRIN on November 14th, 2018.

3 BACKGROUND

The world of EO is changing dramatically driven by rapid advances in sensor and digital technologies. Recent decades have witnessed extraordinary developments in Information and Communication Technologies (ICT), including the Internet, Cloud computing and AI, leading to radically new ways to collect, distribute and analyse big data about our planet. This *digital* revolution is also accompanied by a *sensing* revolution providing an unprecedented amount of data on the state of our planet and its changes.

Europe is leading this sensing revolution in space through the EO missions of ESA, a new generation of meteo missions for Eumetsat, and especially the *Copernicus initiative* in partnership with the European Union (EU). Copernicus is centered around a family of Sentinel missions developed by ESA for the EU so as to enable global monitoring of our planet on an operational and sustained basis over the coming decades.

A new trend, referred to as *New Space* in the US or *Space 4.0* in Europe, is now rapidly emerging through the increasing commoditization and commercialization of space. With the rapidly dropping costs of small sat building, launching and processing, new EO actors including startups (e.g. Planet, Spire), new players (e.g. SpaceX) and ICT giants (e.g. Google, Amazon) are now massively entering the space business, resulting in new constellations of standardized small satellites delivering a new class of data on our planet with high spatial resolution and high temporal frequency.

These new global data sets from space lead to a far more comprehensive picture of our planet. This picture is now even complemented through data derived from millions of smart sensors connected to the Internet (referred to as the Internet of Things IoT) and also from Unmanned Aerial Vehicles (UAVs) systems. Such streams of data offer entirely new possibilities for science but also for entrepreneurs to develop new businesses by turning big data into new types of information services.

However, these *opportunities* come with new *challenges* for scientists, business, data and software providers to make sense of the vast and diverse amount of data by capitalizing on new Big Data Analytics techniques such as AI.

AI driven by Moore's Law and now super-fed by big data is in the midst of a true *renaissance* becoming an integral part of our society, deeply transforming the way we work, operate, and live. Within the report "AI for Earth" presented at the 2018 World Economic Forum [1], AI is even coined to be the new "electricity" fuelling the 4th Industrial revolution.

Until recently AI was mainly a field occupied by restricted experts and scientists. There have also been several hypes of AI in the past decades, but today, AI is routinely used in everyday life without us even noticing it, in applications ranging from recommendation engines, language services (e.g. translation, speech recognition), face recognition, virtual assistants (e.g. Siri, Alexa) and autonomous vehicles (e.g. drones, self-driving cars).

Over the last decade, machine learning went through a major revolution, through the unique convergence of large-scale computing capability (e.g. cloud computing, graphics processing unit architectures, high performance computing), easy access to large volumes of data through the Internet, and the availability of new algorithms enabling robust training of large-scale *deep* neural networks.

In the 2010s, a breakthrough happened: this was the beginning of the deep learning revolution. Scientists started to pay attention to deep learning algorithms as they lead to dramatic improvements in the ability to automatically recognise objects (going below 3% accuracy, well beyond human performance for some tasks only) within the famous ImageNet Large Scale Visualisation Challenge [image-net.org]. Today deep learning, and most recently reinforcement learning, are becoming the workhorses of AI.

4 CHALLENGES AND OPPORTUNITIES

The AI revolution is reinforced by its convergence with other transformative technologies such as IoT, cloud computing and blockchain, which are disrupting entire industry verticals such as automotive, healthcare, transport and banking. For example, the tandem of AI, IoT and blockchain enables the development of new market places to share and monetize AI services. Another example is the emergence of self-driving cars, which are made possible today thanks to the *convergence* and *integration* of technologies such as IoT, cloud computing and AI.

So, will a similar disruption occur in the space sector? Will AI be the key to unlock the potential of the new streams of EO data coming online to better understand changes on our planet and predict its evolution? Will satellite hardware become a commodity with the focus moving to the AI-powered software enabling autonomy and remote upgrade, as is happening with Tesla cars?

In this context, one of the key challenges for the EO community is to harness the full power of new technologies such as AI in collaboration with new players in the ecosystem including ICT companies, startup and data and EO scientists.

In the case of EO, making full use of AI techniques - such as machine learning, deep learning, reinforcement learning or computer vision - requires their adaptation to account for the **specific nature of remote sensing data** being:

- Physically-based and indirect measurement of geophysical parameters estimated from remotely sensed physical quantities across the whole electromagnetic spectrum. This requires significant work for data preparation (e.g. data cleaning/curation, formatting, co-gridding, normalisation), which represents the major task in any AI application. Besides, remote sensing problems involve very high dimensional data (e.g. multispectral, hyperspectral, multiangular, multisensor and multitemporal) which hampers direct adoption of standard deep learning approaches, using pre-trained models, or even designing new topologies. It also requires work to integrate physical principles (e.g. radiative transfer models) into the statistical algorithms underpinning AI. This also requires ground truth data to validate EO products as well as for training AI algorithms. Currently there is a scarcity of training data sets for EO, which is a key limiting factor.
- Diversity of data, derived from a site of sensors including in particular:
 - Multispectral optical sensors acquiring images in several spectral channels. Up to now, AI has been mainly applied to optical imagery, in particular at very high resolution by use of traditional computer vision methods (using mainly RGB bands) but work needs to be done to make full use of *all* available spatial, temporal and spectral EO-derived information at the global scale. This requires work for new development of deep learning schemes but also exploring new methodologies based on reinforcement learning.
 - Radar sensors delivering information on amplitude, frequency, phase or polarization of the collected radar echoes. This requires work to make full use of the "complex nature" of radar data with AI schemes, as well as the full-polarimetric nature of the new generation SAR.
- Volume of data reaching the Petabyte scale, which need to be turned into small actionable information, with additional constraints such as need for near-real time information. This requires work to perform statistically-based big data mining while taking account of EO physics.
- Complexity of data capturing dynamic features of a highly non-linear coupled *Earth System*. This goes well beyond recognising cats and dogs in images where a wide variety of training data sets are available (e.g. ImageNet). It also presents much more potential to unlock socio-economic value of the data by enabling a myriad of applications of "AI for Good" helping to manage our natural resources and inform decisions. This requires work to properly train AI algorithms and adapt them to the variety of spatial/temporal scales inherent to EO data in order to better detect, quantify, understand and predict what is happening on our planet on a routine and automatic basis. This also presents a huge opportunity for the AI community to go "AI for Good" as the EO data have considerable value for our society.

There is also a lot of work in EO with leading European actors, but the intersection of AI and EO remains an emergent field, albeit a rapidly growing one. There has indeed been a lot of work on machine learning applied to EO over the past decade, but with the rapid emergence of deep learning, the field has been growing rapidly as illustrated by the increase in number of publications.

We are now at a *cross road of opportunities*, where on the one hand AI is becoming one of the most transformative technologies of the 21st century, while on the other hand European EO capability is delivering a totally unique and comprehensive picture of the planet, thereby generating big open data sets to be explored by AI. Making the most of this window of opportunity is a challenge, and an urgent one as discussed below. It is also worth noting that this intersection of AI and EO, and is, in fact, a particular case of a wider intersection of converging transformative techs such as AI, IoT and Blockchain, which is beyond the scope of this report, but offers even more opportunities.

5 TIMELY NEED FOR A EUROPEAN INITIATIVE ON AI4EO

Fostering the *intersection of AI and EO* has a lot of potential for both science and society through the development of new knowledge, businesses and jobs. Many European labs, companies and research groups are now exploring a variety of AI4EO applications. But recent studies [1-4] suggest that **Europe is not keeping up with its AI capability**.

Europe is strong but fragmented. Europe has leading research centers such as Cambridge, CNRS, MPI, DFKI, DLR, ETHZ, INRIA and others. New research organisations and initiatives are also rapidly developing, such as AI for Humanity in France, the Alan Turing Institute in UK, the Max Planck Center for Learning Systems in Germany and Switzerland and a recent initiative ELLIS to establish a European Lab for Learning and Intelligent Systems. Nonetheless most of the top labs and centres of AI research and business are located in US and China, where huge investments are made by the government and large ICT corporates.

The US publishes the most AI papers in the world [3], has the most AI patent applications (15.000+ in 2010–2014), has the most AI workers (approximately 850.000+), fuels the most innovative AI startups and accounts for the biggest share of private AI investments. It hosts global companies that lead in AI, such as Google, Apple, Facebook, Amazon, and new AI players, such as Nvidia born as a gaming company and now becoming a world-leading developer of graphics processing unit chips tailored for AI.

China has the ambition of becoming the world leader in AI within the decade, supported by a strong political agenda [4] that aims to build a domestic AI industry worth 20+ B\$ by 2020. China is developing leading research centers in AI and hosts the largest web giants, Baidu, Alibaba, Tencent and Xiaomi (BATX), which are leading cutting-edge developments in AI (e.g. face recognition, recommendation engine, predictive analytics).

Europe currently has no major AI champions, except perhaps SAP, and there is a risk that many European companies critically depending on AI are not perceived as competitive in a global market, because they are not already adopting AI today across business verticals.

AI is leading to an extraordinary boom in business opportunities but again mainly driven by the US, in particular from Silicon Valley. At the intersection of AI and EO, many startups are emerging using computer vision across industry sectors (e.g. retail, agriculture, urban, insurance). There are some promising examples in Europe, such as RSMetrics [rsmetrics.com] for real estate and Earthcube [earthcube.eu] for infrastructure management, but the vast majority of business was developed in US, over the last few years, capitalizing on easy access to capital from Venture Capitalists (VCs). This includes companies like Descartes Labs [descarteslabs.com] predicting crop yield and price, Orbital Insight [orbitalinsight.com] predicting retailers profits by counting cars in the parking lots, SpaceKnow [spaceknow.com] and URSAspace [ursaspace.com] estimating world oil inventory, CapeAnalytics [capeanalytics.com] supporting management of properties, and TellusLabs [telluslabs.com] building map of the food production and economic activity. These companies managed to raise considerable sum of money from VCs. Also, leading EO companies in US such as Digital Globe are now acquiring AI capabilities through acquisition of other companies (e.g. Radiant Group and timbr.io).

This leads to a **sense of urgency to act now in AI acknowledged by the EC** [3], and for EO, a pressing need to develop innovative AI4EO capabilities to fuel the new generation of EO business and scientific discovery in this domain. To strengthen the position of Europe, there is a need to capitalize on what Europe is strong at (i.e. its research capability, its diversity, its talents) and be ambitious (think big).

There is now a unique and timely opportunity to play a major enabling role in the scientific and societal AI revolution that is underway by enabling AI4EO at a large European scale, thereby delivering new Earth science and more competitiveness for the European service industry and space sector. To play that role, it is critical to better:

- Enable rapid transfer of AI knowledge, techniques and expertise from data scientists to the world of EO research (in both directions) and business applications,
- Foster new partnerships with non-space and ICT players in order to jointly co-develop new innovative EO-based solutions,
- Develop and rapidly prototype innovative AI-based EO solutions by providing the necessary digital infrastructure, data and tools,
- Empower the new generation of researchers, data entrepreneurs, and digital startups with AI4EO capability. These are now becoming the main engines of innovation,
- Harness the involvement of citizens through crowdsourcing, by integrating this data stream into innovative solutions,
- Deliver economic impact and create jobs in Europe by creating AI-powered EO solutions addressing real industry challenges.

6 GETTING READY FOR AI - COMMUNITY RECOMMENDATIONS

To deliver on AI4EO, the workshop identified the actions set out below. This list is not comprehensive and will be complemented by further workshops and activities.

6.1 Research and Innovation Activities

There is a critical need to conduct a suite of R&I activities aiming to accelerate the evolution of technical capabilities of European AI4EO research and industry and foster the uptake of AI in the EO community across applications, from EO product to satellite tasking. A "mission-driven" approach [2,6] to applied research is advocated, similar to that employed in climate informatics. Project opportunities should be specifically designed to stimulate inter-disciplinary interactions at European level, and thereby enable European actors to remain at the forefront of relevant technology developments. A series of activities should support fundamental and applied research, as well as technology transfer and commercialization of research (only when using AI makes sense). Actions should adopt open standards to promote sharing of knowledge, access to data sets and algorithms and, where possible, open source software.

6.1.1 Exploratory Activities

There is a need to support *fundamental research* in AI4EO (including Exploratory / Proof of Concept studies) applying/adapting existing AI schemes/architectures (e.g. machine learning, deep learning, reinforcement learning, computer vision) or new hybrid schemes (models/statistics) to EO problems and better understand "*why*" AI schemes operate as they do (so-called XAI, Explained AI). Fundamental research is also essential to develop the methods and tools that will accelerate the application of AI to EO data.

Some classes of EO-specific problems are given below:

- Classification / Recognition
 - Characterising land properties, updating land-cover maps,
 - Integrate electromagnetic transfer scheme in AI classification schemes,
- Detection
 - Identifying hot spots, patterns, features, objects (e.g. ship),
 - Detect changes at the global scale and on an automatic basis,
 - Detecting features using complex information from SAR in machine learning schemes,
- Indexing
 - Querying up large-archive of EO data with scalable search engines for text (e.g. metadata) building on Natural Language Processing (NPL),
 - Develop EO ChatBot to find the data needed, and operate as an automatic helpdesk to provide guidance in using the data,
 - Management of storage of big data,

- Understanding
 - Identifying Earth System processes and feedback loops,
 - Find physical mechanisms underpinning AI statistical relationships (bringing induction and deduction approaches together, Explainable AI, model interpretation),
 - Observational (empirical) causal inference
- Prediction
 - Quantifying Earth System risks and socio-economic conditions,
 - Forecasting essential Earth parameters (e.g. crops, climate, biodiversity),
- Data Fusion
 - Merging diverse EO data and derived parameters, e.g. New Space/Sentinel,
 - Integrating open data, IoT, UAV, social media and crowdsourced data with EO data,
- Optimisation
 - Estimating parameters in an optimal and effect way (e.g. retrieval),
 - Reducing dimensionality for big data (e.g. hyperspectral data compression),
- Inversion problems
 - Development of schemes fusing different spatio-temporal scale (e.g. super-resolution),
 - Development of new generation joint multi-sensor retrieval schemes,
 - Development or improvement of bio-geophysical parameter estimation (e.g. biomass, aerosol concentrations),
- Automation
 - Development of automatic processing workflows, increasing performance, reducing cost, (e.g. optimised planning, scheduling of operations),
 - Development of virtual assistant (bots),
 - Development of on-board processing systems,
- Autonomous Smart systems
 - Development of on-board AI to ensure self-guiding autonomous vehicles / robots,
 - Development integrated "smart" observing systems such as IoT/constellations,
 - Demonstrate Swarm intelligence with EO and IoT.

This suite of exploratory projects also includes cross-cutting activities addressing issues of hybrid AI / optimization schemes, use of Bayesian techniques in deep learning to better understand uncertainty [bayesiandeeplearning.org], issue of certification of / explaining AI schemes (looking deeper into the AI blackbox).

6.1.2 Demonstration Activities

Development of a portfolio of Use Cases in partnership with user / industry players,

- Demo projects addressing real business solutions across industry verticals, which also integrate other types of data (e.g. crowdsourcing) in the value chain,
- Applied-research addressing "sectorial grand challenges" in EO and "mission-driven" focus (matching EO solutions with business problems) is advocated,
- Explore development of new market places for AI services, capitalizing on blockchain technologies,
- Technology transfer of innovative capabilities from R&D for use in applications and commercial exploitation.

6.1.3 Challenges Activities

Development of a suite of AI challenges, addressing data scientist and innovators, using innovation platform like Kaggle [kaggle.com] or other crowdsourcing tools:

- Data Contest Challenges and hackathons addressing new products (e.g. Copernicus Masters), Earth system monitoring and understanding (e.g. fdleurope.org), classification Scheme,

6.2 Capacity Building Activities

Create a data ecosystem for AI4EO, providing researchers, industry, and institutions with the data required, in the form required, in order to remove the bottlenecks of data preparation and training data creation.

6.2.1 Enabling Open Data sets

- Develop training data sets with clean, quality-controlled, and labelled data. This is a pre-requisite to do AI, and there is a pressing need for such data in EO are currently lacking. What is needed is the equivalent of Imagenet for EO. Spacenet is a first step in this direction but work needs to be done to extend the concept to other EO data such as the Sentinels.
- Adapt, develop and integrate solutions that make full use of citizen science and crowdsourcing tools to generate annotated data sets and improve accuracy of labelling, which should contribute to more authoritative training data sets (point outlined above),

- Develop synthetic training data sets generated through simulations (e.g. based on radiative transfer models) with perfectly known labels. This kind of data has many benefits such as speed (fast to generate), accuracy (using physically-based models), tailored for specific needs and scalable. This can also explore how generative adversarial networks can generate "realistic" or physically-based data for the EO community,
- Supporting activities aimed at generating and providing access to harmonized analysis ready, quality controlled satellite data (e.g. radiometrically and geometrically corrected, data cube, tools and APIs) from different missions and operators,
- Develop a common protocol and create incentives for EO related projects, private industry and crowdsourcing campaigns to publish and share annotated data sets for training data purposes, e.g. via the marketplace concept for land cover data currently being developed as part of the LandSense project (lep.landsense.eu),
- Facilitate coordination between European data providers, ground segment operators, and infrastructure providers to define and implement harmonized standards for enabling data sets.

6.2.2 Expertise and Training

- Develop workshops, dedicated training, summer schools and online course for students, academics, and industrial participants on topics related to AI4EO,
- Create a permanent world-class physical venue for AI4EO in Europe, housing researchers, engineers, students, and industry, devoted to EO-specific aspects of AI, with a strong focus on an interdisciplinary approach, bringing academics together within teams of EO (remote sensing) and AI (applied mathematics, computer science), and addressing the whole lifecycle from science to industrial activity, including schemes of visiting scientists (e.g. Φ -lab hosting) and collaboration with industry,
- Elaborate European PhD and MSc program / curriculum in cooperation with universities focusing on disciplines needed to address AI4EO,
- Support creation of startups where they keep their Intellectual Property rights.

6.2.3 Tools and Infrastructure

Set up a shared AI4EO sector virtual research infrastructure to provide secure and tailored access to the various participants in the AI4EO ecosystem, e.g. researchers, companies, and public institutions, to useful data for the development of AI, as well as to software resources and extensive computing infrastructure required. This infrastructure should allow users to develop new functionalities tailored to their individual features, objectives and processes.

- Coordinate with existing infrastructure and AI platform initiatives in Europe, e.g. Copernicus Data and Information Access Services (DIAS), the H2020 ICT-26, and the European Open Science Cloud (EOSC), in order to on-board AI4EO users and use cases
- Establish and maintain a European AI4EO virtual environment – e.g. hub, network, or platform – dedicated to AI for EO, where exploitation activities can be performed, providing access to certified annotated training datasets, and analysis ready data, co-located with the resources and tools required by AI4EO practitioners
- Establish and maintain a European AI4EO sandbox environment, easily accessible with minimum formality, providing full support to participants, and access to data, expertise in AI and EO, existing open tools (e.g. EOLib, EOminer, Neumapper [wiki.services.eoportal.org]) and resources for use in development and experimentation
- Develop and provide access to open source tools required by the AI4EO community.

6.3 Ecosystem Building Activities

The idea here is to build an interdisciplinary ecosystem of European AI4EO actors under a single banner, bringing together research, industry, institutions, and users, and based on the major challenges of AI4EO. Actions aim to foster bi-directional bridging among communities, formulate common objectives, stimulate synergies, conduct common activities, initiatives, and programmes and develop, attract, and retain expertise.

6.3.1 Research community

- Create a bridge between the EO and AI communities, in both directions.
- Engage with leading AI centers of excellence and initiatives, such as AI for Humanity, Learning systems and the recently created Vector Institute [vectorinstitute.ai] in Canada,
- Engage with citizens through a suite of CS activities addressing AI4EO, both for validation and classification of EO data [e.g. geo-wiki.org, laco-wiki.net, lep.landsense.eu, picture pile],
- Define a syllabus and conduct recurring training courses
 - AI4EO – aimed at instructing EO practitioners in AI
 - EO4AI – aimed at instructing AI practitioners in EO
- Establish a recurring forum, such as a yearly workshop, where the community meets under the topic of AI4EO,
- Establish sessions on AI4EO in interdisciplinary conferences

- Establish a European expert group or network on AI4EO, with representative participation from R&D, industry, and institutions, with the goal of establishing communication lines to relevant R&D programmes at the National and European levels,
- Build an alliance of European stakeholders aimed at providing dense, open annotated data sets of certified quality. In this context instigate and participate in a major European activity to benchmark labelled and validated ‘ground truth’ training and validation data, linking existing organizations and initiatives
- Establish dedicated sessions at Big Data from Space (BIDS), EO Open Science, Living Planet conferences and others, showcasing results and setting direction for future work, and addressing technology transfer.
- Attract students from mathematical fields to build a solid foundation for AI4EO.

6.3.2 Private sector

- Conduct focused technology transfer workshops with participants from research and industry specifically addressing technology and requirements transfer,
- Support industry-led initiatives to explore and develop the market potential of AI4EO
- Foster the dialogue with private funding, e.g. VC funds,
- Define and conduct a call and challenges specifically for activities/projects for industry seeking commercialization of AI4EO.

6.3.3 Startup and Investor ecosystem

- VCs have started to invest significantly in the space sector, in particular in US (e.g. SpaceX, OneWeb, Orbital Insight, Planet). It therefore becomes increasingly important to connect with Investors, including Business Angels and VCs,
- Develop a Sandbox approach for startups to be able to rapidly develop and test prototypes of AI-based services, in particular addressing real-world problems,
- Support new types of researchers/entrepreneurs through industry fellowships,
- Create links with the startup ecosystem including an AI accelerator, innovator / challenges, develop trust for investors through an ESA-based label (kind of certification), and development of a European tech pass concept [5].

7 REFERENCES

1. Harnessing AI for the Earth, PwC, World Economic Forum WEF, 2018. http://www3.weforum.org/docs/Harnessing_Artificial_Intelligence_for_the_Earth_report_2018.pdf
2. For a meaningful AI – towards a French and European strategy. Villani et al., 2018. https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf
3. Declaration – Cooperation on AI, EU Member States, 2018. <https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence>
4. AI strategy for European startups - Recommendations for policymakers, Asgard - Roland Berger, 2018, <https://asgard.vc/wp-content/uploads/2018/05/Artificial-Intelligence-Strategy-for-Europe-2018.pdf>
5. Mission-oriented Research & Innovation in the European Union, a problem solving approach to fuel innovation-led growth, M. Mazzucato, 2018
6. Lary D.J. et al. 2018. Machine Learning Applications for EO. In: Mathieu PP., Aubrecht C. (eds) Earth Observation Open Science and Innovation. ISSI Scientific Report Series, vol 15. Springer, Cham, https://doi.org/10.1007/978-3-319-65633-5_8.

ANNEX 1: WORKSHOP PARTICIPANTS

The following participated in the AI4EO R&D community consultation at ESRIN, 26-27 March 2018.

Richard	Bamler	German Aerospace Center	DE
Lorenzo	Bruzzone	University of Trento	IT
Gustau	Camps-Valls	University of Valencia	ES
Gabriele	Cavallaro	Jülich Supercomputing Centre	DE
Thomas	Corpetti	Université Rennes 2 Haute-Bretagne	FR
Mihai	Datcu	German Aerospace Center	DE
Fabio	Del Frate	University of Tor Vergata	IT
Begüm	Demir	Technical University of Berlin	DE
Mark	Doherty	European Space Agency	ESA
Steffen	Fritz	IIASA	AT
Sveinung	Loekken	European Space Agency	ESA
Markus	Maertens	European Space Agency	NL
Alessandro	Marin	Rhea	IT
Pierre			
Philippe	Mathieu	European Space Agency	ESA
Allan	Nielsen	Technical University of Denmark	DK
Konrad	Schindler	ETH Zurich	SUI
Xiaoxiang	Zhu	German Aerospace Center	DE

ANNEX 2: GLOSSARY

Artificial Intelligence (AI) is intelligence exhibited by machines that can observe, perceive and act upon their environment to maximize their chance of success at some goal. It refers to the capacity of an algorithm to assimilate information to perform tasks that are characteristic of human intelligence, such as recognizing objects and sounds, contextualizing language, learning from the environment, and problem solving. AI systems in development today remain mainly "optimizing" tools, operating as specialized expert systems that use a database of knowledge to make decisions (mainly through inference, not really "Intelligence"). Researchers are however working on a "strong" AI where machines can perform the full range of human cognitive capabilities. Within this report, the term "AI" will therefore be used as a generic term to mainly refer to Machine Learning adapted to work with geospatial data.

Machine Learning is a branch of AI relying on algorithms that are capable of learning from both data and through human interactions (e.g. supervision) to enable prediction, but is also used for data mining (i.e. discovery of unknown properties and patterns). Machine learning is a field of statistical research for training computational algorithms that split, sort, transform a set of data in order to maximize the ability to classify, predict, cluster or discover new patterns in target datasets. Machine learning allows computers to learn how to deal with problems without programming. In fact, machine learning generates models by taking some data for training a model, and then makes predictions. Machine learning relies on a wide variety of algorithms (supervised and unsupervised) [6], ranging from simple Symbolic Regression, Neural Network, decision tree, Support Vector Machine (SVM), up to genetic programming and ensemble methods such as random forest.

Deep Learning is a type of machine learning algorithm that aim to solve the same kind of problems by mimicking the biological structure of the brain and construct hierarchical architectures of increasing sophistication. There are a wide variety of network architectures including Convolutional Neural Network, Recurrent Neural Networks, Long-Short Term Memory, Generative Adversarial Networks, Deep Belief Network, and stacked auto-encoders. People started to pay greater attention to deep learning in 2012 when a deep learning model improved dramatically the accuracy of image recognition within the ImageNet competition. Today, deep learning achieves accuracy better than human performance, and could yet substitute handcrafted feature extraction, thereby enabling totally automatic image recognition of big data (including EO) and opening huge opportunities for new science and business.

Reinforcement Learning is an area of machine learning inspired by behavioural psychology, concerned with how software agents should take actions in an environment so as to maximize some notion of cumulative reward, learning from "mistakes".

Big Data Analytics is a suite of analysis techniques aiming to deliver "value" from big datasets, whose Volume, Velocity, Variety, Veracity, and Value is beyond the ability of traditional tools to capture, store, manage and analyse. Within this report, the word big data analytics will mainly address machine learning.

Crowdsourcing is the practice of public participation and collaboration in a common goal. Within this report, the term will also be used as a synonym for Citizen Science when the goal is to do research. Citizen scientists can help in processing / analysing EO data (e.g. visual interpretation of land cover and identification of other features visible from very high resolution images) but also in generating new observations (e.g. air quality measurements using a mobile phone (ispex.nl) or a variety of new mobile apps for recording observations on the ground) for a myriad of applications, ranging from land cover validation to animal tracking to humanitarian response. There are also many emerging synergies between AI and CS, in both directions, where AI can help analyse the data of citizens, and where citizens can train the AI through generation of data sets (e.g. labelled observations).

Computer Vision is a field concerned with the automatic extraction, analysis, and understanding of useful information from a single image or sequence of images (e.g. videos). It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding. Over the last decade, computer vision has dramatically improved its capability enabling challenging tasks, such as identifying cars, buildings or changes, to be done by a machine in an automatic manner. Such potential of computer vision has not yet been fully harnessed in EO where specific problems are still to be explored (e.g. smaller objects, more spectral bands), thereby holding great promises for EO.

Image Recognition Challenges have played a key role in the AI renaissance, development of machine learning algorithms and breakthrough of deep learning schemes. They provide the benchmarking framework necessary to test / rank a suite of algorithms but train them through the availability of labelled data sets (being a pre-requisite to train machine learning algorithms). Here are examples.

1. ImageNet LargeScale Visualisation Challenge [image-net.org] is the most famous as it has provided the community with a very large database of images (14M+) for training and benchmarking of classification algorithms. Such databases have played an instrumental role in improving machine learning schemes and the emergence of deep networks, in particular to recognise objects (like cats and dogs). The database

was made possible partly thanks to the ability to crowdsource labelling on the Amazon Mechanical Turk.

2. Spacenet [explore.digitalglobe.com/spacenet] is a satellite imagery object detection challenge with very high resolution EO data (e.g. WorldView, DigitalGlobe). This competition challenges scientists to submit algorithms that automatically detect specific patterns (e.g. building footprints) in satellite imagery.

3. IEEE GRSS Data Fusion Contests are held each year to encourage researchers to develop new methods related to image processing and fusion of data from multiple sources. The contest winners submit papers to the annual IGARSS conference each year.

Internet of Things (IoT) is a network of small sensors connected to the Internet and exchanging data, including for example mobile devices, vehicles, and home appliances. Experts estimate that, by 2030, about 30 billion objects will be connected to the Internet.