





## Introduction

RISIS2 aims at building a data and services infrastructure supporting the development of a new generation of analyses and indicators.

This project builds on RISIS1 (2014-18), which has demonstrated the relevance of such an approach and opened access to a first set of databases and services.

RISIS2 gathers 19 partners aiming to transform the field of STI studies into an advanced research community.

This step change is achieved by:

- developing an e-infrastructure that supports full virtual transnational access by researchers,
- providing a vastly enlarged set of services tailored to field-specific needs,
- maintaining datasets dealing with firm innovation capacities, public research developments, R&I outputs and projects, and policy learning,
- developing new datasets on 4 key issues for research and policy (social innovation, non-technological innovation, the role of PhDs in society, portfolios of public funding instruments).

## Old and new datasets

RISIS contains 12 datasets and RISIS 2 will integrate it with 4 new datasets dealing with important but lasting research and policy issues. The 16 datasets are on the following themes:

**Firm innovation dynamics:** 3 datasets concerning world largest firms (CIB), European venture capital backed firms (VICO) and fast growing mid-sized firms (CHEETAH). They will be complemented by 2 datasets on social innovation and on trademarks.

**European integration:** 2 datasets dealing with European projects and trans-border programmes (EUPRO and JOREP) and European public research actors (with ORGREG serving as a reference database incorporating the ETER DB on universities, PROs and university hospitals).

**Knowledge dynamics:** output datasets enriched for analysis at actor and metropolitan areas levels. Publications with the WOS Leiden-ranking and Patents and PATSTAT-IFRIS (get rid of accessible through CIB), with a demonstrator dataset for the study of emerging technologies (Nano S&T DB).

**PhD and researcher careers and mobility:** PROFILE and MORE. They will be complemented by a dataset on non academic PhD careers.

**Policy learning:** repository on science and innovation policy evaluations (SIPER) linked to the OECD-World Bank IPP platform. It will be complemented by a dataset on effective portfolios of public funding instruments on research and development.

All together they provide resources for truly new indicator developments, individually and more and more by combining them on problem, space or organization-based approaches. It will avoid having to operate massive simplifications at aggregate level.

## Distant access

We add the possibility of virtual transnational access to the physical access with the opening of the **RISIS Core Facility (RCF)**.

This change will not only simplify access for users independently of their geographical location.

It will also radically enlarge the 'user base' as the research topics that researchers can address will be multiplied through the simultaneous access to multiple datasets and the possibilities of combination offered by our data integration services.

We will connect the community to more and more open data enabling far wider exploitation via the opening to the growing number of open access datasets thanks to the D4Science capabilities articulated with the RCF.

## Freely accessible services

Three data integration services provide a new resource for combining datasets both RISIS datasets and external datasets.

All are open online for all 'accredited' researchers for their own corpuses as well as when using RISIS core datasets.

RISIS 2 include the two actor-based registers of "research performers": **ORGREG** for public research (higher education institutions, public research organizations, research hospitals, research performing NGOs and research funding organizations); and **FIRMREG**.

They are key supports for our integration services and their maintenance and development are strategic for data integration.

RISIS 2 will offer a completely new service based on topic identification (though the ontology tagging service of **GATE**).

This is a major addition because of the rising research interest in analysis at the level of targeted topics not recognisable through existing classifications, for instance in the case of possible future policy missions pursued by European framework programmes.

We extend the realm of semantic analysis tools we offer through **CORTEXT** (based on corpuses) with a newly opened corpus geocoding facility, and via a greater range of visualization services (in particular VosViewer).

