

RISISWEEK AGENDA

January 25th -29th, 2021 – ZOOM Meeting

The RISIS week 2021 will entail: First 4 plenary sessions and a roundtable. It will mix demonstrations (demo sessions) and break-out sessions. Demonstrations will be one hour with questions and answers, break-out sessions are organised by working groups and will last 1,5 hours each.

	Monday 25	Tuesday 26	Wednesday 27	Thursday 28	Friday 29
9					
10	Overall view of RISIS developments	Overview of dataset developments	Interactions with colleagues	Indicator development	Roundtable on publication resources
11	Demonstration of RCF	SIPER developments	Firmreg session	Datasets on national projects & programmes	Overton demo and discussion
12	LUNCH				
13	Firmreg demonstration	Pharma biotech integrated use case	OpenAire demonstration	Demo on machine learning regression & classification	Closing session Take away messages
14	BREAK				
15	DCC breakout session	Session on Knowmak Ontologies	Orgreg session	RCF integrated use case	
16	Demonstration on Trademarks	BREAK	BREAK	Demo: Characterising AI governance	
		Café session	Café talk about sovereignty		
17				Café session	

MONDAY 25/01

10h00-11h00 Overall view of RISIS developments

Convenor: Philippe Laredo. This session is focused on the state of affairs of RISIS2 and the expectations that come from the fact we are supported as an ‘advanced research community’ research infrastructure.

11h00-12h30 Demonstration of RCF

Convenors: Philippe Breucker, Cristian Martinez, Marc Barbier. This session proposes to discover the RISIS Core Facility (RCF) interface. The introduction of this session will recall the trajectory and the principles of architecturing this Facility. The design of the interface will be demonstrated with a user point of view, explaining the various structural components of the interface. Two modes of uses are proposed, one assisted the other one for more advanced users. The notion of « RCF project » will be explained in light with the connection to RISIS dataset portal and RISIS services. Thereby, the presentation of the associated scenario and mock-up will demonstrate how these different tools can be combined.

13h30-14h15 FirmReg demonstration

Convenors: Roberto Camerini & Massimiliano Guerini. The aim of the new redesigned FirmReg structure and perimeter is twofold. First, as integrative element of the RISIS infrastructure, FirmReg allows not only to match all firms included in firm-level datasets, but it also offers the possibility for other non-firm level dataset owners to match other firms included in their non-firm level datasets. Second, it has the potential to become a point of reference to download lists of “top firms” in Europe based on indicators and thresholds applied to RISIS datasets. The demo session aims at explaining the key functionalities of FirmReg to RISIS partners. We will show usage examples, some useful indicators what we can quickly build using this data, and some geographical maps that will show how easily these indicators can be matched with other tools such as Knowmak. The visual representations can provide some interesting insights concerning policy-relevant questions like: How many top R&D firms are located in a certain region? What are the most active entrepreneurial ecosystems in Europe? How are these ecosystems evolving?

14h45-16h15 DDC breakout session

Convenor: Eric Iversen. We lack good (granular, cross-country) metrics for ‘PhD production’ while metrics for ‘labor-market placement of trained PhDs’ are fragmentary at best. DDC addresses the need for new metrics at this important policy intersection by developing a dataset that centers on dissertation-based information but that links in complementary information from other RISIS-resources (ETER Gate, RCF/CorText, and CWTS Publication). The work program consists of a preliminary and a practical phase. Preliminary work has conceptualized the labor-market placement of trained PhDs from a theoretical point of view, leading to a Wiki. The scope to extend current approaches has also been explored. Centering on the ongoing Careers of Doctorate Holders (Light) work, the potential for extensions appears greatest for countries that have register-based approaches (e.g. Norway, but also Austria and Netherlands) and that make them available to the scientific and policy communities. The WP showcases the potential of this approach to map non-academic carriers based on an example. The second phase of work has been dedicated to the practical work of designing and implementing a pilot dataset. Focused data from six countries have been assembled to elaborate the 4-step DDC’s workflow. This Session will review developments made and open a discussion for future directions based upon the pilot made for five (6) countries and results obtained on a sample.

16h15 - 17h15 Demonstration on Trademarks

Convenor: Rainer Frietsch. In the session, we will introduce the two trademark databases - EUIPO and USPTO -, with a strong focus on the former one. The session will offer results and analyses based on the FirmReg-Matching, but also provide data at the national level. Results and trends based on the newly developed classification of product and technology fields will be shown. In addition, trademark data was linked to patent data in two ways. On the one hand, the FirmReg-Matching allows data analyses at the micro level. Based on this, company profiles and correlations with patents will be analysed. On the other hand, a concordance between the trademark classification of products and technologies with the WIPO classification of patents was established. This is used for analyses of national trends and profiles. In addition, data access and ready-to-use tables on trademarks by country, year and WIPO classes will be introduced.

TUESDAY 26/01

9h30-11h00 Overview of dataset developments

Convenor: Thomas Scherngell. This session will, first, summarize the main developments and advancements of RISIS core datasets (datasets maintenance and integration for indicators production, datasets deepening to respond to upcoming questions), and the project-based access to these datasets via the RISIS datasets portal. In this context, we will summarize accesses in the first two years of RISIS (WP8), and briefly highlight new developments in the portal (e.g. multiple accesses).

Second, the session intends to provide a compact overview on new datasets developments (WP10), comprising the European Social Innovation Database (ESID), the Trademarks database, the database on national R&D programmes and instruments (EFIL) and on careers of PhDs based on CV information (DCC). Each of these four new datasets will provide a short overview on the status of its development, the workplan, integration potentials with other RISIS datasets and the perspectives for usage and opening.

11h00-12h30 SIPER developments

Convenor: Susanne Buhner-Topcu. SIPER is a unique open access data base of evaluation studies in the field of R&I policy. Evaluations are conducted to assess the effects of policy instruments. Each evaluation on its own but also analysis across evaluations constitute a valuable resource for policy learning. However, evaluations are published in a large variety of locations and are not all readily traceable. This is where SIPER aims to make a difference through offering access in one place (www.Siper.eu) and thus allowing a meta-analysis. The aim of this breakout session is to present and discuss last year's development of SIPER. This includes a) the launch of the new Website, b) an overview of the recent number of collected and coded evaluation studies, c) the presentation of a new approach to automatically code evaluations, d) the ontology that was developed to assess the performance of evaluated measures. The session addresses researchers conducting and/or using evaluations and will help to refine the upcoming works.

13h30-14h15 Pharma biotech integrated use case

Convenors: Antoine Schoen, Patricia Laurens, Alfredo Yegros & Thomas Scherngell. This research project mobilises data from RISIS Patent, Leiden Ranking, EUPRO and tools and methods from Cortext (for geolocating activities).

Research focus: To which extent do corporate publications (i.e. publications with an address from a large multinational firm) foster corporate patent activity? We look first at the role of corporate publications in the overall publications of metropolitan areas, and then at the relationships with corporate patents. The focus is on urban and rural areas (URA, mostly metropolitan areas) covering the US and Europe: 3155 URAs out of which around 1000 are considered 'active' (more than 1 corporate publication per year). The corporate population is made of 631 CIB firms from the pharma and chem sectors, for which we have identified about 550 000 publications (over 20 year from 2000 to 2019) and about 600 000 patents. URA are qualified by their overall publications, their chem pharma specialisation, network characteristics (URAs centrality in the EU projects – with a first focus on Europe) and a spatial matrix with all geodesic distances between URA. The Spatial Autoregressive model (SAR model) will be used for analysing the territories' patent output.

14h45-16h15 Session on Knowmak Ontologies

Convenor: Benedetto Lepori. The session will discuss the future develop of the KNOWMAK ontology and its alignment with new-KETs and Sustainable Development Goals. The revision work undertaken so far, and the overall revision strategy will be presented. Moreover, preliminary results on the SDG 'Sustainable Cities and Communities' will be presented and validated with the participants. Finally, the session will be devoted to organizing further work in this area.

16h30 Café Session

Open discussion of shared matters of interest, emerging bottom-up

WEDNESDAY 27

09h30-11h00 Interactions with colleagues

Convenor: Emanuela Reale. The future of RISIS will depend on the extent to which we interact with colleagues and demonstrate our value added to stakeholders. The session will review all the means through which we interact with colleagues to better see their complementarities and enhance their synergies: research accesses/research seminars/Zenodo (new additions, downloading)/training activities/advanced methodological seminars. This will include a review of our Interactions with wider public (colleagues from other fields, policy analysts, policymakers): data online access (ORGREG/KNOWMAK for the time being); policy sessions; information via website and social media.

11h00-12h30 FirmReg session

Convenor: Massimiliano Guerini. The breakout session will deal with the technicalities of FirmReg. During the last year SPRU has redesigned the technical structure of FirmReg, in terms of perimeter and tables. This was done in order to accommodate all the novelties that will be introduced in the new release of FirmReg, namely: 1) the 'backend': which includes the subsidiaries owned by each firm in the FirmReg and their ultimate owners. All this information will help other partners to match their firms with FirmReg; 2) the 'thresholds': a new feature of FirmReg that allows users to access to top firms in Europe according to a number of indicators. This new structure was done in collaboration with POLIMI and then validated with Joanneu. The session aims at clarifying three important points in order to implement the register during 2021: 1) definition of use cases in order to detail the technical functionalities and requirements for the online interface of FirmReg; 2) access rights and how to set them in the on-line interface; 3) validation of the update protocol.

13h30-14h15 OpenAIRE demonstration

Convenor: Paolo Manghi. This session will present two EOSC services operated by the OpenAIRE infrastructure:

- The OpenAIRE Research Graph, one of the largest metadata collections of scientific products (publications, data, software) worldwide (graph.openaire.eu). We will first describe how the graph is constructed and then short demo its possible usages and applications supported on top, from discovery across disciplines (explore.openaire.eu), by discipline/community (connect.openaire.eu), to monitoring (monitor.openaire.eu).
- The OpenAIRE Research Community Dashboard, providing Community Gateways for research infrastructures (communities) on top of the OpenAIRE Research Graph. OpenAIRE Community Gateways offer tools to identify community-related products and tools to search/discovery and monitor statistics in the given domain (connect.openaire.eu). In particular, the demo will showcase the RISIS Community Gateway, how it is constructed and how the community can set the criteria driving its content.

14h45-16h15 OrgReg session

Convenor: Benedetto Lepori. This session will build the ground for the development of the OrgReg3.0 version and, specifically, of the new indicator modules connected with KNOWMAK. Following main points will be discussed: Review of the work done so far (launch of the OrgReg2.0 version); The selection of the indicators to be provided in OrgReg; The design of the user interfaces; First results of the preliminary analysis of data and discussion of the issues to be addressed; The revision of the OrgReg handbook; The plan for dissemination and promotion; The working plan for 2021.

16h30 Café talk about sovereignty

Convenor: Rainer Frietsch. This topic of (technological and also digital) sovereignty is high on the agendas of many national policy makers, but also of stakeholders at the EU and the OECD level - but there is still a need for conceptual clarity and empirical facts. The intention of this café session at the RISIS week is to openly discuss the possible conceptual and especially empirical contributions of RISIS (and its datasets) to the current European discussion on technology sovereignty. As an impulse to this discussion we will have a short introduction of a paper published by Fraunhofer ISI in late 2020; however, we will only take this as a starting point from where we will jointly try to define sovereignty in our context and operationalise this definition with means and perspectives by the RISIS research community and the RISIS datasets. This café session is to be seen as the initiation of a process that equips us with the tools to write a RISIS policy brief on technology sovereignty, to hold a policy discussion before the end of the year, and maybe to develop a regular analytical tool for Europe. For those who are interested in the ISI-paper, you can find an abstract, a short version (executive summary) and the full version here: <https://www.isi.fraunhofer.de/en/presse/2020/presseinfo-11-Technologiesouveraenitaet.html>

THURSDAY 28

9h30-11h00 Indicators development PLENARY

Convenor: Benedetto Lepori. This session will present for discussion the main directions for an indicator program within RISIS. A preliminary overview of the indicators made available by the RISIS datasets will be presented. On this basis, a structuration of the indicators by analytical level will be presented and areas for further integration within RISIS will be proposed. Finally, the session will specifically present the new indicator module on research organizations and proposed work around the integration of regional indicators.

11h00-12h30 Datasets on national projects & programmes

Convenors: Emanuela Reale, Thomas Scherngell. The exploration of project-based R&D activities funded by public research funding organisations has become a core research issue in science and innovation policy studies. RISIS responds to this demand from the research community but also policy makers by maintaining and developing datasets on project-based R&D (EUPRO) and R&D funding programmes across European countries (EFIL). While EUPRO shifts – within the new NATPRO module – attention to the collection and integration of national R&D projects as an important complement to the existing European ones, EFIL is a completely new dataset currently under development collecting data on national R&D policy programmes at the level of policy instruments. This session will discuss progress, planned activities and problems related to EFIL and the new NATPRO module of EUPRO. Emphasis will be shifted to the results from the pilot phase and the data collected so far on the pilot countries, with specific attention to the situation in the new EU member states that have been taken into account from the beginning. Finally, integration issues between NATPRO and EFIL as well as other datasets will be discussed, as well as planned next steps including an opening perspective.

13h30-14h15 Demo on machine learning regression & classification

Convenor: Giovanni Cerulli, CNR. Using the Stata/Python full integration available in Stata 16, I recently programmed two Stata commands - R_ML_STATA and C_ML_STATA - for carrying out machine learning regression and classification. These routines are higher-level wrappers of the popular Python platform Scikit-learn for machine learning computation.

I propose a DEMO of these two commands applied to the merging of the two datasets EUPRO (for EU-FP projects) and ETER (for HEIs). I will show how to obtain predictions and feature importance indices in a series of applications on this database. I will discuss the potential of comparing/exploiting different learners and/or their ensembles to predict relevant STI outcomes starting from selected inputs. I will embed the applications within the “meta-learning” architecture I recently presented in Cerulli (2020).

14h45-16h00 RCF integrated use case

Convenors: François Perruchas, Philippe Breucker, Marc Barbier. The objective of the session is to present the first RCF integrated scenario with a known use case, that we have generalized to provide a tool to the RISIS research community, and to demonstrate the capability of the new RCF-beta integrated platform. Following the work of François Perruchas during his PhD [<https://doi.org/10.1016/j.respol.2020.103922>], we developed an analysis scenario for RCF around patent analysis. The goal of the scenario is to generate and analyse empirical evidence on the distribution of environmental patenting activity by countries, over time and by domain of specialization. Starting from the selection of IPC/CPC patent classes, the scenario consists on a set of data integration and analysis tasks, and data visualization output, available in the RCF user workspace. Following the experience of the CorText Platform's users (see <https://www.cortext.net/publications/>), we propose to demonstrate a scenario of Corpus scientometrics analysis based on regular scripts already in use within the Corpus Manager interface (bibliographic field frequential analysis, network mapping, contingency matrix). We shall propose to run this scenario on two WoS and Scopus corpora.

16h15-17h00 Demo: Characterising AI governance

Convenor: Bilel Benbouzid (UGE). AI-REG is a new project exploring the emerging governance of artificial intelligence regulation. We start this exploration from European Union institutions, but we integrate also every place where positions confront each other, where irreversibilities are constructed and where market 'frameworks' are gradually being set; all different kind of actors contributing toward the formation of the rules of engagement that constitutes AI domains. AI-REG uses quantitative methods (text mining and social network analysis mainly) to explore the emergent 'technological zone' of AI. The AI-REG dataset will provide curated data on different actors, places and resulting networks of AI governance at worldwide level. It aims at building a new RISIS resource to assemble content and citation network of heterogeneous documents from academia, policy consulting, legislative body, administration and business. This dataset contributes to STI research by focusing on the dynamics by which knowledge about governance of AI can be “characterized as balancing tentative and definitive forms of governance ” ([Kuhlmann, Stegmaier](#) and Konrad 2019).

17h00 Café Session

Open discussion of shared matters of interest, emerging bottom-up

FRIDAY 29/01

09h30-11h00 Roundtable on publication resources

Convenor: Ed Noyons.

Multiple new sources have been or are being developed and we need to discuss how to position ourselves and which resources are better fit for different problems. We also need to keep in mind the need for one core stable resource for indicator building that will be as all other RISIS resources accessed at a distance through the RCF. We are also in search of new resources on policy documents that could help in the design of problem-based ontologies.

11h00-12h30 OVERTON demo and discussion

Convenors: Ed Noyons, Euan Adie, Zichou Fang, Jonathan Dudek, Alfredo Yegros and Juan Pablo Bascur

Demonstration of the possibilities offered by OVERTON and a discussion about potential collaborations.

13h30-14h15 Closing session

Discussion on the take away messages' that will come from each session.

