DIGITAL SCIENCE AND INNOVATION POLICY AND GOVERNANCE (DSIP)

Project update

CSTP meeting, October 2017

Michael Keenan, Dmitry Plekhanov (STI/STP),
Fernando Galindo-Rueda, Daniel Ker (STI/EAS)

OECD
BETTER POLICIES FOR BETTER LIVES

Context

- STI activities, including their planning, performance, and outputs, increasingly leave "**digital footprints**"
- The digital footprint of science emerging as a growing component of the **statistical infrastructure** on STI - featured prominently at ***BlueSkyIII***

- **Datasets** becoming larger, more complex, and available at higher speed
- These datasets can be **analysed** – using e.g. big data analytics – to *complement traditional policy data sources i.e. official R&D and inno. data*

➢ Country-level initiatives incorporating and analysing **data** on individual researchers, research organisations, project/grant awards, and research outputs

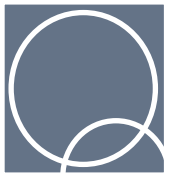
- Private sector also active – especially large academic publishers exploiting their data to develop tools for research funders and performers

Definition of DSIP initiatives

DSIP initiatives refer to

- the adoption or implementation by public administrations, with responsibilities for science and innovation,
- of new (or re-used) practices, requiring an intensive and novel use of digital technologies and/or data resources,
- with the aim of supporting the formulation of science and innovation policy or the delivery of public services in this area.

Promises of DSIP initiatives



Anticipatory intelligence:
Identify emerging research areas, technologies, industries, policy issues, etc.

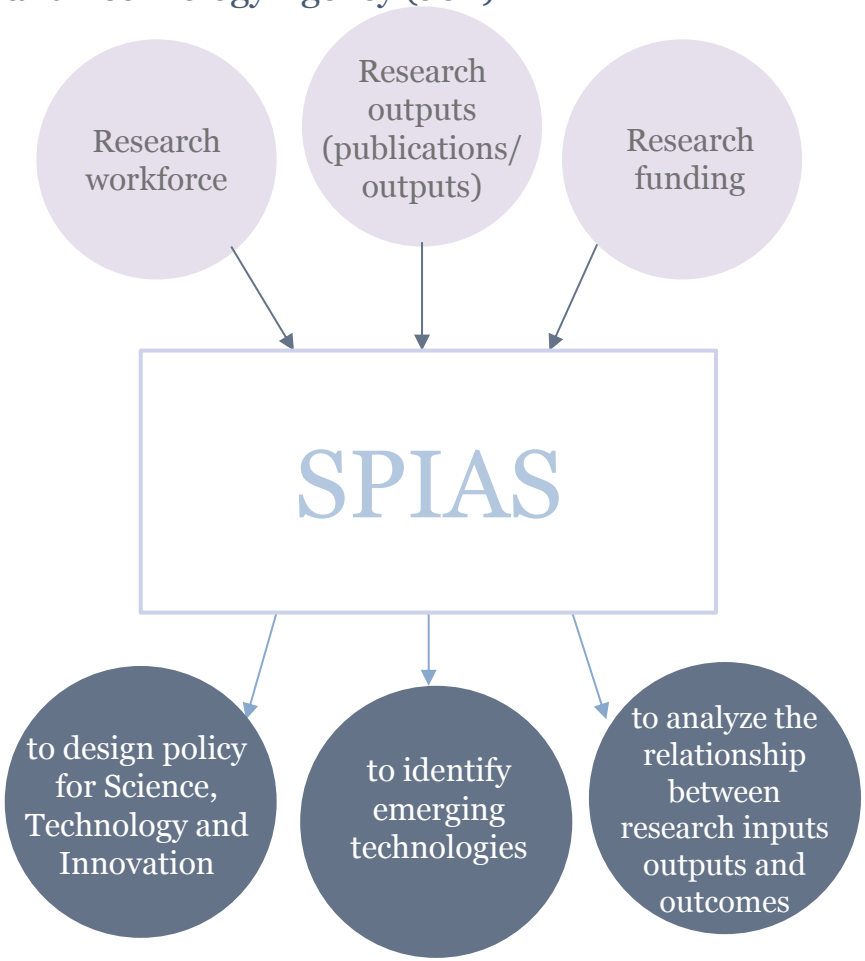
***Policy formulation and design*:**
more granular analysis → contextual variations in policy → boost policy effectiveness.

***Performance monitoring and management*:**
possibility to collate real time data →improved insight into policy impacts → more agile policy adjustments

***Optimisation of administrative workflows*:**
digital tools → streamline administrative procedures → efficiency gains within agencies and their various “customers”.

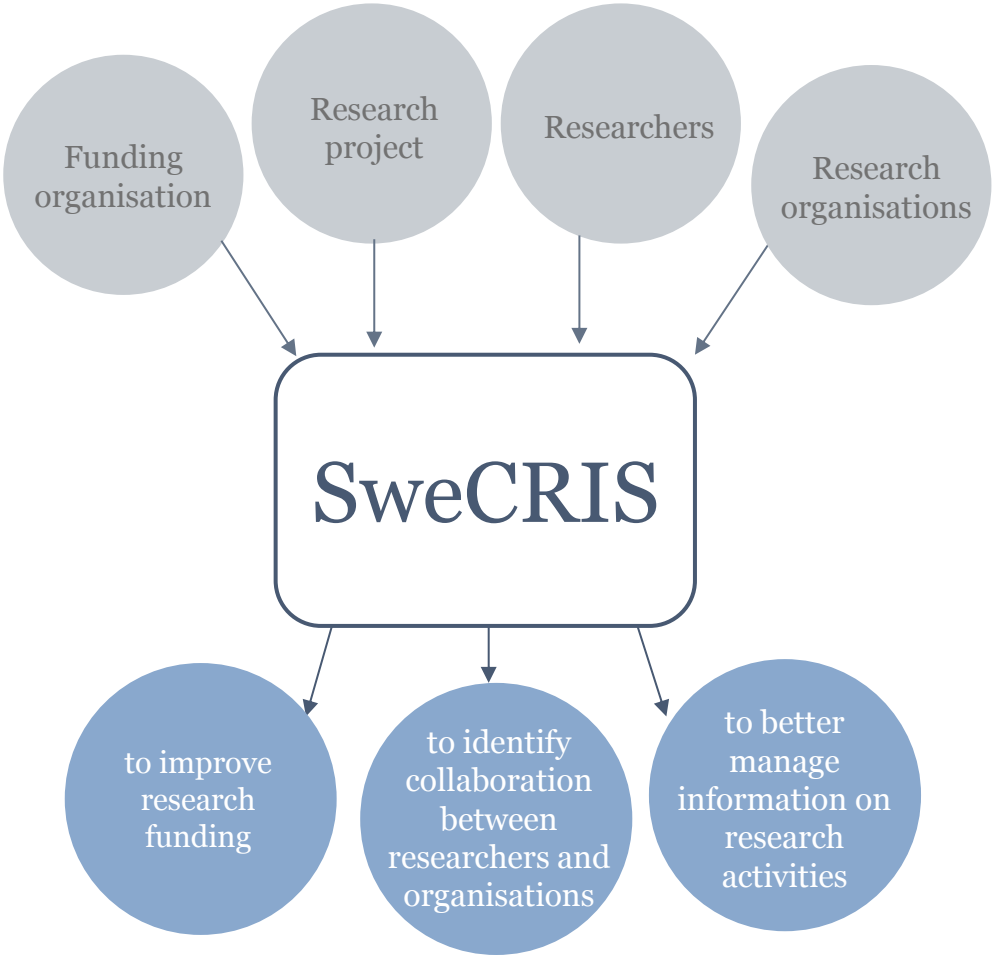
DSIP initiative in Japan

SPIAS (SciREX Policymaking Intelligent Assistant System), a digital system created in cooperation of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), National Institute of Science and Technology policy (NISTEP), National Graduate Institute for Policy Studies (GRIPS), Japan Science and Technology Agency (JST)



DSIP initiative in Sweden

SweCRIS, a digital platform hosted by the Swedish Research Council



CSTP project

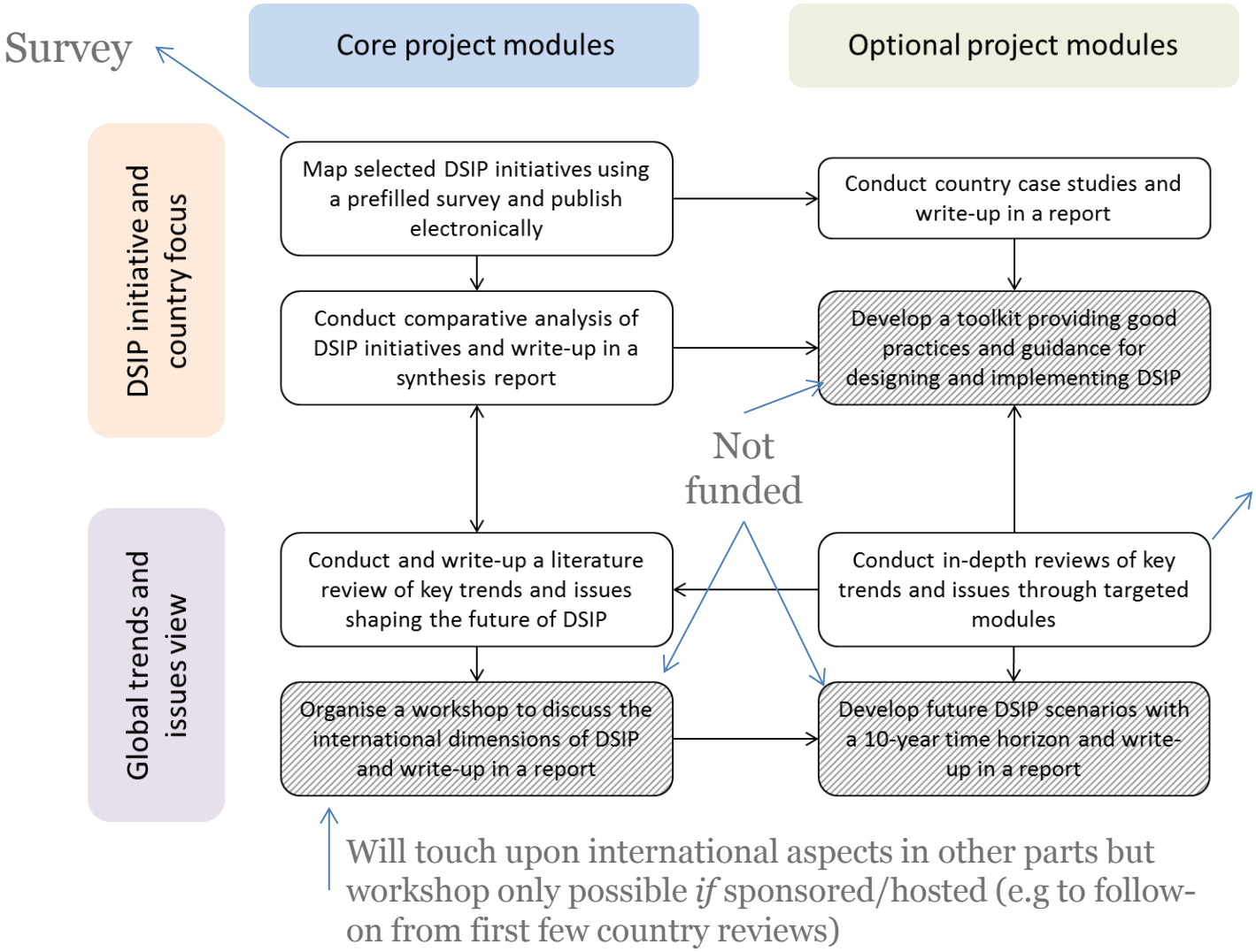
The objective of the project is to **provide CSTP with means to make an informed assessment of the transformational potential and possible pitfalls of DSIP**:

- What should STI policy makers and administrators know about new ways of using data about science and innovation?
- How can they influence this environment?

Why is this important?

- digital information-based content and processes starting to **transform** the evidence base for S&I policy design, operational delivery, and governance arrangements
- DSIP was identified at Blue Sky III as a major driver of current and **future statistical infrastructure** on S&I

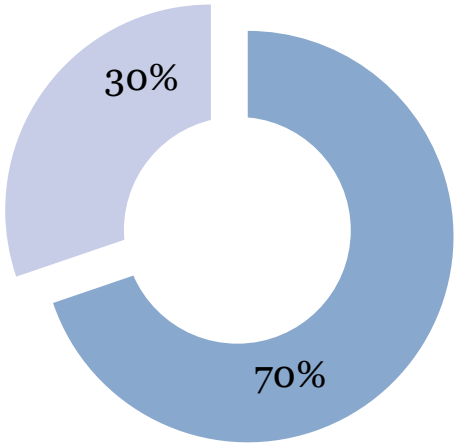
DSIP project modules



Final report – proposed chapters

1. Overview of DSIP
2. Results of survey of DSIP initiatives
3. Meeting the interoperability challenge
4. DSIP and the future of research assessment
5. The private sector in DSIP

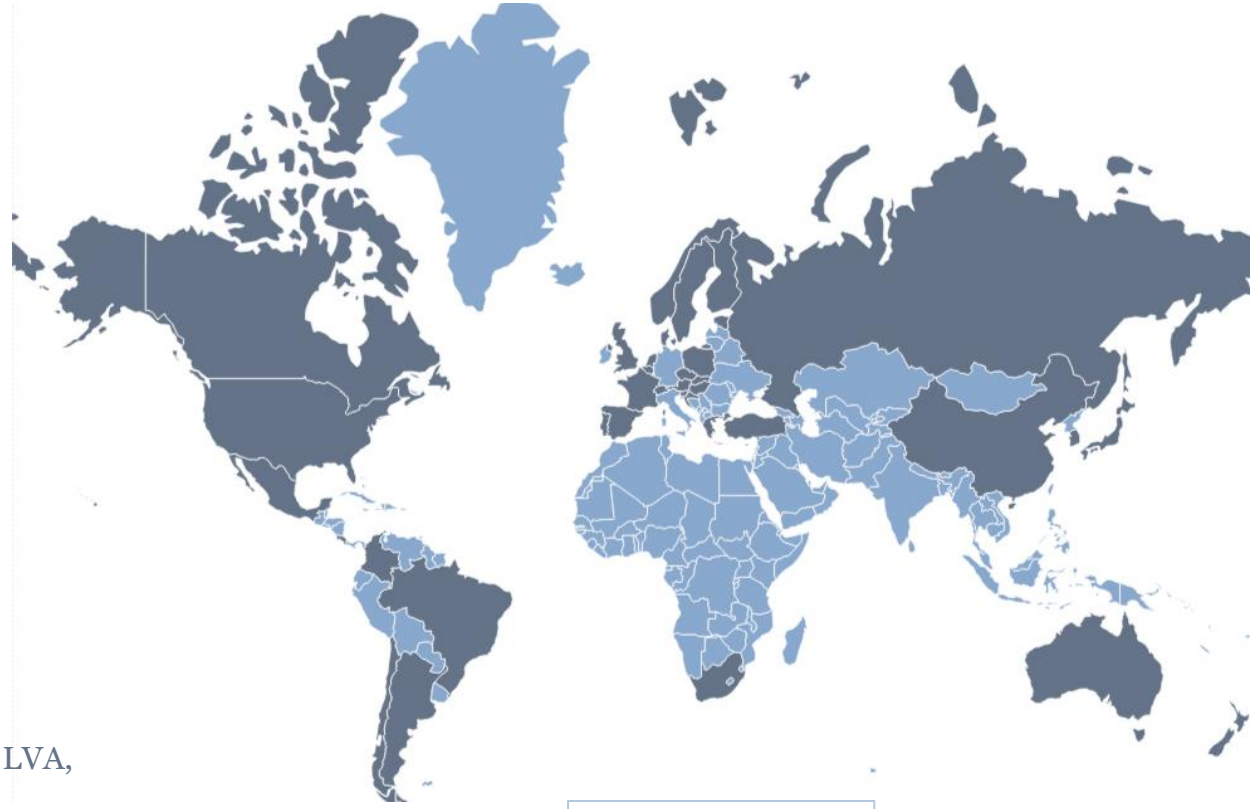
Mapping of DSIP initiatives



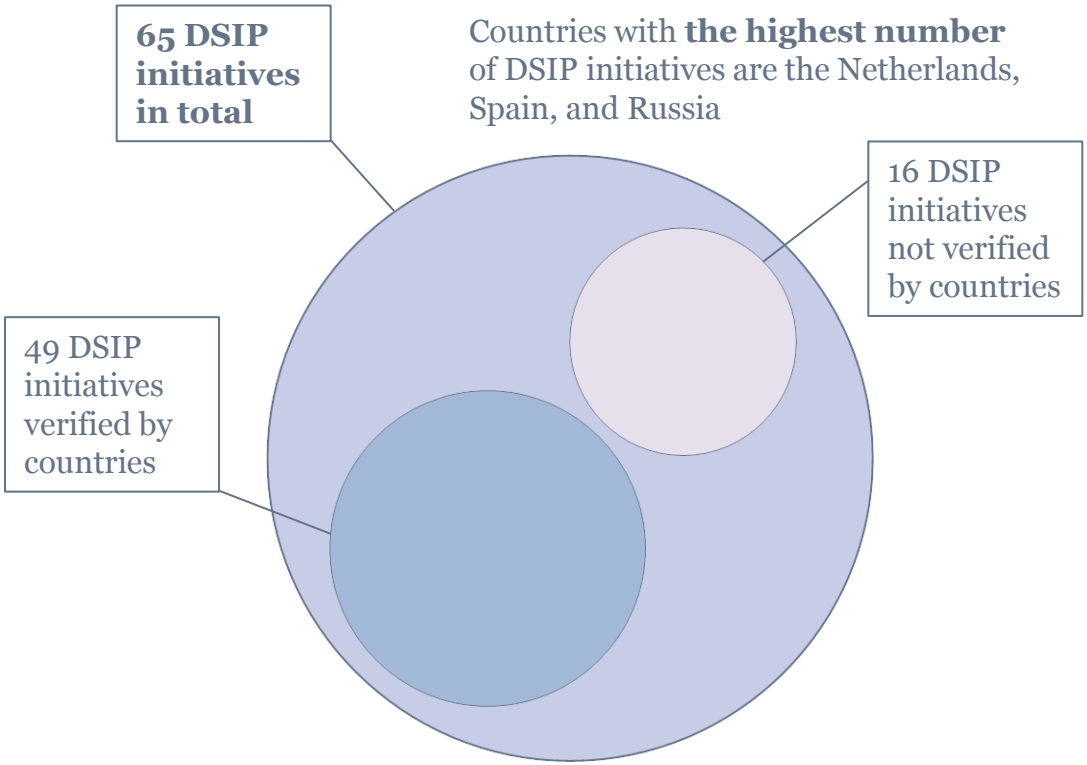
- Contries that responded
- Countries that did not respond

(CZE, FRA, GBR, HUN, ISL, ISR, ITA, LVA, LUX, TUR, POL, SVN, SVK)

Countries with known DSIP initiatives



Mapping of DSIP initiatives



Most frequent types of data analysed by identified DSIP initiatives: research outputs, human resources, funding, organisations, and research projects

Survey

- **Goal:** to explore rationales for developing DSIP initiatives, sources of funding, sustainability models, stages of development, technological architecture, challenges and problems impeding design and implementation of DSIP initiatives.
- **Building blocks:** online questionnaires and telephone interviews
- **Coordinators of 65 DSIP initiatives** will be invited to participate in the survey in **late 2017 - early 2018**.

1. Study the mechanisms and main rationales behind the emergence of DSIP initiatives (e.g. systems of incentives, the role of the private sector)
2. Map functions of DSIP initiatives
3. Identify main challenges impeding further development and implementation of DSIP initiatives and the approaches of coordinators of digital systems to solve them
4. Study funding schemes and sustainability models of DSIP initiatives

Private sector roles in DSIP initiatives

Clarivate Analytics

DIGITAL science

Google

Microsoft

NAVER

Baidu 百度

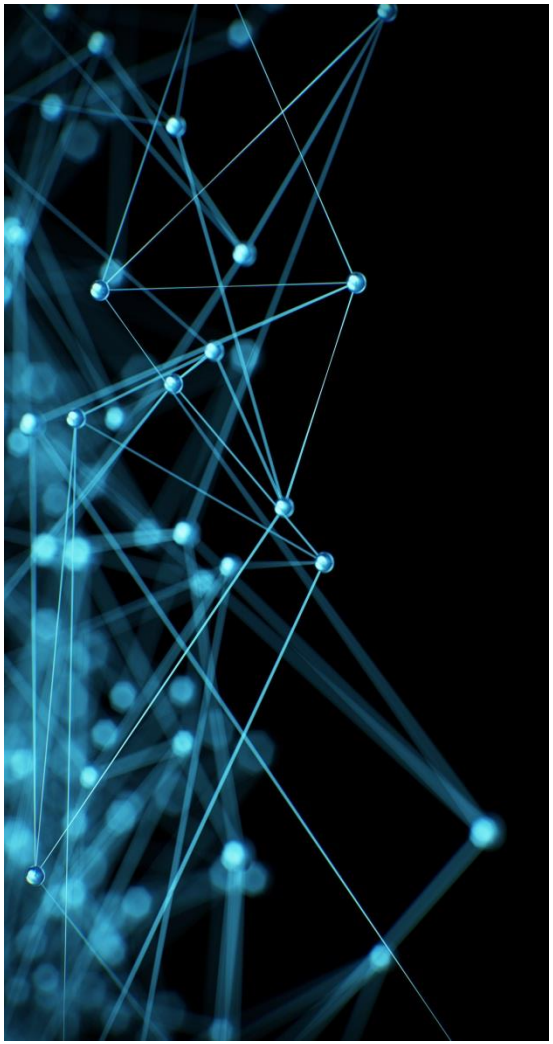
ResearchGate

Alfred P. Sloan Foundation

Bill & Melinda Gates foundation

- **Private sector** and **public sector** jointly develop new methods for analysis of research impact, design and implement digital tools for policy-making purposes.
- **Main players:** academic publishers (Holtzbrink Publishing Group), information analytics companies (Elsevier, Clarivate Analytics), and technology companies (Google, Microsoft), social networks for academics (ResearchGate, Academia.edu), charities (Alfred P. Sloan Foundation)
- **Benefits:** Provision of off-the-shelf, full-fledged digital solutions and building blocks for DSIP
- **Potential pitfalls:** discriminatory access to data, overreliance on proprietary solutions, biases and inaccuracies in data analysis

DSIP initiatives in research assessment



- DSIP:

→changes in monitoring, assessment, and evaluation of research activities

→Informed decisions on research funding allocation

- In-depth review focussing on:
  - "altmetrics"
  - potential and challenges of DSIP approaches
  - limits and pitfalls of digitalisation

Unique, Persistent, and Pervasive Identifiers (UPPIs)

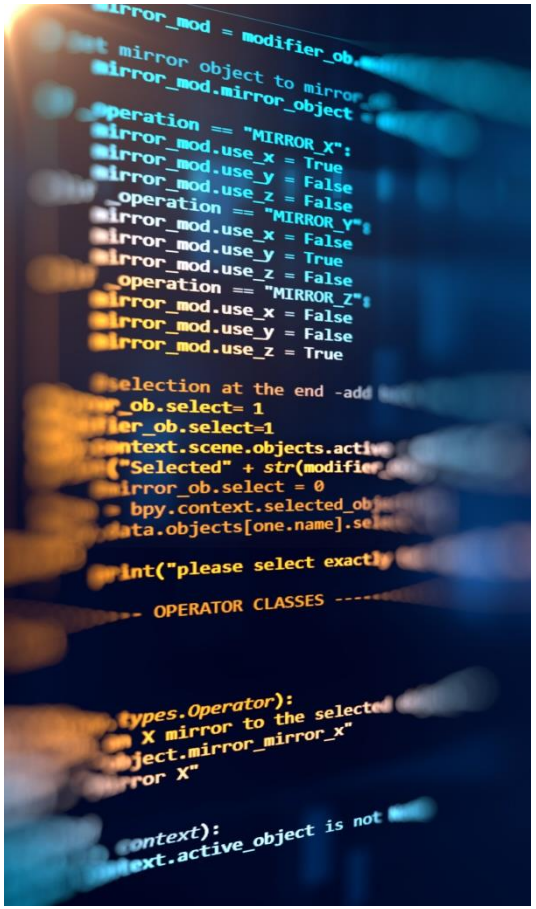
- NSOs have long used systems of unique identifiers
  - Often shared across multiple statistical domains e.g. via a central business register

  →Allow different datasets to be compared and combined
- With digitalisation, many other organisations now hold large, and relevant, datasets
  - Not just government administrative data; businesses, universities, etc. too
- These data can unlock powerful new analyses which complement existing NESTI work and support policymaking
  - For example, R&D survey tells us a company received gov R&D funds…
  - Personnel and grant administration data might allow us to identify/verify grant vs contract funds *and* to learn what kind of R&D personnel time was charged against the grants

→Need to be able to match data relating to same entities across a wide range of data sources

→ potential for systems of identifiers that go beyond NSOs/Government

= **Pervasive** as well as **unique** and **persistent**



An example: ISNI

- ISO standard 22729
- Register operated by an independent, UK-based, non-profit institution
- 9.5m (+counting) individuals and organisations contributing to “creative works”
  - includes 2.6m researchers
  - 600 000 organisations (e.g. 0000 0004 1936 8948 Oxford University)
  - Relational information e.g. associated companies, employee←→employer
  - Global scope
- Openly available for use – API
- Funded through registration fees
- **ORCIDs are ISNIs – and allocated free of charge**

→ various other systems, many proprietary/commercial (primarily organisational identifiers)

UPPIs – not without challenges

Most notable is **coverage**

Especially in "opt-in" systems

→ scope for policy to help drive coverage…if sufficient support

e.g. by requiring an entity to have an UPPI in order to get funding, qualify for tax credits, etc.

Also quality:

- Verification of claims, relational information, etc.
- IDs applied at persistent organisational level (e.g. enterprise)
- Managing changes in organisational structure

→ NSOs have experience in managing such issues, establishing systems of identifiers which work.

Case study of Norway

The case study will explore what improvements could be made in termsof

- **interoperability** and connecting different data sources,
- facilitating **international comparisons** and co-operation,
- building **a robust infrastructure** that enhancesthe valueof the data.

Overall, the case study is expected to contribute to formulating a new **“national DSIP strategy”.**

DSIP components *and key (potential) sources*

Framework setting			
Regulation	Inter-operability	Data access	Skills (digital, management, interpretation..)
Trust and privacy	Organisational dynamics	Data quality	

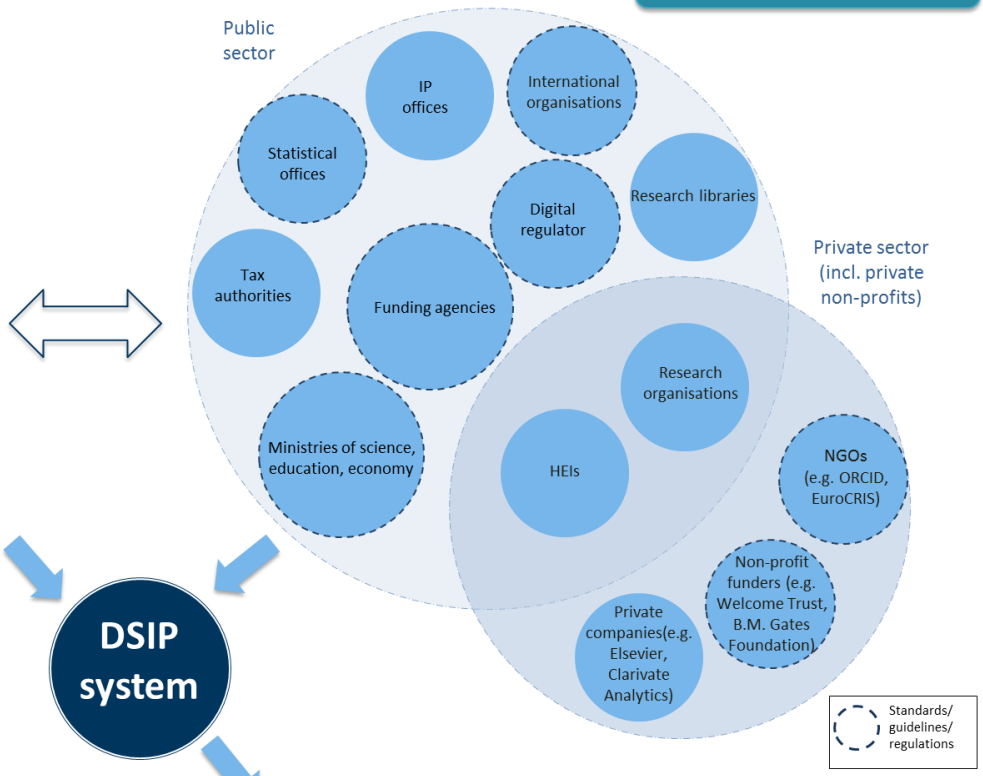
	Funding		Doing		Outputs		Impacts	
Data sourcing and use	Funds allocated/ disbursed	*Funders Buyers*	HR data (info on the people involved in R&D)	*S&I performers*	Patent filings, approvals	*Government (IP offices)*	Patent renewals	*Government (IP offices)*
	Own funds	*S&I performers*	Project management info: (progress reports, etc.)		Publications	*S&I performers Publishers*	Citations	*Publishers, Google scholar etc.*
	Funds received	*National Statistical Agencies*	Transactions (e.g. accounting data, charges against funds received)	*Funders*	Blog posts, MOOCs, etc.	*Host services*	altmetrics	*Social media platforms*
	Tax support	*Tax authorities* *S&I performers*	Facilities/ infrastructures	*Operators*	Events	*Adverts, reports*	Local/ national economic impacts	*National Statistical Agencies*
			CV data (qualifications, experience, etc.)	*Job sites, networks*				

Technology adoption		
	Unique, persistent, and pervasive identifiers (UPPIs)	*ORCID, ISNI, GRID, Ringgold, etc.*
	Current Research Information Management Systems (CRIS), Research Information Management Systems (RIMS)	*Private providers* *In-house developers*
	Big data analysis Semantic web Machine Learning Artificial intelligence, Neural networks Natural language processing Blockchain Privacy-supporting tech. etc.	

DSIP development cycle

Although DSIP systems vary quite widely, most go through an iterative development process. Systems can be compared in terms of their progress through this cycle.

Key DSIP actors



DSIP system

Outputs

- Connected research inputs and outcomes
- Disambiguated data
- Data visualisations
- Long-term availability of data

Challenges in use/adoption

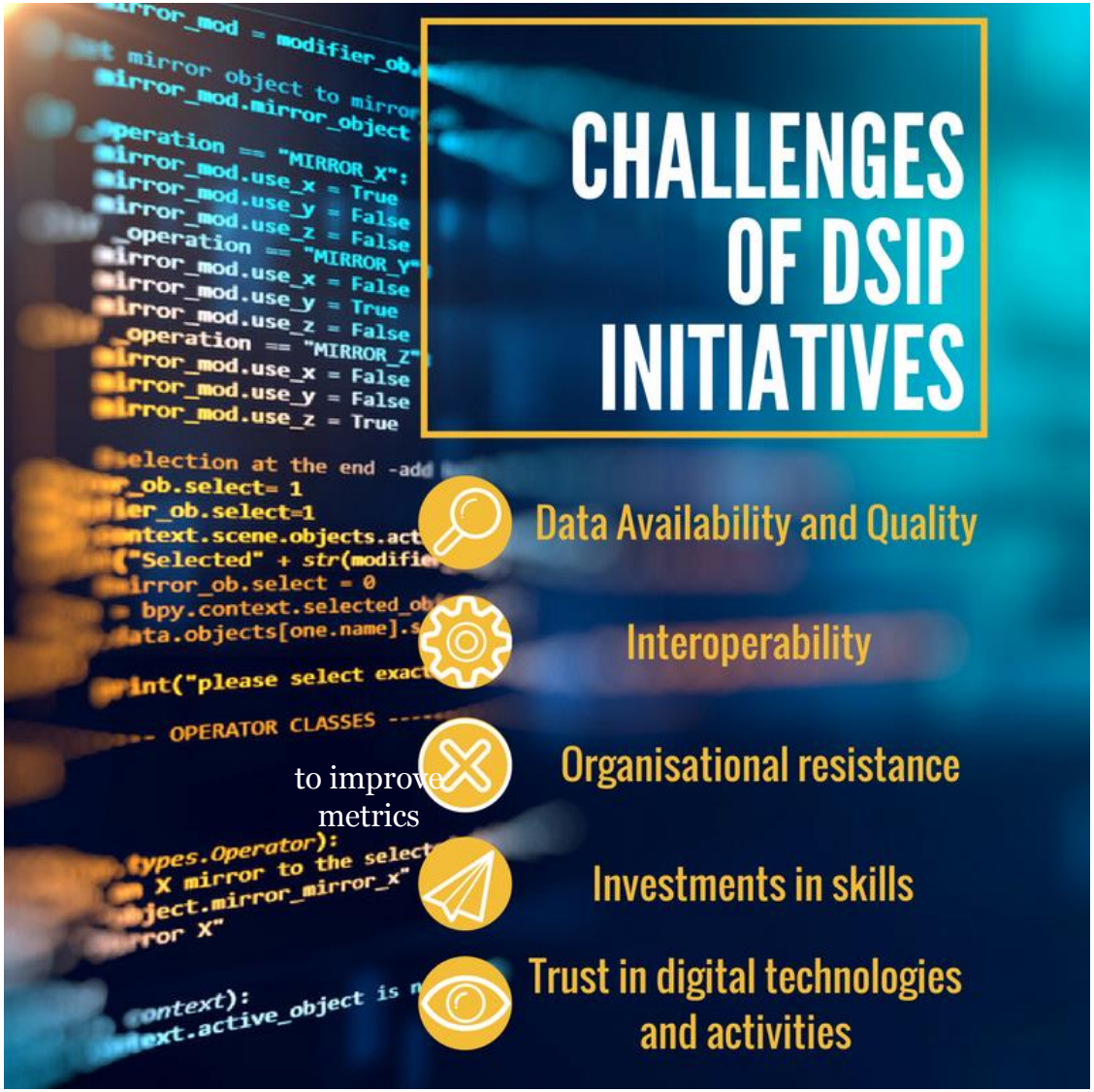
Skills and capabilities
Funding, sustainability and business models
Organisational change
Data quality
Regulatory uncertainty

1. Policy formulation and design
2. Performance monitoring and management
3. Anticipatory intelligence
4. Optimisation of research workflows

Key questions, challenges, issues…

- **Country studies aim to better understand:**
  - Data access – how are relevant data identified, what are the challenges of gaining access?
  - Coverage – what entities, transactions, etc. are the subjects of the system and how is population coverage ensured?
  - Usage – what outputs are generated and how are they used? Who are the customers of DSIP systems?
  - Quality – how are quality metrics (and underlying data) ensured?
  - Privacy – how is privacy managed in DSIP systems?
  - Funding and sustainability – how are systems funded
  - Market power issues – relating to “monopoly” systems
  - …and more

Key challenges



to improve metrics

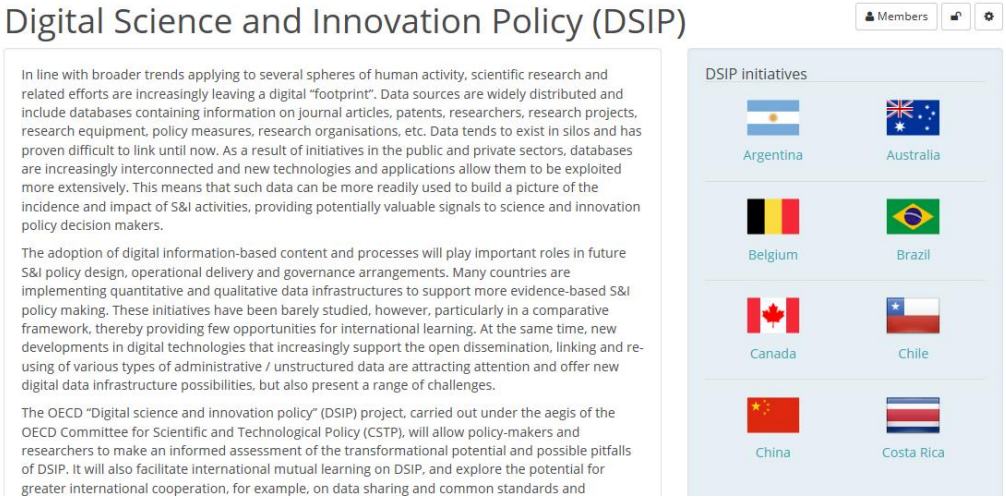
Getting involved

- Participate in the survey of DSIP initiatives
- Lend support to in-depth reviews
- Commission a country case study

Benefits

Most DSIP initiatives are relatively new and there's currently scant knowledge about them. The project offers opportunities for mutual learning and support for countries developing DSIP systems

Keep up-to-date: Community of Practice on the Innovation Policy Platform



Latest information on project updates

Information on DSIP initiatives

A platform for discussion and engagement of experts

project website: **oe.cd/DSIP**

Key takeaways

- The survey of initiatives is about to get underway (questionnaire + interviews):
  - Check that we are covering DSIP initiatives in your country – see the table on the project website
  - The Secretariat will copy CSTP delegates into messages to coordinators of DSIP initiatives
- Countries should indicate to the Secretariat as soon as possible their interest in a case study of their national DSIP ecosystem