Public policies and evidence: reflections based on Ivàlua's experience

Frederic Udina,
Chairman of the
Governing Council of
the Catalan Institute
for the Evaluation of
Public Policies (Ivàlua)
and Director at the
Statistical Institute of
Catalonia (Idescat)

Marc Balaguer,
Executive Director at
the Catalan Institute for
the Evaluation of Public
Policies (Ivàlua)

Abstract

The purpose of this article is to explain what evidence-based public policies are and to identify, on the basis of Ivàlua's experience and the examination and analysis of the way different evaluation agencies work, which factors contribute to establishing the use of evidence as normal practice for public administrations.

The article is structured into three parts; the first explains the current use of evidence in Catalonia and Spain, the second defines the link between public policies and evidence, and the third describes some elements that contribute to the greater use of evidence by public administrations.

The development of evaluation in Spain and Catalonia

The presence of an evaluation culture in public administrations is increasingly common, although the situation is still very uneven across the OECD countries. When countries are ranked in terms of their use of evidence – see Jacob et al (2015) and Furubo J-E, Rist RC and Sandahl R. (eds.) (2002)[1] –, Spain is toward the bottom of the list. This is due to the lack of an evaluation culture, something that is not helped by the limited presence of evaluation agencies[2].

Barriers to the development of this evaluation culture include: a) poor awareness of evaluation and a lack of incentives for senior people to promote it; b) confusion among administration staff about the difference between evaluation, monitoring and economic, organisational and/or management audits; c) the varying ability of administrations to promote and

The presence of an evaluation culture in public administrations is increasingly common, although the situation is still very uneven across OECD countries

---

1. Jacob et al (2015): Updating the International Atlas of Evaluation 10 years later. Evaluation 2015, Vol. 21(1) 6-31; Furubo J-E, Rist RC and Sandahl R (eds) (2002) International Atlas of Evaluation. New Brunswick, NJ and London: Transaction Publishers.

2. With the conversion in July 2017 of AEVAL (State Agency for the Evaluation of Public Policies and Quality of Services) into the General Sub-Directorate Institute for the Evaluation of Public Policies, forming part of the Secretariat of State for Public Administration, that is, the Ministry of Finance and Public Administration, Ivàlua became the only evaluation agency in Spain.

define evaluation processes; d) the under-use of administrative data for evaluation purposes; and e) a lack of mechanisms for transferring knowledge from the academic world to administrations.

Generation and use of evidence in evaluation

The European Commission (EC, 2007)[3] defines evaluation as "the judgement of the interventions of public authorities according to their results, impacts and the needs they aim to satisfy and aimed at providing a rigorous evidence base to inform decision-making".

Different types of evaluation deal with different public policy issues and, therefore, generate knowledge (evidence) about them.

3. European Commission 2007: **Communication to the Commission from Ms Grybauskaité in agreement with the President: Responding to Strategic Needs: Reinforcing the use of evaluation**. SEC 2007 (213).

Types of evaluation carried out and the issues they address

Evaluation types	Issues they address
Evaluation of design	Do the public intervention's objectives and activities reflect the theory behind the evaluated programme?
Evaluation of needs	What problem/problems need to be addressed? What information is necessary to generate evidence? How is this scaled? Which public services and resources are necessary to address the problem(s) identified?
Evaluation of implementation	To what extent does the theoretical design of a policy correspond to its actual functioning in practice? Does the target audience receive the services provided?
Evaluation of impact	What are the effects of the intervention on the recipients of the public policy and the general public? Has the intervention led to the expected solutions to the problem initially raised?
Economic evaluation	What is the economic evaluation of the impact achieved? To what extent does this impact offset the costs of the policy? To what extent can the policy be considered efficient?

Source: own elaboration.

Elements that contribute to a greater use of evaluation by public administrations

Aspects relating to the generation and development of evidence

Obtaining the answer to each of these questions through evaluation is a necessary but not sufficient condition for this knowledge to be incorporated into policies or, in other words, for it to be used.

Certain strategies contribute to breaking down the aforementioned barriers, resulting in an effective use of evidence in public decision-making. Some elements are important in the evaluation programme preparation phase while

others are important during the launch of the evaluations or after they have been completed.

Evaluation programme preparation phase

1) Areas for evaluation: In environments where evaluation is more normal, it extends to a larger number of areas and sectors. In this type of environment there are evaluation agencies specialising in the policy area they evaluate and even in the type of evaluation they carry out. A clear example of this is the United Kingdom with agencies such as Nesta, specialised in the quasi-experimental evaluation of innovation policies, and the National Institute for Clinical Excellence (NICE), specialised in health-based evaluations. In environments where the use of evaluation is more limited, agencies tend to have a more generic approach in terms of sectors and methods; this is the case for the LIEPP (Laboratoire Interdisciplinaire d'Évaluation des Politiques Publiques) in Paris and also for Ivàlua.

2) Collaboration between the supply of and demand for evaluation: To encourage the use of evaluation, it is important for there to be bridges between the supply of and demand for evaluation. Here the Nordic model stands out, providing public financing to research centres so that they can prepare reviews of evidence in departmental areas identified by the government. The government then takes that evidence into account when prioritising the

financing of different public intervention alternatives.

3) Alliances between agencies: Networking is a way of multiplying efforts to extend the evidence culture. One interesting example of such collaboration is the Alliance for Useful Evidence, in which agencies such as Nesta, the Economic and Social Research Council and the Big Lottery Fund have agreed to promote the greater use of evidence by the public administrations. In Catalonia, the different public agencies specialising in evaluation, namely the Agency for Health Quality and Assessment of Catalonia (Agència de Qualitat i Avaluació Sanitàries de Catalunya, AQUAS), the Catalan University Quality Assurance Agency (Agència per a la Qualitat del Sistema Universitari de Catalunya, AQU Catalunya) and Ivàlua, have recently started to work together.

4) Collaboration with institutions and key players within the administration: Collaboration between the evaluation agencies and the management units that have shown the most interest in evidence is important; this improves the quality of the demand for evaluation and encourages its use. In this regard, the agreement between the Public Administration School of Catalonia (Escola d'Administració Pública de Catalunya, EAPC) and Ivàlua is noteworthy in Catalonia. This has made it possible to identify the degree of expertise among the managerial positions in different fields, including evaluation, and, based on this,

The Nordic model provides public financing to research centres so that they can prepare evidence reviews in departmental areas identified by the government

identify areas for improvement in the programming of management training.

5) Evaluation programmes: Building complete evaluation programmes must look beyond a one-year period. This allows for the long-term programming of evaluations and for these to be scheduled strategically. This can be seen in the strategy adopted by the **Area for the Evaluation of Public Policies** (Àrea d'Avaluació Econòmica de Polítiques Públiques) which forms part of the Department of the Vice-presidency and of the Economy and Finance (Departament de la Vicepresidència i d'Economia i Hisenda) in the Government of Catalonia (Generalitat de Catalunya).

The performance of evaluations and advisory services

1) Multidisciplinary teams: Having a multidisciplinary team allows the evaluation to be approached from different viewpoints and contributes to narrowing the gap between the demand for the evaluation and the results it produces.

2) Collaboration between evaluation professionals: Collaboration with universities and evaluation professionals can contribute to generating a community of evaluation professionals. The American Evaluation Association stands out as an example of a very active community of professionals.

A multidisciplinary team allows the evaluation to be approached from different viewpoints and contributes to a narrowing of the gap between the demand for evaluation and the results it produces

3) Policy briefs: It is important for evaluation reports to include a short summary and visual content so that the results and recommendations can reach a wide audience of public decision-makers and become relevant. The Policy Briefs of the Campbell Collaboration, those of 3ie and those prepared by IPA all stand out.

4) Promoting advisory services as an element of the evaluation culture: The OECD defines evaluability as the extent to which an activity or programme can be evaluated in a reliable and credible fashion. Entities such as **Better Evaluation** emphasise the usefulness of carrying out evaluability studies in those environments where evaluation is at an early stage, meaning that any currently non-evaluable policies may become evaluable in the future thanks to the recommendations included in these diagnoses.

The administration's capacity to incorporate evidence

It is necessary to train the people who work in the public administrations so that they adopt the culture of evidence on a daily basis. For this reason, the formulation and generation of training resources is essential when promoting the evaluation culture in public institutions.

1) Training at all levels: As stated in **Evidence for Policy Design** (EPoD) at Harvard Kennedy School, it is important for training activities to be aimed at all public decision-makers and players who

It is important for evaluation reports to include a brief summary and visual content so that the results and recommendations can reach a wide audience of public decision-makers

It is necessary to train the people who work in the public administrations so that they incorporate the culture of evidence on a daily basis

participate in the different levels and stages of a public policy.

2) Practical training: Entities offering evaluation training highlight the need to offer training that incorporates practical cases and group work, if possible with a range of profiles in the room, in order to achieve complementarities. This contributes to the generation of proposals that may later become evaluations.

The dissemination of evidence

Likewise, it is also essential to publicise the evidence and the focus of evidence-based public policies.

1) Public dissemination of results: It is necessary to publicly disseminate evaluations using content, channels and formats adapted to different audiences, and to promote debate around the results obtained and any possible consequences. It is especially important to systematically publish all available evidence and organise public presentations of the results involving the organisation responsible for the original programme.

2) Influence the perception of evaluation: It is important to make evaluation and its policy benefits known to politicians, directors, managers and technical staff, representatives of economic and social interests and the public in general. This is key to the progressive incorporation of evaluation into public agendas.

Conclusions

- Carrying out evaluations, evaluation training and disseminating evidence cannot be seen as isolated activities, but rather as interrelated activities necessary to progress in the generation and use of evidence.

- One of the objectives of the evaluation of public policies, and possibly the most important, is the generation of evidence to allow for an improvement in policies and informed decision-making. Therefore, it is essential to systematically monitor the extent to which evidence is incorporated into public decision-making, as well as any possible advances in this regard.

- It is important to work with specialised institutions both internationally and in our country, as well as with the organisations responsible for evaluation projects, the academic world and the community of professionals who may implement them. This contributes to the generation of an evaluation community in our country. ▌

It is essential to systematically monitor to what extent evidence is incorporated into public decision-making

Knowledge Management for Policy:
a new collective skillset

Lene Topp
& David Mair,
Joint Research Centre
of the European
Commission
(European Union)

Abstract

Knowledge Management is about ensuring that the most useful and robust facts are provided and understood in good time for them to be taken into account by decision-makers. Effective Knowledge Management for Policy, scientific advice to policy and evidence-informed policymaking are synonymous values which underpin a need for a new profession with a collective skillset. The Joint Research Centre (JRC) has mapped the essential skills of the new knowledge brokers (researchers and policymakers) in boundary-spanning organisations at the evidence-policy interface.

The European Commission has proven its commitment to evidence-informed policymaking through the establishment of the Regulatory Scrutiny Board and the development of the better regulation tool box. At the Regulatory Scrutiny Board Conference in 2017, Frans Timmermanns, First Vice President of the European Commission, stated "we have worked hard to embed better regulation into the DNA of the European Commission, installing a priority driven, evidence-based, disciplined, transparent and above all inclusive policy process". The Joint Research Centre, as the European Commission's knowledge and science service, therefore has a very clear mandate to support EU policies with research evidence. The use of evidence for evaluation and to yield impact of EU policy is a means to grant legitimacy to the work of the European Commission.

At the same time, the current set of economic and social crises have brought to a head a crisis in the relationship between evidence and policy and between the scientific and academic community and the world of policymakers and politicians. While the growing complexity and inter-connectedness of societal challenges and policy problems calls out for robust evidence to help make difficult trade-offs

The Joint Research Centre, as the European Commission's knowledge and science service, has a very clear mandate to support EU policies with research evidence

between competing interests and values, the current crises instead appear to be shrinking demand for evidence to support policymaking. The recent emergence of "post-fact" or "post-truth" politics (perhaps stimulated by the "filter bubble" effect of social media) and the increasing disregard for expertise seems to be making it harder for democracy to reconcile policymaking grounded in evidence with the emotional pull of values-driven politics.

The presence of these issues in the debate should be seen against a considerable degree of consensus about what needs to be done by both scientists and policymakers to change the way they interact. The debate has clearly moved on from the days of the "deficit model", seeing the problem as a lack of supply of knowledge, inability of policymakers to understand or opportunities for scientists to engage with policymakers. It is now much clearer that there is a considerable over-supply of knowledge, not all of it

robust, compared to the very limited bandwidth of policymakers. There is also a clear understanding that policy problems are increasingly "wicked" (complex and inter-dependent) calling for more cooperation between different policies and different disciplines (although inter-disciplinary work is more called for than practised). The policy process is now better understood as "messy", rather than structured and linear. There is also clearly a problem in synchronising the supply and demand so that the facts are available at the right time and place for decision-making. Finally there is lack of mutual respect, understanding and empathy between scientists and policymakers, as well as a lack of self-knowledge about behavioural biases, which at worst leads to hubristic behaviour and at best an inability to change minds through evidence.

The way forward to address these common issues can perhaps best be described as "Knowledge Management for Policy" with "knowledge brokers". Knowledge Management for Policy is about ensuring that the most useful and robust facts are provided and understood in good time for them to be taken into account by decision-makers.

Skills for the science-policy interface

Effective Knowledge Management for Policy, scientific advice to policy and evidence-informed policymaking are synonymous values which underpin a need for a new profession with a collective skillset.

There is a lack of mutual respect, understanding and empathy between scientists and policymakers, as well as a lack of self-knowledge about behavioural biases

Changes in individual behaviour will not be sufficient if enabling organisational and contextual environments do not follow. Paying close attention to the needs and skills of policymakers is equally important to those of researchers

The JRC is in a good position to experiment with approaches for more evidence-informed policymaking. This has resulted in a mapping of skills essential for researchers and policymakers active in the science-policy interface. The primary focus is on the **collective skillset of the new knowledge brokers (researchers and policymakers) in boundary-spanning organisations at the evidence-policy interface**. Changes in individual behaviour will not be sufficient if enabling organisational and contextual environments do not follow. Paying close attention to the needs and skills of policymakers (including civil servants at strategic and operational level, but not politicians) is equally important to those of researchers (including researchers and research managers). Whether producing scientific knowledge or developing policy, both sides involved with evidence require a collective skillset to address the challenges of this field.

Overview of the 8 skills clusters

The skills framework consists of 8 clusters of skills, with each cluster addressing a specific part of the collective skillset required to increase the impact of evidence on policymaking:

• **Understanding Policy & Science.**
Effective researchers understand the key drivers of the policy process – which can never be as simple as a policy cycle – and adapt their evidence presentation strategies to the policy context. Effective policymakers anticipate what evidence will be needed in the future.

- **Interpersonal Skills.** Being able to interact well with others – using verbal and non-verbal communication skills – is essential to building trust and solving problems that occur in creating and applying knowledge to policymaking.

- **Synthesising Research.** Effective knowledge management will provide policymakers with access to more robust and fit-for-purpose evidence. Effective researchers employ methods and tools to make better sense of the wealth of knowledge ("secondary research") available on a given topic.

- **Managing Collaborative Expert Communities.** "Communities" of experts, sharing a common language or understanding, are fundamental to creating and applying knowledge to complex problems. Effective researchers develop networking and facilitation skills, through digital and physical interactions, to reduce disciplinary and policy divides.

- **Communicating Scientific Knowledge.** The communication of research to a non-scientific audience requires effective communication skills, using content-related tools like infographic design, succinct writing, public speaking and data visualisation tailored to the audience.

- **Advising Policymakers.** Effective science policy advisors go beyond simply communicating research evidence towards identifying options, helping to understand the impact of policy choices during and after implementation, and providing policy advice from a scientific viewpoint.

- **Engaging with Citizens & Stakeholders.** Engagement with the public and stakeholders is essential to build trust and legitimacy of evidence to be used in policy.

- **Monitoring & Evaluation Framework.** Monitoring and evaluating the impact of research evidence on policymaking is a specific skill needed to continuously improve the influence of evidence on policymaking.

Conclusion

Knowledge Management for Policy requires a new collective skillset in organisations working at the science-policy interface. If developed and applied effectively by organisations, these skills ensure that the most useful and robust facts are provided by scientists and understood in good time for them to be taken into account by policymakers throughout the policy cycle. As the skills framework contains eight skills, we believe that neither "pure scientists" nor "professional politicians" can be successful in this work. Knowledge brokers are essential: scientists with a feel for policy and policymakers understanding how to manage science and scientists. "Knowledge brokers" support scientists in stimulating policymakers' appetite for evidence. Likewise, they help policymakers understand evidence and the impact of it. The collaboration between scientists and policymakers may not come natural, thus the training associated with the skills framework is rather a career choice than a quick fix. ▮

The collaboration between scientists and policymakers may not come natural, thus the training associated with the skills framework is rather a career choice than a quick fix

"It takes a system-wide effort to get to the stage where government actors will speak the language of evidence"

Rohini Pande,
Professor of International Political Economy at Harvard Kennedy School. She co-directs the *Evidence for Policy Design* Initiative (USA)
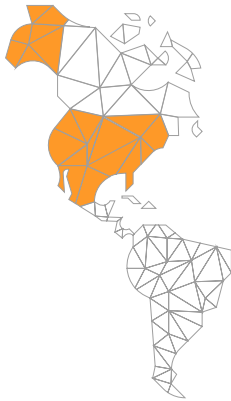
What are the greatest challenges you have identified to get public institutions to incorporate evidence into their decision-making processes?

The biggest challenge is operating in systems where the incentives of public servants are not always aligned with

citizen needs. This is a fundamental characteristic across many governments, and we are trying to understand it and identify ways to improve this alignment. We also believe that improving this alignment will strengthen the demand within public institutions for evidence on what citizens need and what works in filling these needs. When bureaucrats answer only to their supervisors and are deeply insulated from the people who benefit or suffer based on their performance, their motivation may be to maintain the status quo. It is hard for someone in this position to innovate systems to serve the public better. It might mean exposing the poor performance of the ministry, or appearing to try to outshine the boss.

So, it takes a system-wide effort of aligning incentives across many state actors to get to the stage where government actors on different levels will speak the language of evidence.

We understand that training civil servants is an important piece of it. What skills should people working in public administration have in order for them to be able to incorporate evidence into public decision-making?

Some of the most important work comes before we even approach any technical

When bureaucrats answer only to their supervisors their motivation may be to maintain the status quo

> An important first step is getting past the idea that consulting evidence will be another burden but rather that it will lead to improvements and net gains in efficiency

skills. It's imparting an appreciation of the power of evidence to help public administrators in their jobs and to improve outcomes in government. The civil servants we train are busy people working in ministries that often have entrenched ways of doing things. Time is short and inertia is strong. So, an important first step is getting past the idea that consulting evidence will be another burden – another requirement to fill or box to tick – but rather that it will lead to improvements and net gains in efficiency.

For nearly all trainees, we start with the basics of using descriptive evidence for decisions and the use of systematic policy frameworks. After that point, the skills themselves depend on the area the civil servant is operating in. We typically give higher-level decision-makers content on commissioning evidence, and on-the-ground administrators more technical training on how to consume evidence (by, for instance, being able to learn from impact evaluations).

In India and Pakistan, where our trainings are the most ingrained (they've become part of the curriculum at national civil service academies), we're beginning to hear mid-level bureaucrats say that they've been able to use evidence because their superiors already know its value and speak the language. We consider that a big success.

In the last 7 years, EPoD has trained many public servants in the use of evidence. What content should these kind of training sessions have in order to be effective? In what format should training be provided?

Again, we've found that the most effective content depends on the nature of a public servant's job, but – across the board – imparting a fundamental appreciation of evidence is the first hurdle, and not as easy as it may seem. When trainees demonstrate increased motivation to use evidence in decision-making in their work, it's a huge success in our book.

As far as the format, we've developed a blended-learning model where part of the training takes place online and part in the classroom. Both are critical: the online content to save resources and allow learners to work at their own pace, and the classroom part to ensure engagement and the ability to contour the course to the needs of the group. Our online program evaluates the students' learning as they progress, and feeds that information directly to instructors before they even enter the classroom. This allows them to give certain subjects more time if necessary and avoid wasting time on content that is elementary or irrelevant to the audience.

Is the lack of connection between politics and evidence solved solely by training civil servants? What other actions should accompany training?

Not by a long shot. The use of evidence depends on demand for evidence. In a

The use of evidence depends on demand for evidence. In a democracy, if the public doesn't require the government to operate on evidence, the government won't

democracy, if the public doesn't require the government to operate on evidence, the government won't. Effective public service can be subverted when powerful groups can campaign for their own needs above others, or when current events conspire against evidence. A good example is immigration policy over the last few years. The evidence may say that the economy of a particular country benefits from immigrants because they provide cheap labor, feed the tax base, and free nationals up for higher-productivity activities. But when voters see an influx of foreigners, the evidence may not matter, and opportunistic politicians can take advantage of that. So, in cases like this you will see the gap between politics and evidence widen without civil servants playing any role whatsoever.

A challenge that civil servants might face when using evidence is that pieces of evidence that refer to their local context might not be available. However, in this situation, we should be able to transpose evidence produced in other countries. Can you give us some clues on how to do it?

You're right, we should be able to make use of information from other countries, but a main finding from our group's research on how bureaucrats use evidence is that they have a heavy bias toward local evidence – even when compared to locations within the same country. We found in surveys that bureaucrats placed more weight on weaker evidence from their own city than on stronger evidence from another city in the same country. As with

other biases, this can be overcome through education.

But I'd also say that maybe skepticism toward evidence from elsewhere should be counteracted not just by importing evidence but by producing more locally. In many of the places we work, we find valuable evidence – produced by other government departments, for example – going unused. More and more we're moving toward performing data analytics with existing local evidence to inform policymaking. After all, some small amount of skepticism toward outside evidence is warranted: the way policy works is heavily dependent on the local context, and no one knows that better than bureaucrats.

Finally, what role should evaluation Institutions and the academy have when it comes to bringing the use of evidence closer to public administration?

We should be asking civil servants what they need, and responding to that.

Too many academics come to the table thinking they already have the answer, before the question has even been asked. A key part of our mission at Evidence for Policy Design (EPoD) is to take several steps back in the policy-research process, collaborate with policymakers to identify their priorities, and search out the windows where evidence can make a difference. Then comes a period of evaluating possible options. Only after that do you think about solutions. ▮

More and more we're moving toward performing data analytics with existing local evidence to inform policymaking. The way policy works is heavily dependent on the local context

Our mission is to take several steps back in the policy-research process, collaborate with policymakers to identify their priorities, and search out the windows where evidence can make a difference

"Academia is key but not enough. Nobody can have a monopoly of the evaluation of public policies"

Étienne Wasmer,
Professor at New York
University in Abu Dhabi,
Founder and
co-director of
Sciences Po – LIEPP,
2011-2017 (France)
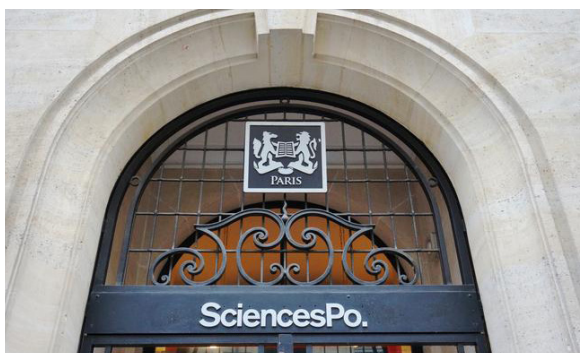
LIEPP is an interdisciplinary research center for public policy evaluation; why do you think that an interdisciplinary approach is so important for policy evaluation?

LIEPP is an academic initiative with recurrent public funding (the French Labex programme) which gives us the luxury to choose our own projects, hire assistant professors, and finally to finance our own research programmes. The *I* (standing for interdisciplinarity) in LIEPP is important for many reasons. Let me start with the one reason we thought was not a good one. We did not think that there should be a quest for interdisciplinarity just for itself. Researchers need to practice their own field and get evaluated by their peers, as economists, sociologists or political scientists. I would be concerned by transdisciplinarity if borders between disciplines are abolished: it's already hard to evaluate the academic quality of good work in one field, so it would be even more difficult to evaluate it in two or three different fields.

Instead, the good reason for interdisciplinarity is that, when cross-examining policy programmes or reforms, one single look would generally not be enough. It is often the case that economists carefully run regressions in randomised experiments or

A very careful description of the context of a policy by actors coming out of careful sociological work is not sufficient for policymakers: they need a quantitative answer

quasi-experiments and establish a robust result, but cannot always explain all their coefficients and all the mechanisms, or have difficulties in establishing the external validity of their results. On the other hand, it is also often the case that very careful description of the mechanisms, of the context, of the appropriation of a policy by actors coming out of careful sociological work, is not sufficient for policymakers: they need a quantitative answer. We therefore thought that having two or three different approaches to a single policy question would be needed: each approach would complement the others. When the conclusions of each approach are similar, we have more confidence in the result. This is what came out, for example, of the **evaluation by LIEPP of the CICE**: researchers from different fields concluded the programme had a very limited impact during its first few years. When the conclusions differ strongly or simply

diverge slightly, we can be more cautious in our policy prescription: this is a school of modesty.

How does LIEPP ensure an interdisciplinary approach through the evaluation process? Do you experience any difficulties?

This is not easy because of the contradiction in terms: we want people to publish in their own field to achieve excellence, and also to talk or collaborate with others to cross over expertise and methodologies. However, this is a problem that can be overcome: social scientists are passionate about their topics and like to learn from each other. Nevertheless, it calls for strong incentives: LIEPP, for instance, incentivized by providing more money to interdisciplinary teams. It also requires the creation of a common language. Let me give you an example: sociologists, and to an even greater extent political scientists, think for instance that the concept of "optimal policy", so familiar to economists, is at best naïve and in the worst case a dangerous concept, because it seems to forget that a policy is often (if not always) a way to arbitrate between diverging interests. Once we grasped this reciprocal misunderstanding, we could go further. In my view there is no contradiction between us: an optimal policy is a normative concept which is not a good description of what is actually implemented. Instead, we invented the concept of the policy circle: some social scientists have a view of what would be preferable for society, let's call it a policy ideal X0, and they may

actually put a policy proposal X1 up for public debate that already drifts away from that ideal, because they know that X0 would not be feasible. This idea passes on to a platform in a party and becomes X2; what is eventually voted for is X3, after many compromises; X4 is implemented by the administration because it controls the application and the decrees; the social scientists evaluate X4, find it does not work well, or could be improved if it had been closer to X0/X1, or more rarely find that X4 works well and change their own view of X0. In each stage the policy evaluation triangle, the *ex ante* methods, the monitoring and the *ex post* methods, all play an important role. We discussed these ideas with Cornelia Woll in 2011 to develop LIEPP.

Academia is key but not enough. We cannot always answer the questions asked: either the data are not available or they do not even exist

Your Center is at the Sorbonne; what should academia's role be in order to foster evaluation and its utilization?

We are part of Sciences Po, itself part of a larger group of universities called Université Paris-Sorbonne-Cité. This is different from Paris-Sorbonne. The French like to make things complicated and then complain that nobody understands. More seriously, academia is key but not enough. It's key because we have time to evaluate projects and the independence to make points that may be critical of the action of the current or previous governments. However, we cannot always answer the questions asked: either the data are not available or they do not even exist, or there is no good econometric identification of the impact of the policy. In contrast, the

administration often has access to the data and knows more about them. However, the administration is sometimes under pressure not to publish results that would be too sensitive in the political context as the administration is supposed to be neutral. But this is not black and white: academics may be politicized. This is why nobody can have a monopoly of the evaluation of public policies. Different viewpoints are useful and should be compared.

I would also add that having public funding is key: as I said before, this allows us to choose the agenda. We cannot rely exclusively on external funds to pay for our staff and researchers. If as a director I needed to accept contracts with conceptual difficulties (lack of data, lack of identification, lack of freedom to publish

Nobody can have a monopoly of the evaluation of public policies. Different viewpoints are useful and should be compared

and so on) to make sure that the center survives, I would be in a weak position. And we cannot allow this to happen: the stakes are too high. So, a budget of say 75% public funding and 25% from contracts seemed to us the best compromise. We were also lucky to obtain several external contracts on our own terms, and with great partners such as *Societé du Grand Paris* (SGP). SGP gathered a world-class scientific committee and gave us three years to evaluate the impact of the Grand Paris Express (an additional 200 km of metro lines). We developed new theoretical tools and new datasets for them, and delivered interesting new results that we expect to publish in academic journals. Our results have already contributed to policy debates and media coverage while also being presented in many academic departments, from London to Barcelona and from Berkeley to Vancouver.

What are the main barriers for effective use of evaluation?

I'm tempted to say that we cannot claim to have the truth. We're a bit more than suppliers or sub-contractors of policymakers, but we are not the ones who decide. In a democracy, the final word goes back to Parliament and the government. So as academics, we only need to gain in credibility, defend our independence, play fair, not reach firmer conclusions than what the data really tell us, and perhaps limit our media exposure to what is necessary to improve the public good. We are not politicians, or if we become politicians, that should be clear for people listening to us

in the best interest of science's credibility and independence. Membership of political parties or philosophical clubs should be made public along with our financial interests. I have the feeling that we as economists have got better in the second aspect but not really in the first one.

What are your recommendations in order to enhance the role of evidence in policymaking?

As a quantitative economist, I strongly believe that evidence-based policy is key. However, over the years I have also come to think that this is not enough; the dialogue we fostered with sociologists, especially non-quantitative sociologists, and political scientists showed that we can be less naïve when faced with other researchers. Again, this lends greater credibility to our recommendations, which can only be good for evidence-based policy. A last dimension is that somebody should take a serious look at what experts have said over the decades, and an independent ranking of the credibility of experts may be a great complement to their academic curriculum. Both are key ingredients of good policy evaluation, and sometimes judgment and common sense are needed, something which cannot be evaluated by academic journals. ▌

As economists, the dialogue we fostered with sociologists and political scientists showed that we can be less naïve when faced with other researchers

Interview with Geoff Mulgan,

Chief Executive Officer, Nesta (United Kingdom)

Geoff Mulgan is Chief Executive of Nesta and has been in post since 2011. Under his leadership Nesta has launched a range of new initiatives in investment, programmes and research, and has implemented a new strategy involving partnerships with foundations, governments and companies in the UK and internationally. Nesta is an innovation foundation and its priority fields are health; education; the creative economy, arts and culture; innovation policy and government innovation.

"In every public sector we now need to ensure that we are constantly gathering data to find out in real time if there are surprising patterns. If a teaching method or a welfare policy which we thought would work finally doesn't work, let's ensure we are able through data to spot that quickly so we can change and adapt."

"There is a big culture change in some governments which is saying that instead of introducing a policy at the level of a whole nation, it may often be more efficient to try it out on a small scale before you take it to a large scale."

"The commitment to serious evaluation goes in cycles in governments and in the last few years in many countries there has been big pressure from austerity and financial crises. It means that politicians have seen less value in evaluation."

"Evaluation is perhaps more and more away from being a big fact report which comes at the end of the project to be based on how day-to-day public sector management works."

"Governments need many more data scientists, computer scientists, people who understand what can be done with digital tools. But there are also other missing skills, and design methods are still very weak in many governments."

"We need quite a comprehensive overhaul of what it means to be a 21st century public servant and most of these things are learnt not sitting in a classroom but doing things and then learning by the experience of doing with the help of a specialist professional."

"Many public servants have a quite negative self-image and yet good public service has done more for human wellbeing than anything else. We need to restore our confidence, our faith in the vocation of public service."

Good practices

🌐 Commission on Evidence-Based Policymaking (United States)
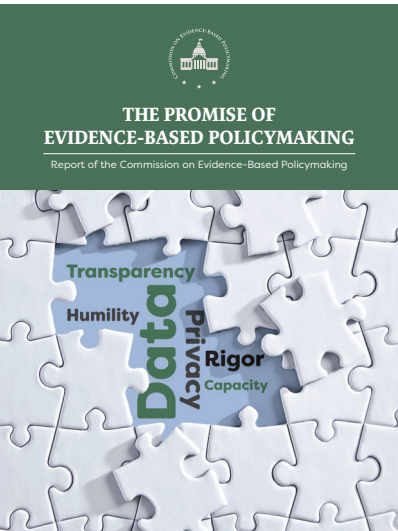
The **Commission on Evidence-Based Policymaking** was established by the Evidence-Based Policymaking Commission Act of 2016 passed by the United States Congress and signed by President Barack Obama on 30 March 2016. Over a year the Commission was tasked with studying how data, research and evaluation are currently used to build evidence and designing a strategy to promote evidence-building in the country. In September 2017 the Commission presented its **final report** in which it set out a series of recommendations in three key areas: improving secure access to data, improving privacy and transparency about the uses of data for evidence building, and enhancing institutional capacity to support evidence building. Measures of this third kind include establishing a Chief Evaluation Officer in each department to coordinate evaluation and developing multiyear learning agendas for each department to plan the evaluations and research to be performed. The strategy's purpose is to efficiently create rigorous evidence as a routine part of government operations and ensure it is used to construct effective public policy.

Swedish Agency for Health Technology Assessment and Assessment of Social Services (Sweden)

The **Swedish Agency for Health Technology Assessment and Assessment of Social Services** (SBU) was founded in 1987 by the Swedish government and tasked with assessing healthcare and social service interventions from a broad perspective covering medical, economic, ethical and social aspects. The SBU is one of the pioneers in Health Technology Assessment (HTA). Initially responsible for assessing medical treatments, in 2015 its mission was expanded and it also took over assessing measures used in the social services. The definition of health it works with means that the SBU covers prevention measures, such as strategies aimed at promoting physical activity and the influence of the workplace environment on health. The SBU's main objective is to provide useful information to public decision-makers and practitioners in social and health services. It does this largely by drawing up and publishing systematic literature reviews and analysis of reviews carried out by other institutions (such as the **Cochrane Collaboration** and the **Campbell Collaboration**) to confirm that their conclusions fit into the Swedish context. In addition the SBU also has a quick query service for responding to specific requests from practitioners. Finally, another of the SBU's aims is to identify knowledge gaps so as to guide practitioners and public decision-makers and also research funding institutions. The SBU currently has a permanent team of 65 people and a large network of collaborators involved in drawing up its systematic reviews.

Study of the Management Skills of People in Public Management Positions (Catalonia)

The General Sub-directorate for Research and Training in Senior Management at the **Public Administration School of Catalonia** (EAPC) has sponsored a study by the **Catalan Institute of Public Policy Evaluation** (Ivàlua) on the skills of people in public management positions in Catalonia. The study was drawn up in 2017 to identify the areas in which the skills of public managers need to be improved so that starting in 2018 the EAPC can map out appropriate training strategies to address them. The study points out that the areas that need to be improved most are strategic planning, evaluating public policy and recognising and rewarding the teams' effort and working in networks. This is the first time that a systematic study of the training needs of people in public management positions based on their competency profile has been carried out in Catalonia. The initiative comes in response to the specification of the professional framework for the managerial function with the aim of building more efficient government and will include a new Management Training Plan. General competence, main roles and cross-cutting skills are some of the issues addressed. Updating of professional skills, strategic vision, results-orientation, identification with the organisation and negotiating skills are picked out as some of the aspects required.

⊕ BetterEvaluation (Australia)

BetterEvaluation is an international initiative to share a large number of practical resources on methods, perspectives and applications with evaluation practitioners and other stakeholders. It is a great window onto the world of evaluation. It supports people to apply new knowledge to their particular situations and to share their experiences in order to improve processes. Its primary platform is its corporate website which went live in 2012 and its core team is based at the **Australia and New Zealand School of Government** (ANZSOG). Its contents include an interactive guide on evaluation for public managers setting out the various steps that should be followed, a resource bank, recommendations about which evaluation method is best for each need and the latest developments in the sector. It also offers readers a compilation of thematic areas that may be useful before conducting an evaluation process such as impact evaluation, gender analysis, feminist evaluation, evaluation of policies addressed to childhood and evaluation of an organisation's performance. In five years of operation it has had more than 3.8 million web visits and has almost 7,000 collaborating members around the world.

The What Works Network (United Kingdom)

This initiative aims to improve the way government and other organisations create, share and use high quality evidence for decision-making. It supports more effective and efficient services across the public sector at national and local levels in the United Kingdom. What Works is based on the principle that good decision-making should be informed by the best available evidence. If evidence is not available, decision-makers should use high quality methods to find out what works. The network is made up of 7 independent What Works Centres and 2 affiliate members. Together these centres cover policy areas which receive public spending of more than £200 billion. What Works Centres are different from standard research centres. They enable policymakers, commissioners and practitioners to make decisions based upon strong evidence of what works and to provide cost-efficient, useful services. The centres help to ensure that thorough, high quality, independently assessed evidence shapes decision-making at every level by: (1) collating existing evidence on how effective policy programmes and practices are; (2) producing high quality synthesis reports and systematic reviews in areas where they do not currently exist; (3) assessing how effective policies and practices are against an agreed set of outcomes; (4) sharing findings in an accessible way; and (5) encouraging practitioners, commissioners and policymakers to use these findings to inform their decisions. The Centres are funded by a combination of government and non-government sources.

Campbell Collaboration (Norway)

The **Campbell Collaboration** is an international network which produces systematic literature reviews and other evidence syntheses about the effects of social interventions to foster evidence-based professional practice and public policy. It conducts evidence reviews in justice, education, international development and social welfare. In addition to producing these reviews its Knowledge Translation and Implementation group seeks to ensure that the systematic reviews are used by policymakers and researchers. One of its key actions is drawing up Policy Briefs written in plain language which summarise the findings of the literature reviews in an accessible way. With the vision of "Better evidence for a better world", the foundation was established in 2000 and has offices in Oslo (Norway) and New Delhi (India). Its mission is to promote positive social and economic change through producing and using systematic reviews for evidence-based public policies and actions. This enables the Campbell Collaboration to contribute to the implementation of better informed decisions with a higher degree of effectiveness for public and private services worldwide.

New trends

🌐 Randomised Controlled Trials

A **Randomised Controlled Trial** (RCT) is an experimental form of impact evaluation in which the population receiving the programme or policy intervention is chosen at random from the eligible population, and a control group is also chosen at random from the same eligible population. It tests the extent to which specific, planned impacts are being achieved. The distinguishing feature of an RCT is the random assignment of units (e.g. people, schools, villages, etc.) to the intervention or control groups. One of its strengths is that it provides a very powerful response to questions of causality, helping evaluators and programme implementers to know that what is being achieved is as a result of the intervention and not anything else. An RCT measures the effect of a programme or policy intervention on a particular outcome. The key feature of an RCT is that it uses random assignment of an intervention. This design is called an experimental design. An RCT is only useful for measuring impact in certain scenarios such as when a large sample is available; the intended impacts of the programme or policy intervention can be readily agreed and measured; and the RCT is planned before an intervention begins. To be successful it's important that the programme is well defined and has clear objectives. Data collection and analysis should be based on a strong theory of change, allowing the evaluators to assess the pathways through which the intervention will be successful in achieving its objectives.
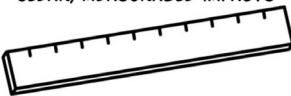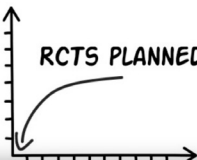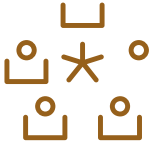
New avenues for partnership between evaluation agencies and governments

Government is constantly innovating to tackle innumerable social problems and challenges. However, very few of these projects end up being evaluated and hence their effectiveness is unknown. To foster rigorous impact assessments through Randomised Controlled Trials (RCT) and thus increase the corpus of evidence, new forms of partnership are emerging between evaluation agencies and government. This is the case of the **J-Pal State and Local Innovation Initiative**, which after **two years of learning** has become firmly established.

This initiative supports US state and local governments in using experimental design evaluations to generate new and potentially widely applicable lessons about which public policies work, which work best and why. The initiative consists of a competitive competition for local and state leaders in the United States who are interested in designing and implementing experimental design evaluations and using the evidence generated to inform their decision-making. In 2016 five governments were assisted and in 2017 three more joined the initiative. These governments are currently partnering the J-Pal North America team in implementing **experimental design evaluation projects** in areas such as drug addiction treatment and intervention with homeless people. The winners of the third competition will be announced in the summer of 2018.

Data Labs

In many countries, social intervention programmes are outsourced to charities or private companies. These organisations and governments alike have an interest in evaluating the impact of the interventions they perform. However, this involves accessing the government's data without compromising data protection and also means they need to have the skills required to do so. Data Labs are being launched in particular in the United States and the United Kingdom to cater for these two requirements.

Data Labs are institutions formed by teams of data analysts reporting to government or who work directly for it and they enable administrative data to be used for research and evaluation purposes. In the United Kingdom the British organisation **New Philanthropy Capital** has been sponsoring the creation of Impact Data Labs since 2011. The first to begin operating was the **Justice Data Lab** run by the Ministry of Justice. It started out as a pilot in 2013 and became a permanent service in 2015. The initial aim was to help charities working with offenders to evaluate the impact of their programmes. Up to October 2017 the Justice Data Lab had produced 167 reports for charities, social enterprises, other private organisations and public entities. These reports are **posted on its website** and feed the corpus of evidence about what works and what doesn't in this intervention area. New Philanthropy Capital is continuing its efforts to create new Impact Data Labs in **education**, **employment** and **health**. Drawing on what they have learned, they have published a **paper** featuring useful ideas for anyone who looking to set up Data Labs.

Improving the accessibility of statistical data

Governments create huge amounts of data which have enormous potential for generating knowledge in basic research and also for evaluating public policies. Difficulty in accessing government databases for evaluation purposes is one of the common barriers to doing this and more and more governments are putting measures in place to tackle the problem.

Accessing government data in order to evaluate policies has been easier in Catalonia since last year. **The Statistics Plan of Catalonia 2017-2020 Act 5/2016, of 23 December, and amending the Statistics of Catalonia Act 23/1998**, includes for the first time the objective of promoting the use of official statistics for research in social sciences and the evaluation and improvement of public policies. Accordingly the Act urges the Statistical Institute of Catalonia (Idescat) and the institutions and agencies in the Statistical System of Catalonia to promote and facilitate the use of available statistical information to improve the evaluation of government public policies in Catalonia and also to monitor and programme policy cultures to be implemented.

Scientific research institutes, researchers, research centres and public law entities engaged in designing, planning, programming, monitoring and evaluating public policies are allowed to access confidential data for scientific purposes protected by statistical confidentiality as long as people cannot be directly identified from the data in compliance with data protection and statistical confidentiality legislation. This is the first law to enable government data to be used for public policy evaluation in Spain.

The public's evaluation of public services

New technology provides great potential for future developers, organisations and the general public when it comes to overseeing government action. More and more tools are available to evaluate the effectiveness and impact of the resources invested by government in public services and to determine whether or not to maintain them. One example – as described in issue 2 of the EPuM – is the *gencatgram* app sponsored by the Catalan Government which records all the data and activity of digital channels in the same environment. It will contain data from social media, mobile apps, websites, procedures, messaging and any other public service channel. In addition to acting as a file and directory, this database has open access to data in order to build other products via API (Application Programming Interfaces), a module that allows one computer programme to communicate with another one. Examples include command and control dashboards, investigation and research data sources, widgets and other website components.

The first product is the social media activity website which the Catalan Government has set up to open all the data it has on its social media. This first service illustrates how *gencatgram* is an innovative and disruptive app. In other words, it exemplifies the uses, products and services which can be developed with *gencatgram* to evaluate these resources, help digital channel managers improve their presence and learn from the ones that work best.

Evaluating the impact of public policies

The **France Stratégie** ideas laboratory aims to evaluate public policies, anticipate future economic, social and technological changes, confer with French and international experts and players and put forward recommendations to national, regional and European public authorities. In the field of evaluation it sponsors seminars and content in partnership with other research centres and institutions. In 2016 it published the user's guide for managers and practitioners **"How can we evaluate the impact of public policies?"** (in French) where it explains how to measure the impact on beneficiaries; how to generalise the results of an evaluation; how to explain the success or failure of a policy; how to compare the effects of different policies; and how to move from evaluation to recommendation. The publication notes how efficacy and efficiency have become crucial factors for most developed countries given the tightening of public budgets and the public's growing distrust of institutions. The guide shows how a public policy which from the outset provides for measuring its effectiveness gains in legitimacy, although objective and transparent evaluation methods are needed as well. The authors conclude that the results of evaluations must be thorough yet also understandable for decision-makers and public opinion. The promoters' impartiality and the evaluators' independence also have to be ensured. The France Stratégie portal additionally features other useful content about evaluating public policies.
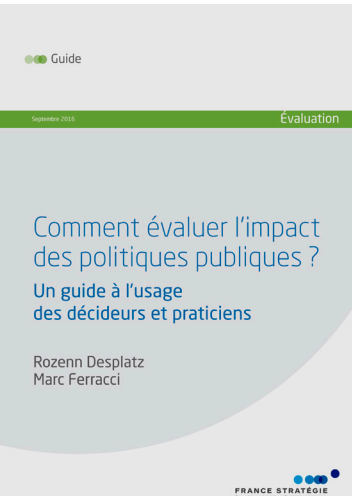
Newsflash

⊕ Barcelona, candidate for a European 5G technology laboratory

5G technology will lead to a technological shift and the creation of smart environments and advanced industries. It will provide access to the Internet of Things, opening the way for the generation of new disruptive services in areas such as autonomous and connected vehicles, drones, Industry 4.0 and access to remote services. Catalonia is at the forefront of this new technology. Since the start-up in 2014 of the 5G PPP initiative, part of the European Commission's H2020 programme, several research centres have obtained funding to take part in 22 of the 37 awarded projects. This figure confirms Catalonia's international leadership in the 5G ecosystem. **5GBarcelona** will establish a European 5G Digital Hub based on an open experimental infrastructure in the metropolitan area which will serve as an urban, citizen-led and technological laboratory for validating 5G technologies and services. It will be an open innovation environment based on teamwork between public and private players which will generate synergies with the **Mobile World Congress** (MWC). In 2018 the European Commission is to select a small number of projects in order to create 5G testing and validation environments and Barcelona hopes to host one of them.
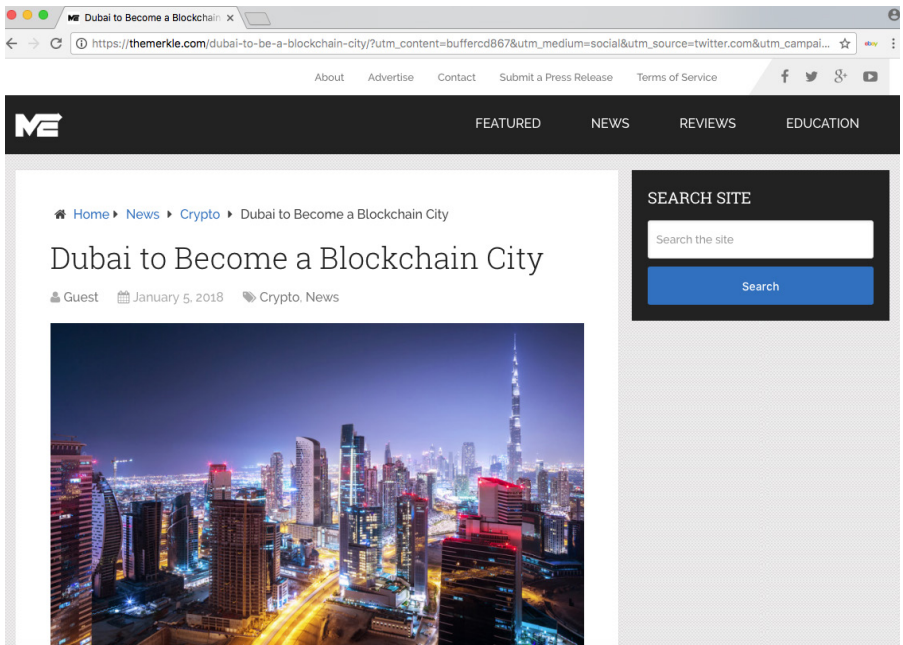
Digital Local Government Congress 2018

On 21-22 March the Centre of Contemporary Culture of Barcelona will host the 3rd **Digital Government Congress**, a leading forum in Catalonia for digital government arranged by the **Open Government Consortium of Catalonia** (AOC) and the **Localret** consortium with the support of the Public Administration School of Catalonia and other organisations. This time round the programme looks at the local level and more specifically at the digital transformation of local authorities' dynamics, processes and ways of working in order to meet the new challenges posed by 21st century society. The idea is to promote a platform for exchanging experiences and transferring knowledge to highlight local government's leadership and efforts towards achieving an open and innovative society. The organisers also intend to share examples of inter-governmental partnerships which foster the principles of transparency, participation and accountability. Over 1,000 people are expected to attend and the programme will cover five main areas: (1) generating savings, simplification and wealth; (2) innovation in new disruptive public services; (3) designing citizen-centred services; (4) evaluating impact on society; and (5) generating trust and transparency.

⊕ Dubai and Greece join blockchain technology

Dubai in the UAE is emerging as the world's first government based on blockchain technology. Its leaders forecast that by 2020 most municipal documentation will be powered by this type of algorithm including visa applications, bill payments and licence renewals. It is about doing away with much of the bureaucracy government generates by fostering financial and human resources savings. The real possibility of promoting a local digital currency, similar to the Bitcoin but more appropriate for an Islamic culture, is also on the table. Blockchain is proving to be a feasible technology which can meet many of the challenges currently faced by governments worldwide. One example is **accreditation of university qualifications**: three Greek universities are driving an open source project through which their graduates can get their degree certificates electronically. To date there has only been the option of certified photocopies, but this allows for the possibility of fraudulent or fake documents. Using the new technology increases transparency, the certificates' reliability is guaranteed, a large amount of paperwork is eliminated and the graduates themselves are responsible for administering their records, thus cutting bureaucracy and workload at the universities.
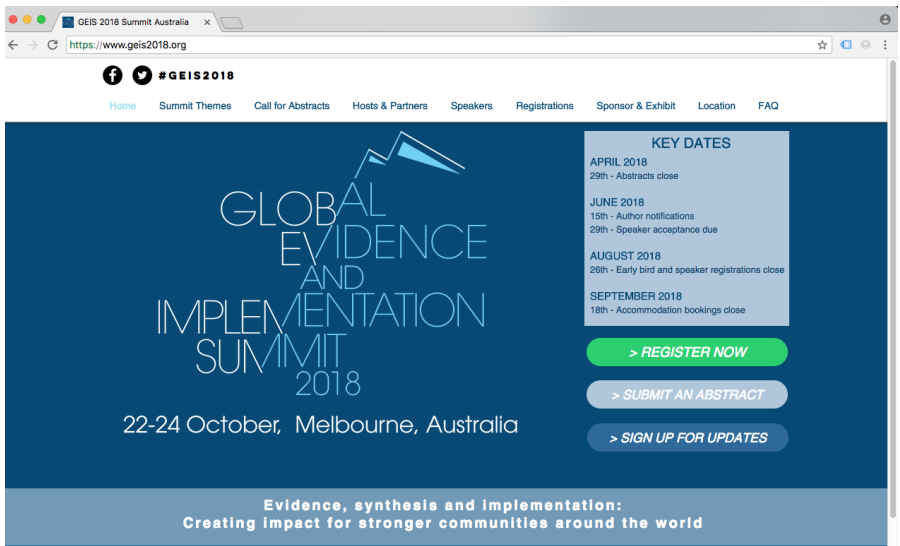
⊕ Quality of public administration. European Commission

The European Commission has recently published the **"European Semester Thematic Factsheet. Quality of Public Administration"**. This publication is about the horizontal aspects of the functioning of public administration. It looks at achieving results and improving accountability, policymaking, structures and processes, human resources and service delivery. The key challenges detected are: (1) Achievement and limitations of recent reform efforts; (2) Executive capacity; (3) Developing employee potential in public administration; (4) Quality of public services; (5) Online service delivery; (6) Open data; (7) Public administration in times of fiscal consolidation; and (8) Public administration and societal challenges. Other recommendations and opportunities are strengthening multi-level governance, encouraging intermunicipal cooperation; streamlining and simplifying processes; attracting new recruits; enabling mobility; creating stimulating workplaces and creating a broader framework for performance management. The factsheet also includes other useful resources and case studies and sets out the main societal changes and resulting challenges for public administration which are: (1) Globalisation; (2) Demographic change; (3) Climate change; (4) Technological change; (5) Economic trajectories; and (6) Public trust in government.

🌐 Global Evidence and Implementation Summit 2018. Melbourne

The **Global Evidence and Implementation Summit 2018** (GEIS) organised by the Centre for Evidence and Implementation and the Campbell Collaboration is one of the year's big events in evidence building and using it for better public policies. It is to take place in Melbourne, Australia, on 22-24 October and will bring together experts from across the world to share their experiences in building and implementing evidence to improve public policies. The summit will explore case studies in the design, development and effective evaluation of public programmes and services. Expected to reach 700 delegates from 30 countries, GEIS 2018 will share the latest research and strategies for improving the lives of individuals, families and communities worldwide. Its programme is divided into five major thematic areas: (1) Understanding what works; (2) Achieving scale and sustainability; (3) Methods for impact and implementation evaluation and synthesis; (4) Using evidence for better policy, programmes and practice; and (5) National and global cooperation and partnerships.

⊕ Digiwhist, European big data against corruption

Digiwhist is a European big data project for detecting fraud and fighting corruption. It is a tool for processing indicators and public data that works in partnership with civil society organisations and seeks synergies with institutions to improve its usability and the quality of the information it provides to users. One of its key areas is public procurement, and after three years of research it includes 7.66 million records from across Europe. The Digiwhist consortium is led by the University of Cambridge and has investigated the documents available in 35 jurisdictions (the 28 EU member states, Norway, the European Commission, Iceland, Switzerland, Serbia, Georgia and Armenia) in order to build a powerful **public access database**. The project has also set up the **Opentender** portal to make public tenders more transparent with the option of including indicators about the efficiency of procedures and creating numerical information displays. Digiwhist's resources include accessing and comparing public European regulations and mechanisms (financial disclosure, conflicts of interest, freedom of information, public procurement and political financing) and making it easier to track European public tenders. The project was presented on 29 January in the European Parliament under the title **"Big data with local impact: Using open data to improve public procurement"**.

European Public Mosaic (EPuM).
Open Journal on Public Service

Free subscription

EP_uM
BACK ISSUES

EP_uM

European Public Mosaic

OPEN
JOURNAL
ON PUBLIC
SERVICE