Digital and green transformation of European regions and greenhouse gas emissions

INNOVA MEASURE IV online Workshop

Stefano Bianchini [a] Giacomo Damioli [b] Claudia Ghisetti [c]

[a]BETA - Université de Strasbourg, France
[b]Eurpean Commission, Joint Research Centre
[c]Università Cattolica del Sacro Cuore, Milano

The context

- Policy need of benchmarking the EU in the development of new scientific research and technological digital and green knowledge
  - ✓ 2030 Agenda for Sustainable Development (UN General Assembly, 2015)
  - ✓ European Green Deal (EC, 2019)
  - ✓ Science, Research and Innovation Performance of the EU report (EC, 2020) "*A fair, green and digital Europe*"
  - ✓ Wide range of investments, financing tools for environmentally-friendly technologies, support to industry innovations, etc.
- Digital transformation generates enabling / reinforcing knowledge and technologies for meeting societal challenges (Cockburn et al., 2018)
- Lack of documentation and data to support policy decisions

The objectives

- Unique dataset on regional digital and green scientific (publications) and technological (patents) knowledge and greenhouse gas emissions for the 2000-2018 period in the European Union (EU28), Norway and Switzerland at the NUTS3 regional level

- Range of metrics and statistics for a list of selected variables to illustrate the potential use of the data for policy and academic purposes

- First step of a more comprehensive research project to address the policy need for benchmarking EU regions in the development of scientific research and technological digital and green knowledge

Data sources

- Scientific research:
  - ✓ Web of Science (WoS) publications in digital scientific fields (DS)
  - ✓ WoS publications in green scientific fields (GS)
- Technological knowledge:
  - ✓ European Patent Office (EPO) patent applications in digital technologies (DT)
  - ✓ EPO patent applications in green technologies (GT)
- Emissions of selected pollutants:
  - ✓ European Pollutant Release and Transfer Register (E-PRTR)
  - ✓ Emissions related to industrial production (no consumption and transport)

Knowledge and technologies

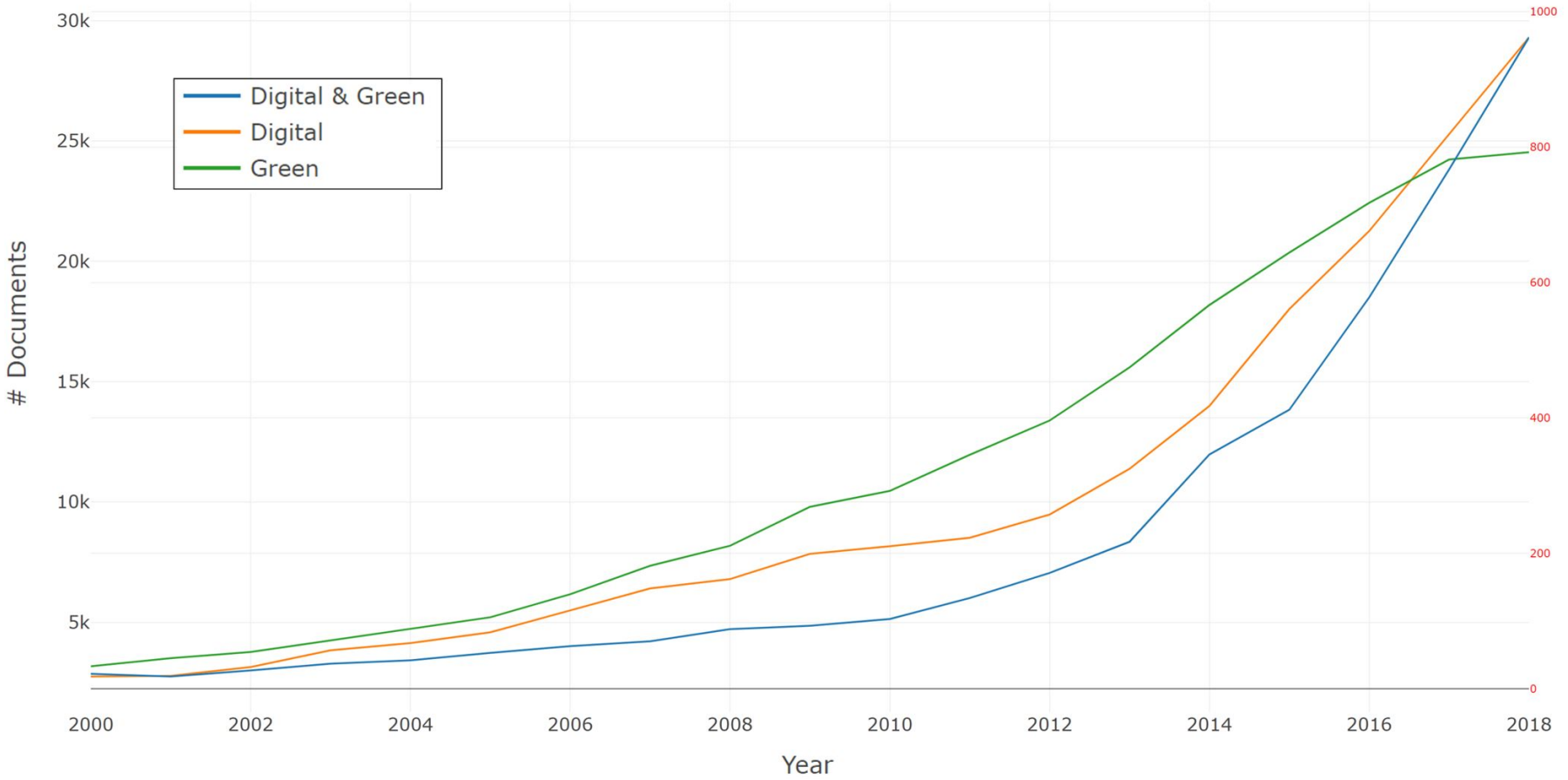
- Digital: Keyword-based approach recently proposed in Van Roy et al. (2020), complemented by a few additional terms referring to the ongoing digital transformation (e.g. IoT, Additive manufacturing) Backup
  - ✓ 221,284 WoS documents (103,853 published articles; 103,120 conference proceedings; 14,311 other) + Geolocalization using authors' affiliation
  - ✓ 14,095 European Patent Office applications

- Green: OECD classification of environmental technologies (OECD Envtech), refined identifying the sub-categories related to GHG emissions in addition to the overall green-tech domain + Keyword-based approach for WoS Backup
  - ✓ 241,088 WoS documents (161,642 published articles; 61,304 conference proceedings; 18,142 other) + Geolocalization using authors' affiliation
  - ✓ 97,493 European Patent Office applications

Emissions

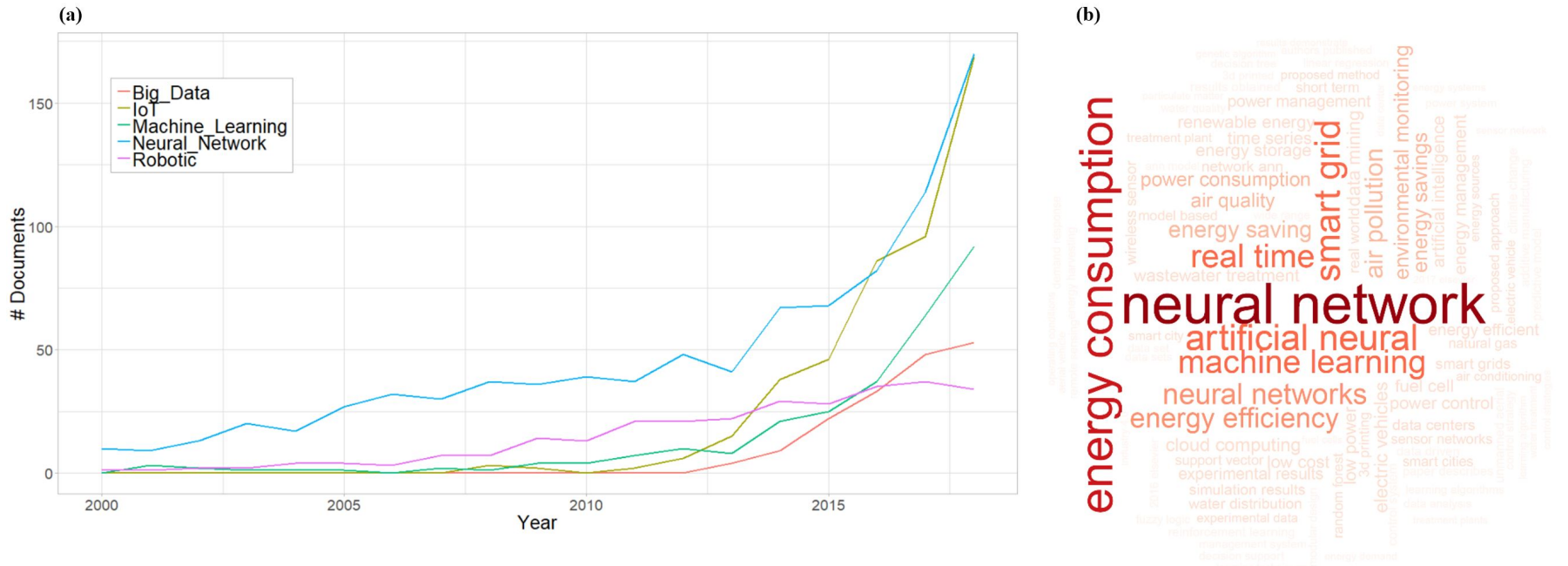
- Air emissions of selected pollutants from the European Pollutant Release and Transfer Register (E-PRTR) dataset
  - ✓ Nitrogen oxides (NOx), nitrous oxide (N20), sulphur oxides (SOx), greenhouse gases (GHG), ammonia (NH3), carbon dioxide (CO2) and particulate matter (PM10)
- E-PRTR data cover compulsory reporting of emissions for the period 2007-2017 by EU Member States; obligation for firms to register emissions when they exceed the applicable capacity thresholds specific to the sub-sector and the substance emitted
  - ✓ Emissions related to industrial production
  - ✓ 6,876 facilities (28% have reported data on emissions of greenhouse gases throughout the full period)

Publication activity, by category



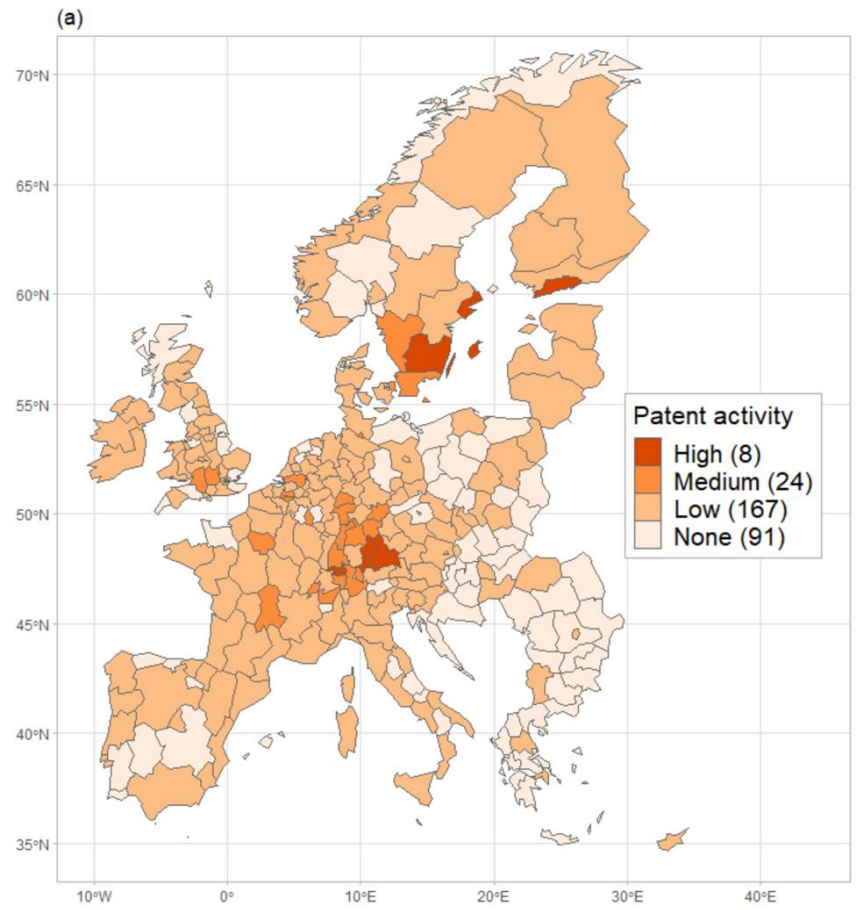
Notes: The left-hand scale refers to the categories "Digital" and "Green". The right-hand scale refers to the category "Digital & Green".

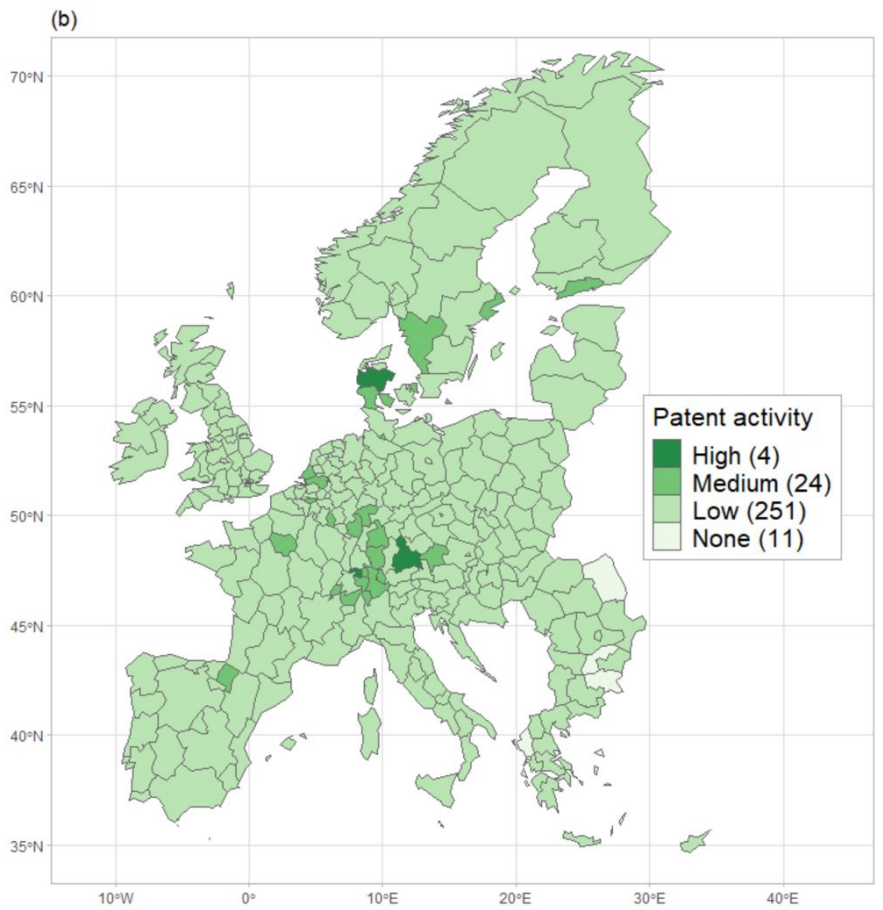
Trends in digital-green scientific knowledge



Notes: Trend in most frequent keywords related to "digital-green" scientific knowledge (panel a). Most frequent bigrams in "digital-green" publications (panel b).

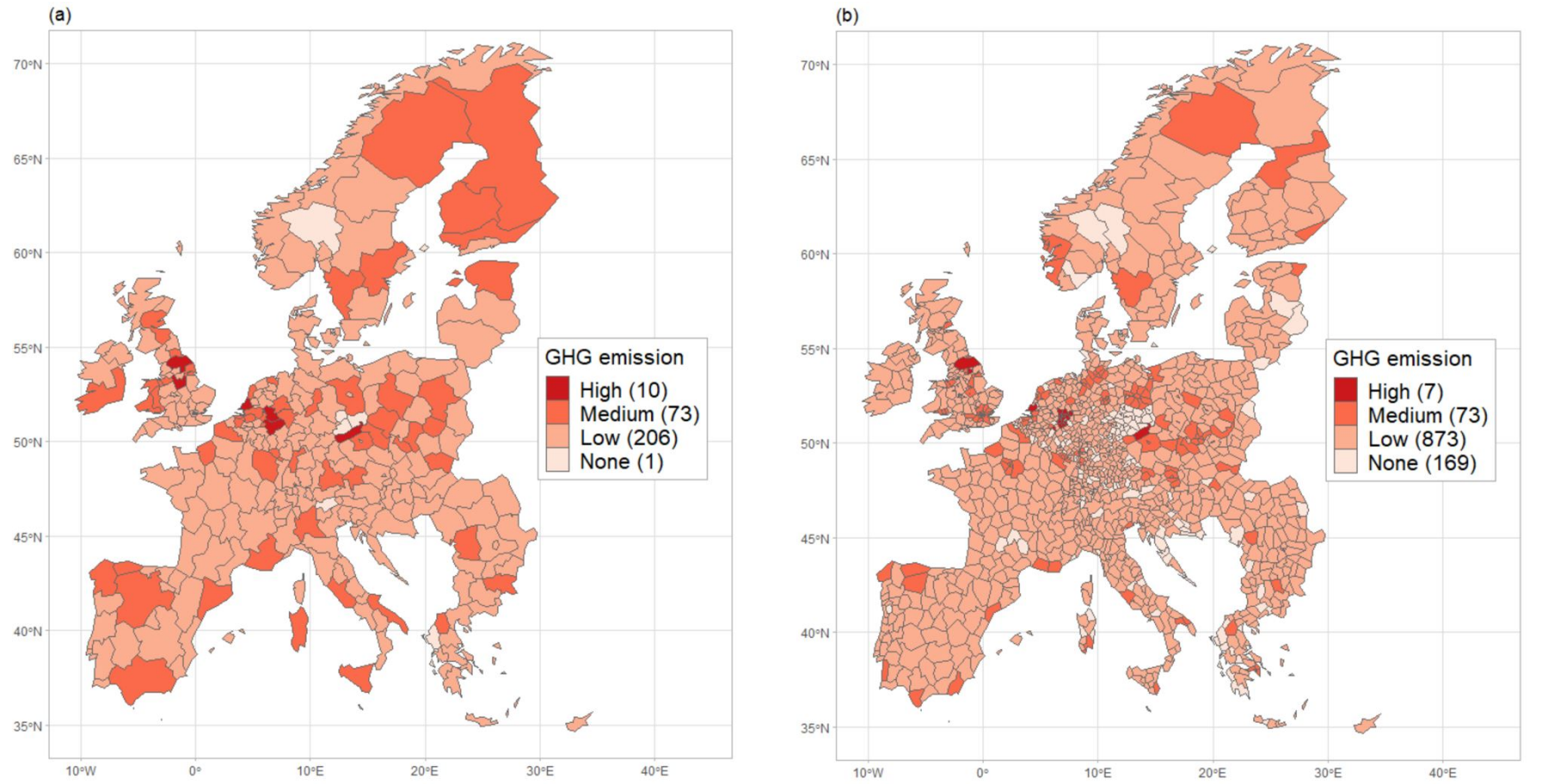
Patent activity, by NUTS2





Notes: Average number of digital (panel a) and green (panel b) patents per 100,000 inhabitants. Categories based on Jenks natural breaks optimization. Number of NUTS2 regions in parenthesis.

GHG emissions, by NUTS2 and metroregion



Notes: NUTS2 (panel a) and metro-regions (panel b) Categories based on Jenks natural breaks optimization. Number of NUTS2 / metro-regions in parenthesis.

Research agenda 1/2

- Taking stock of regional strengths and weaknesses for a European digital and green transition

  ✓ Where is scientific and technological excellence located in Europe today? What type of collaborations (e.g. international networks, public-private links) are at the basis of knowledge creation and dissemination? Which factors (including policy ones) can explain this geographical distribution?

- Upsurge (and decline) of digital and green knowledge poles/hubs

  ✓ Which factors explain the emergence of a region as a scientific and/or technological pole/hub? Are regional endowments in terms of physical and/or human capital able to explain this upsurge? Does the presence of DT knowledge hubs in certain regions facilitates the generation of environmental technologies of those regions?

Research agenda 2/2

- The influence of scientific and technological knowledge on GHG emissions
  - ✓ Which are the returns of digital and green scientific and technological knowledge on the regional environmental impact? Is the combination of DT and GT helping in reducing the environmental impacts of production and the related emissions of the regions in selected pollutants? And is it displaying any additional returns with respect to GT taken in isolation?

- Public demand, public funding and procurement for the digital and green transformation
  - ✓ Can public demand drive and foster the digital and green transformation? Which role for EU Horizon funding? And which role for public procurement in explaining such transition? Does the public funding lead to better quality outcomes?

Ongoing research

- Identify fine-grained cohesive sub-fields composing the digital and green knowledge fields, using text corpus of the title and abstract of publications and patents Backup
- Econometric regional analysis of the direct relationships between green and digital knowledge/technologies and GHG (production) emissions
- Issues: address reverse-causality in the econometric setting
- Possible econometric future extension: account for cross-regional interaction in a spatial setting (e.g., using geographical distance, co-patenting and co-publishing as possible weighting matrices)

Supplementary material

List of keywords related to digital scientific and technological knowledge

Keyword	Type	Keyword	Type
Deep learning	AI learn	Neuro-Linguistic Programming	AI other
Machine learning	AI learn	Object detection	AI other
Reinforcement learning	AI learn	Predictive modelling	AI other
Statistical learning	AI learn	Probabilistic modelling	AI other
Supervised learning	AI learn	Random Forest	AI other
Transfer Learning	AI learn	Self driv	AI other
Unsupervised learning	AI learn	Sentiment analysis	AI other
		Smart glasses	AI other
Artificial intelligence	AI other	Speech Recognition	AI other
Artificial intelligent	AI other	Unmanned Aerial Vehicle	AI other
Artificial reality	AI other	Unmanned aircraft system	AI other
Augmented realities	AI other	Virtual reality	AI other
Augmented reality	AI other	Voice recognition	AI other
Automatic classification	AI other		
Autonomous car	AI other	Internet of things	IOT
Autonomous vehicle	AI other		
Bayesian modelling	AI other	Big data	Infrastructure
Computational neuroscience	AI other	Edge Computing	Infrastructure
Computer Vision	AI other	Fog Computing	Infrastructure
Data mining	AI other	Hyper-Connectivity	Infrastructure
Data science	AI other	5G	Infrastructure
Decision tree	AI other		
Evolutionary Computation	AI other	Additive manufacturing	Additive manufacturing
Face recognition	AI other	3D printing	Additive manufacturing
Facial recognition	AI other	3D scanner	Additive manufacturing
Gesture recognition	AI other		
Holographic display	AI other	Humanoid robot	Robot
Knowledge Representation	AI other	Robotics	Robot
Machine intelligence	AI other		
Machine to machine	AI other	Industry 4.0	IV0
Mixed reality	AI other		
Natural Language Processing	AI other	Blockchain	Blockchain
Neural Network	AI other		

Notes: New keywords compared to Van Roy et al. (2020, Table 2) classification are highlighted in grey.

Supplementary material

(Customized) OECD Envtech classification, IPC and CPC codes

- ENVIRONMENTAL MANAGEMENT
  - AIR POLLUTION ABATEMENT
  - ~~WATER POLLUTION ABATEMENT~~
  - ~~WASTE MANAGEMENT~~
  - ~~SOIL REMEDIATION~~
  - ENVIRONMENTAL MONITORING
- ~~WATER-RELATED ADAPTATION TECHNOLOGIES~~
- CLIMATE CHANGE MITIGATION technologies related to ENERGY generation, transmission or distribution
- 4.1. RENEWABLE ENERGY GENERATION
- 4.2. ENERGY GENERATION FROM FUELS OF NON-FOSSIL ORIGIN
- 4.3. COMBUSTION TECHNOLOGIES WITH MITIGATION POTENTIAL (e.g. using fossil fuels, biomass, waste, etc.)
- 4.4. NUCLEAR ENERGY
- 4.5. EFFICIENCY IN ELECTRICAL POWER GENERATION, TRANSMISSION OR DISTRIBUTION
- 4.6. ENABLING TECHNOLOGIES IN ENERGY SECTOR 1
- 4.7. OTHER ENERGY CONVERSION OR MANAGEMENT SYSTEMS REDUCING GHG EMISSIONS
- CAPTURE, STORAGE, SEQUESTRATION OR DISPOSAL OF GREENHOUSE GASES
- 5.1. CO2 CAPTURE OR STORAGE (CCS)
- 5.2. CAPTURE OR DISPOSAL OF GREENHOUSE GASES OTHER THAN CARBON DIOXIDE (N2O, CH4, PFC, HFC, SF6)
- CLIMATE CHANGE MITIGATION technologies related to TRANSPORTATION
- CLIMATE CHANGE MITIGATION technologies related to BUILDINGS
- ~~CLIMATE CHANGE MITIGATION technologies related to WASTEWATER TREATMENT OR WASTE MANAGEMENT~~
- CLIMATE CHANGE MITIGATION technologies in the PRODUCTION OR PROCESSING OF GOODS

Supplementary material

Digital knowledge (word embedding - WoS corpus)

Robotics	Infrastructure	Machine Learning		Applications		
		Predictive analytics / Deep learning	**NLP / Speech recognition**	**Environment**	**Health**	**Industry 4.0**
Control	Network	Learning	Data_mining	Water	Patient	Manufacturing
Robot	Architecture	Algorithm	Content	Species	Clinical	Additive_manufacturing
Sensor	Cloud	Classification	Speech	Environmental	Treatment	3D_printing
Hand	Service	Features	Language	Urban	Surgery	Geometry
Motion	Communication	Training	Semantic	Soil	Disease	Fabrication
Virtual_reality	Mobile	Machine_learning	Text	Plant	Diagnostic	Defects
Body	Internet	Neural_network	Extraction	Satellite	Cancer	Manufactured
Controller	IoT	Datasets	Representation	Remote_sensing	Laparoscopic	Machining
Movement	Cloud_computing	Segmentation	Speech_recognition	Forest	Minimally_invasive	Selective_laser
Sensing	5G	ANN	Natural_language	UAV	Hospital	SLM
Arm	Protocols	Deep_learning	Words	Land	Surgeons	Implants
Augmented_reality	Connectivity	Support_vector	Ontology	Vegetation	Therapeutic	Rapid_prototyping
Localization	Edge	Time_series	Topic	Climate	Breast_cancer	Surface_roughness
Manipulation	Node	Decision_tree	Corpus	Plants	Magnetic_resonance	3D_printer
Orientation	Blockchain	Random_forest	Documents	Spatial_resolution	Robot_surgery	Manufacturing_process
Cameras	Mobile_devices	Cross_validation	Translation	Ecological	Risk_factors	Material_properties
Controllers	Antenna	Feature_selection	Automatic_speech	Land_cover	Postoperative	Fused_decomposition
Assistance	QOS	Convolutional_neural	Sentiment_analysis	Crop	Clinical_practice	Tissue_engineering
Kinematic	Routing	SVM	Twitter	Agricultural	Computed_tomography	Laser_beam
Humanoid_robot	Fog	Supervised_learning	Information_retrieval	Emissions	Computer_assisted	Packaging

Notes: Clusters of semantically related *n*-grams obtained via K-means on word embedding (Word2Vec). Optimal number of clusters determined through gap statistic.

Main

Supplementary material

Green knowledge (word embedding - WoS corpus)

Environmental Assessment & Policy	Emissions (Air)	Green chemistry	Waste and water treatment	Energy efficiency	Biotechnology
Environmental	Concentrations	Properties	Treatment	Energy	Biofilm
Management	Levels	Formation	Concentration	System	Biofilms
Global	Pollution	Reaction	Waste	Power	Bacterial
Assessment	Emission	Particles	Removal	Heat	Biofilm_formation
Climate_change	Air_pollution	Electrochemical	Degradation	Network	Microorganisms
Strategies	Pollutants	Oxidation	Wastewater_treatment	Voltage	Proteins
Implementation	Ozone	Oxygen	Wastewater	Fuel_cell	Enzymes
Sustainable	Collected	Catalyst	Organic	Load	Organisms
Scenarios	PM	Solid	Compounds	Electrical	Antimicrobial
Framework	Atmospheric	Adsorption	Treated	Battery	Cultures
Policy	PM2.5	Electrode	Separation	Grid	Tissue
Impacts	Air_quality	Synthesis	Treatments	Building	Genetic
Project	Contamination	Electrodes	Effluent	Energy_consumption	Diseases
Measures	PM10	Membranes	Organic_matter	Solar	Fungal
European	Trends	Catalysts	Recycling	Electric	Colonization
EU	Pollutant	Deposition	Waters	Heating	Microbiological
Indicators	NO2	Anode	Anaerobic	Batteries	Pathogen
Scenario	Particulate_matter	Electrolyte	Fermentation	Cooling	Skin
Challenges	Mortality	Reactions	Organic_material	Networks	Yeast
Policies	Environmental_monitoring	Catalytic	Filtration	Converter	Fungi

Notes: Clusters of semantically related *n*-grams obtained via K-means on word embedding (Word2Vec). Optimal number of clusters determined through gap statistic.

Main