

# Impact assessment and the Strategy Evaluation Protocol

Ludo Waltman  
Centre for Science and Technology Studies (CWTS)  
Leiden University

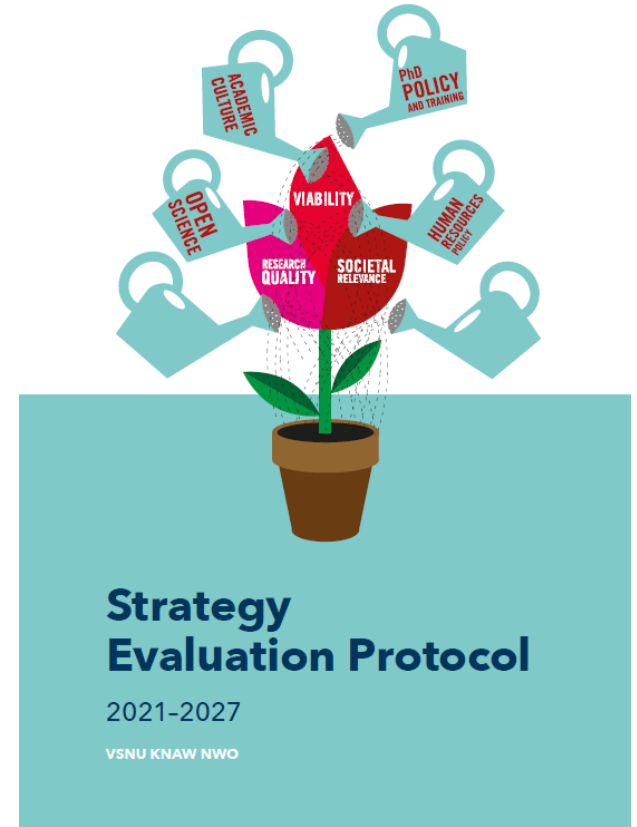
Impact Assessment Workshop  
Helsinki, Finland  
March 16, 2023

# Outline

Strategy Evaluation Protocol

Our SEP experience at CWTS

Conclusions and lessons learned



# Strategy Evaluation Protocol

# Strategy Evaluation Protocol (SEP)

Joint protocol:

- Dutch Research Council (NWO)
- Royal Netherlands Academy of Arts and Sciences (KNAW)
- Universities of The Netherlands (UNL)

“Main goal of SEP is to **maintain and improve quality and societal relevance** of research as well as to **facilitate continuous dialogue about research quality, societal relevance and viability** in context of research quality assurance”

“Main goal of a SEP evaluation is to **evaluate a research unit in light of its own aims and strategy**”

SEP is not about funding allocation!

This article is more than  
2 years old

ANALYSIS | 2/07/20

# Dear REF, please may we have a SEP?

What should replace the REF? Elizabeth Gadd is looking to the Netherlands



Image: Shutterstock



Elizabeth Gadd

Elizabeth (Lizzie) Gadd is  
Research Policy Manager at  
Loughborough University.

**A**mong all the recently research-related news, we now know that UK universities will be making their submissions to the Research Excellence Framework on 31 March 2021.

And a **series of proposals** are in place to mitigate against the worst effects of COVID-19 on research productivity. This has led to lots of huffing and puffing from research administrators about the additional burden and another round of 'What's

## Going Dutch

One of the research evaluation approaches I've **often admired** is that of the Dutch Standard Evaluation Protocol (SEP). So when I saw that the Dutch had published the next iteration of their **national research evaluation guidance**, I was eager to take a look. Are there lessons here for the UK research community?

I think so.

## Formative not summative

Of course the biggest win from a SEP-style process rather than a REF-style one is that you end up with a forward-looking report and not a backward-looking score. It's **often struck me as ironic** that the REF prides itself on being "a process of expert review" but actually leaves institutions with nothing more than a spreadsheet full of numbers and about three lines of written commentary. Peer review in, scores out. And whilst scores might motivate improvement, they give the assessed absolutely zero guidance as to how to make that improvement. It's summative, not formative.

The SEP feels truer to itself: expert peer review in, expert peer review out. And not only that but "The result of the assessment must be a text that outlines in clear language and in a robust manner the reflections of the committee both on positive issues and – very distinctly, yet constructively – on weaknesses" with "sharp, discerning texts and clear arguments". Bliss.

The SEP assessments help boards and units alike to monitor and improve the quality of research conducted by the research unit as part of the ongoing quality assurance cycle. Additionally, the assessments of the research quality and societal relevance of research contribute to fulfil the duty of accountability towards government and society.

The diagram consists of two light blue rectangular boxes with blue borders, one above the other. The top box contains the word 'Formative' in bold blue text. A blue arrow points from the right side of this box to the text 'quality of research' in the paragraph on the left. The bottom box contains the word 'Summative' in bold blue text. A blue arrow points from the right side of this box to the text 'contribute to fulfil the duty of accountability' in the paragraph on the left.

**Formative**

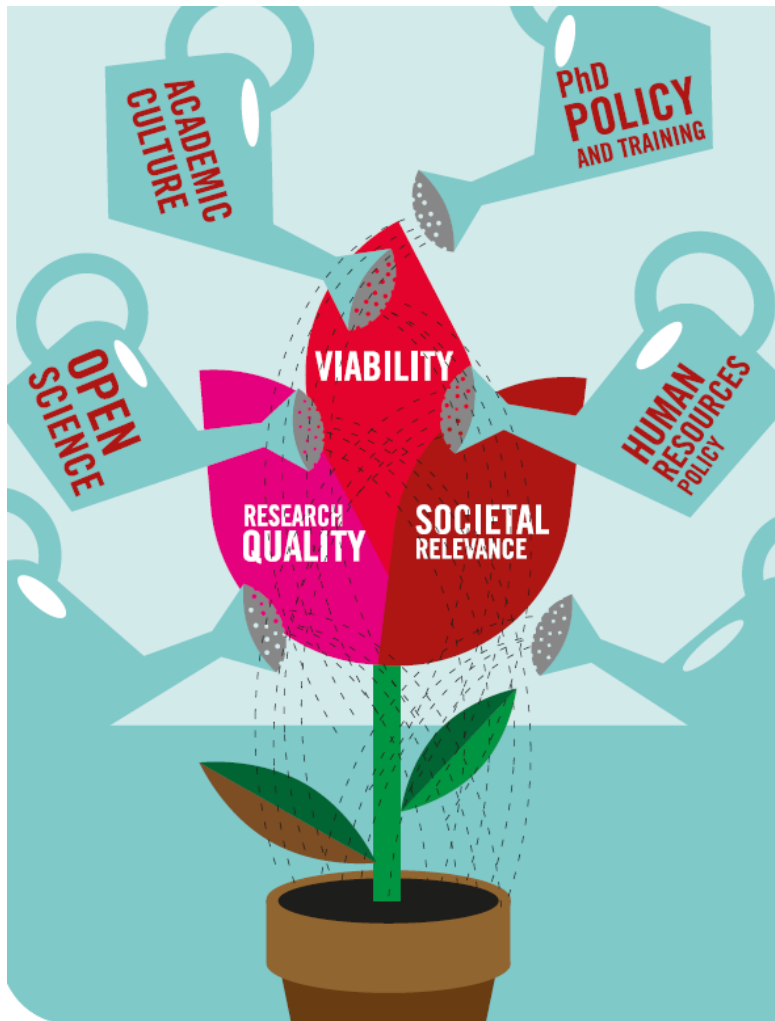
**Summative**

Academic research in the Netherlands is evaluated every six years on a rolling basis. The executive board of the relevant university, the board of NWO or the board of KNAW is responsible for these assessments. The board decides which research units are to be evaluated in which year. 'Research units' refer to institutes, departments, research groups or multidisciplinary clusters with their own research strategy, or other relevant units as defined by the board that commissions the evaluation. The main goal of a SEP evaluation is to evaluate a research unit in light of its own aims and strategy, including the sufficiency or appropriateness of the aims and strategy.

Each research unit is evaluated once every six years

Research units are defined at the level at which research strategies are designed

Evaluation is done in the light of the strategy of a research unit



Main assessment criteria:

- Research quality
- Societal relevance
- Viability

Specific aspects:

- Open science
- PhD policy and training
- Academic culture
- Human resources policy



Board defines the research unit and appoints an assessment committee

Research unit performs a self-evaluation

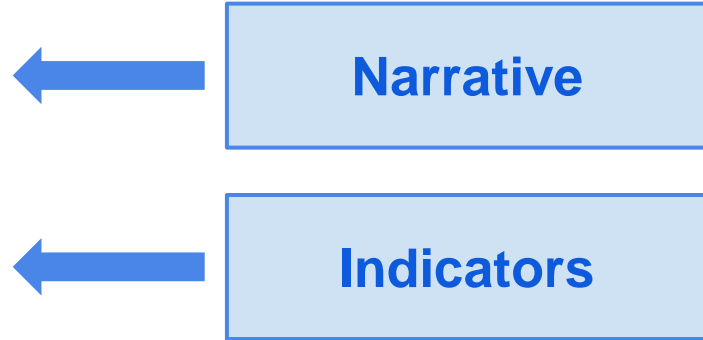
Research unit organizes a site visit for the assessment committee

Assessment committee submits an assessment report to the board

Unit and board respond to the assessment report; all documents are published

# Self-evaluation

For the **past six-year period**, the achievements are documented in the shape of a narrative argument, wherever possible supported with factual evidence (where appropriate, the unit can use quantitative indicators). The unit should choose indicators that are justified in the narrative argument to underpin the scientific achievements of the unit properly, in the context of the national or international research field, its societal relevance in terms of impact and engagement, as well as the way in which these scientific and societal achievements are related. The narrative argument is further illustrated by one or more case studies (see Appendix E3).



**Table E1: Categories of evidence for the quality domains of research quality and relevance to society**

		Quality domains	
		Research quality	Relevance to society
Assessment dimensions	Demonstrable products	1. Research products for peers	4. Research products for societal target groups
	Demonstrable use of products	2. Use of research products by peers	5. Use of research products by societal target groups
	Demonstrable marks of recognition	3. Marks of recognition from peers	6. Marks of recognition by societal target groups

4. Examples of indicators:

- Books
- Patents
- Films for a professional audience
- Websites for professional visitors
- Software for general users
- Lectures for a general audience
- Blogs for general readers

5. Examples of indicators:

- Projects with societal parties
- Contract research
- Use in education
- References in professional and public domains

6. Examples of indicators:

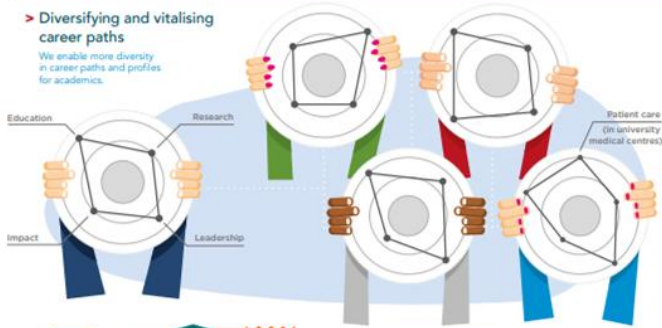
- Financial and material support by society
- Membership of civil society organizations
- Appointments within civil society organizations
- Public prizes

# Room for everyone's talent

towards a new balance in the recognition and rewards of academics

## > Diversifying and vitalising career paths

We enable more diversity in career paths and profiles for academics.



## > Achieving balance between individuals and the collective

We assess academics based on both their individual and their team performance.



## > Focusing on quality

In our assessments of academic performance, we increasingly focus on quality, content and creativity.

## > Stimulating open science

We encourage academics to share their research outcomes with society.



## > Stimulating academic leadership

We stimulate good academic leadership at all levels.

In addition to the SEP, the Netherlands also has a country-wide initiative ('Recognition & Rewards') focused on improving the way academics are recognized and rewarded

## Nieuwe Erkennen en waarderen schaaft Nederlandse wetenschap

Opinie | door gastauteurs

19 juli 2021 | Een groep van 171 wetenschappers, waaronder 142 hoogleraren, waarschuwt in deze open brief dat het nieuwe Erkennen en Waarderen de Nederlandse wetenschap schaaft. Zeker de medische, exacte en levenswetenschappen dreigen door het nieuwe Erkennen en Waarderen hun internationale topositie te verliezen omdat niet meer duidelijk is waarop wetenschappers worden beoordeeld.



"Een groep van 171 wetenschappers, waaronder 142 hoogleraren, waarschuwt dat het nieuwe Erkennen en Waarderen de Nederlandse wetenschap schaaft." Beeld: Ruben Nijveld (Bron: Flickr CC BY 2.0)

**Leonie van Drooge** says:  
Jul 7 2020 at 3:07 pm

This contribution indeed identifies some of the crucial choices and characteristics of the Dutch protocol. I am member of the standing working group SEP. In reaction to some remarks and passages:

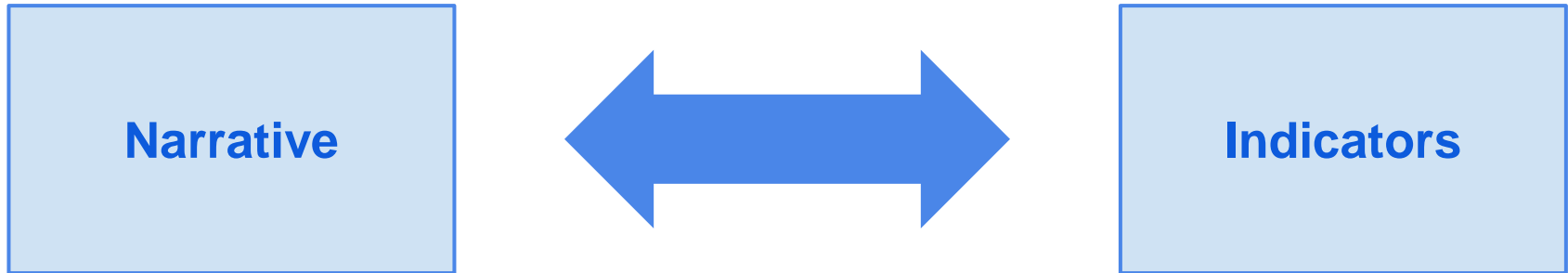
The new protocol is the next in a series of protocols that were dedicated to formative evaluation, that emphasized the context of the unit, etc. In practice, the protocol was sometimes applied, let's say, somewhat different than intended. Researchers, whether they are being evaluated or are evaluating, seem to have a reflex to think in rather simplified and quantitative measures for research quality. Also, societal relevance is a criterion that not all researchers feel comfortable with. And it seems some boards are only interested to know whether a research unit is excellent, or not. Whatever that means.

The proof of the pudding is in the eating. This time around we organise training sessions (and publish a video). In the training we explain the intention of the protocol. But we have also scheduled ample time for the participants to discuss how to ensure that evaluation will proceed as intended. After all, some elements are not standard practice and for some aspects, including Open Science, there are no commonly agreed and consolidated definitions nor measures yet. Those involved in an evaluation, in whatever role, might stumble upon uncharted territory, for which there are no instructions in the protocol. Not even if the protocol were more extensive, say 53 or 86 pages.

I know researchers in the Netherlands and the UK. Those in the Netherlands seem to look at a SEP evaluation with somewhat less fear and aversion, as compared to their UK colleague and the REF. But a perception of unwanted management interference is often present. I think that will always be the case. Also, researchers find the process time consuming. Writing a concise self-evaluation report, in a narrative style, that includes strategy and context is not an easy task. Was it Mark Twain, who apologized for writing a long letter, because he didn't have time to write a short one? On the other hand, and with hindsight, researchers tend to value the discussions that took place throughout the process, both internally as well as with the assessment committee.

Our SEP experience at CWTS

# Performing the self-evaluation



<b>1</b>	<b>Introduction</b>	<b>6</b>	<b>Appendices</b>	<b>31</b>
<b>2</b>	<b>Organisation</b>	<b>6</b>	<b>A Evidence and indicators</b>	<b>31</b>
2.1	Governance	6	A.1 Indicators, Quality Domains and Assessment Dimensions	31
2.2	Research organisation	7	A.2 Evidence and indicators: rationale and justification	32
2.3	CWTS BV	8	<b>B Case Studies</b>	<b>40</b>
2.4	Human Resources	8	Case study 1: The Evaluative Inquiry	40
2.5	Finances	8	Case study 2: Toward a globally inclusive understanding of the science system	41
			Case study 3: Using VOSviewer visualisations for contextualisation and participation	44
<b>3</b>	<b>Mission</b>	<b>9</b>	<b>C Tables with Chosen Indicators</b>	<b>50</b>
3.1	Period 2016-2019	9	C.1 CWTS Personnel	50
3.2	Period 2019-2021	10	C.2 CWTS Funding	50
			C.3 CWTS PhD Candidates	52
			Total overview of PhD Candidates	52
			Active Internal PhD Candidates	53
			Active External PhD Candidates	53
			Supervision of PhD candidates	54
			C.4 Overview of the institute projects 2015–2020	55
<b>4</b>	<b>Accomplishments - Strengthening the science system</b>	<b>11</b>	<b>D Stakeholders' Consultation meeting</b>	<b>57</b>
4.1	Research governance and evaluation	11	Session One	57
	Engaging in research governance and evaluation	12	Session Two	57
	Understanding research governance and evaluation	13		
4.2	Science and society	14		
	From measuring impact to tracing interactions	15		
	Responsible Research and Innovation	15		
	Societal needs, socio-ecological issues, and sustainable development goals	16		
4.3	Open science	17		
	Open scholarly communication	17		
	Open infrastructures	18		
	Open science policy making	18		
4.4	Methods and tools	19		
	Advanced technical infrastructure	19		
	Contextualization of scholarly activities	20		
	Conversational and participatory approaches	20		
<b>5</b>	<b>Accomplishments - Practising what we preach</b>	<b>21</b>		
5.1	Academic culture	21		
5.2	PhD training	22		
5.3	Open science and data management	24		
5.4	Commissioned projects and CWTS BV proposition	25		
<b>6</b>	<b>Reflection and strategy for the future</b>	<b>26</b>		
6.1	SWOT analysis	26		
	Strengths	26		
	Weaknesses	27		
	Opportunities	27		



# Digital Response Analysis - User analysis based on links



Navigation menu with three items: Google Analytics Data, Google Analytics (2) - Number of users per type, and Ahrefs Data. The second item is highlighted.

## Google Analytics - Top 100

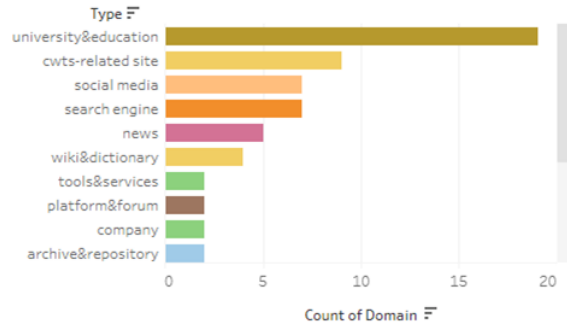
Meta-Category: (All)

Type: (All)

### Top 100 - Number of users, domain and type

Users	Domain	Type
23624	m.facebook.com	society
23549	t.co	society
14229	baidu.com	society
12823	vosviewer.com	peers
10453	linkedin.com	society
10321	en.wikipedia.org	society
9033	researchguides.uic.edu	peers
8098	leidenranking.com	peers
7746	facebook.com	society
6128	cn.bing.com	society
3438	researchgate.net	society
3400	l.facebook.com	society
3258	en.m.wikipedia.org	society
2792	jornal.usp.br	society
2740	cwts.nl	peers
2571	skku.edu	society
2094	duckduckgo.com	society

### Categories - Number of domains



### Categories - Number of users



Assessment dimensions	Quality domains	
	Research quality	Relevance to society
Demonstrable products	<b>1. Research products for peers</b>	<b>4. Research products for societal target groups</b>
	1.1 Number and list of peer reviewed scientific publications	4.1. Blog posts and other media outputs, policy briefs, etc.
	1.2. Number and percentage of Open access publications	4.2. Number and list of data & software produced
	1.3. List of training and education courses organised by CWTS	4.3. Number of CWTS BV contracts/reports and total funding.
	1.4. PhDs supervised at CWTS	4.4. Number of courses for professionals
	1.5. Networks and collaborations at CWTS	4.5. List of podcasts from the FluidKnowledge project
Demonstrable use of products	<b>2. Use of research products by peers</b>	<b>5. Use of research products by societal groups</b>
	2.1. Number of citations received by CWTS scientific publications	5.1. Total number of policy citations received by CWTS outputs
	2.2. Highly cited publications produced by CWTS	5.2. Total numbers of news and blogs mentions of CWTS outputs
	2.3. Readership by academic users	5.3. Total number of original tweets to CWTS outputs
	2.4. Number of participants in academic courses organised by CWTS	5.4. Amount of online traffic directed to CWTS from societal groups (e.g. news, publishers, etc.)
		5.5. Number of participants of professional BV courses

### 3. Marks of recognition by peers

- 3.1. List of institute projects funded
- 3.2. List of academic prizes and awards received by CWTS researchers
- 3.3. List of invited talks and lectures, and invited positions
- 3.4. List of academic memberships (boards, advisory roles, etc.)

### 6. Marks of recognition from societal groups

- 6.1. Number of links made to CWTS digital sources by non-academic platforms
- 6.2. Number of likes and retweets to tweets mentioning CWTS publications
- 6.3. BV projects with clients repeating a project with CWTS BV



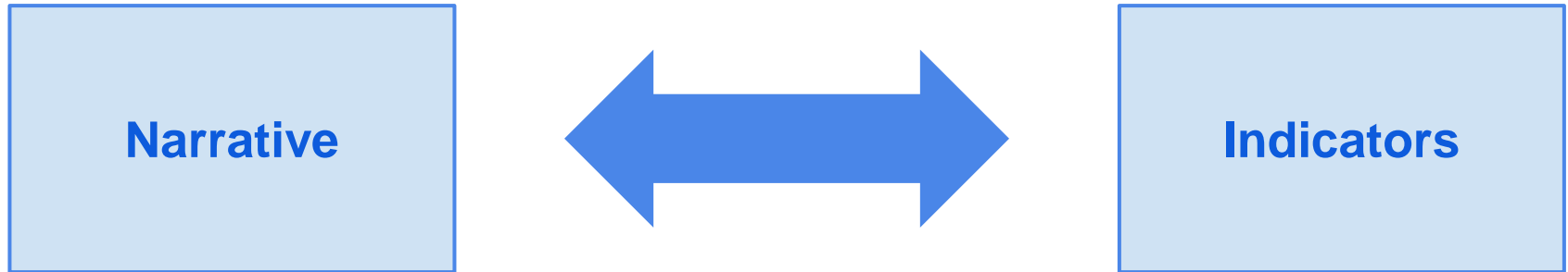
Strategy Evaluation Protocol:  
CWTS self-evaluation report

2015–2020

March 2022



# Providing evidence - Connecting narratives and indicators



# Assessment committee



The composition of the committee was as follows:

- Pierre-Benoît Joly (chair), Director of Research at the National Institute of Research for Agriculture, Food and the Environment (INRAE) and Chair of the INRAE Center Occitanie-Toulouse, France
- Anssi Mälkki, Director of Research Management Development, Research Services, University of Helsinki
- Arianna Becerril Garcia, Professor at the Autonomous University of the State of Mexico, Executive Director of Redalyc
- Steven Hill, Director of Research at Research England, UK Research and Innovation
- Liz Allen, director of Strategic Initiatives at F1000, Taylor & Francis Group & Visiting Senior Research Fellow, Policy Institute, King's College London.
- Ivan Veul, PhD candidate at Radboud University (PhD student member)

**Note the prominent role of societal stakeholders  
in the assessment committee**

“Thank you for preparing this dashboard for us; it looks really impressive!”

“However, we are actually not entirely sure what to do with it”

### 3.4 Societal Relevance

#### *General assessment*

To assess the societal relevance of CWTS's research activities, the committee considered the interactions that CWTS has with society and the associated strategy, the involvement of CWTS in open science, and various output for societal target groups, such as blogs, commissioned projects, courses and the use thereof.

The committee concludes CWTS has highly invested in the societal relevance of its research. CWTS researchers are often invited for influential policy discussions on the research system, such as the European Expert Group on Indicators for Researchers' Engagement with Open Science, the Open Science Advisory Committee of UNESCO and the Dutch Taskforce on Responsible Management of Research Information and Data. The institute takes care to make its work policy-relevant and -sensitive, working in partnerships with research funders and policy makers. Furthermore, the institute is in the process of setting up a UNESCO Chair: a UNESCO-supported research position related to one of UNESCO's priority areas. This Chair will focus on diversity and inclusion in global science. CWTS has a strong network of international partners in Low- and Middle-income Countries and emerging economies, such as in Mexico and Brazil, and through the Participatory Research in Asia (PRIA) network. Another special chair has been established with the Rathenau Institute, the Netherlands national institute for technology assessment. These special chairs strongly connect CWTS's research with relevant stakeholders, which the committee considers a strong asset of CWTS.

Overall, CWTS has an impressive network of stakeholders related to research evaluation. Many collaborations are established through the services for research evaluation it provides through CWTS BV. This allows CWTS to bring its expertise into practice, and use input collected through commissioned projects to further develop its ideas, methods and tools. The committee considers that interactions between CWTS and its in-house company are very fruitful and that they accelerate translation between research and practice. CWTS is also successful in working with commercial partners, such as the major publishers, which are influential parties in the research evaluation system. The committee considers these collaborations to be very important when striving for responsible research evaluations.

CWTS is also very active in direct communication with societal stakeholders. The institute has a very influential blog with frequent contributions by CWTS researchers on current debates in research evaluations and new findings from CWTS research, as well as [a newsletter with over 200 subscribers](#). The methods and tools that CWTS develops are publicly available, with opportunities for extra courses and training through CWTS BV. The committee is impressed by these activities and the efforts CWTS makes to create societal impact.

# Conclusions and lessons learned

# Conclusions and lessons learned

We are quite happy with our Strategy Evaluation Protocol

Societal relevance is still challenging

Recognition & Rewards is another major challenge



# Conclusions and lessons learned

What kind of **assessment** do we need? Summative vs. formative

What kind of **evidence** do we need? Narrative vs. indicators

What kind of **infrastructures** do we need?  
Where is the data? And how to make sense of it?

Thank you for your attention!