Qualitative Evaluation of Academic Careers in Computer Science at CNRS

Pierre Senellart

Global Forum on Development of Computer Science, Tsinghua U., 12 April 2025

PSL University

PSL University (Paris Sciences et Lettres) is one of France's leading research-focused universities comprising smaller institutions, in the heart of Paris:

- ENS – École normale supérieure | PSL
- Dauphine–PSL, Mines Paris–PSL, Observatoire de Paris, and others
- Over 17 000 students ($\frac{2}{3}$ of graduate students) and 2 900 researchers

In Computer Science:

- Internationally renowned ENS Computer Science Department (DI ENS), joint between ENS-PSL, Inria, and CNRS
- Strong research teams around theoretical CS, AI/ML, cryptography, software verification, etc.

CNRS

CNRS (France's national center for scientific research):

- The largest fundamental science agency in Europe
- A multidisciplinary institution covering many fields of science and humanities
- Employs over 11,000 tenured researchers and 13,000 engineers and technicians throughout France (and even some abroad)
- CNRS researchers have full-time research-only positions, but may be working in research units embedded in universities

Evaluation of CNRS researchers

CNRS researchers get evaluated:

- when they get hired (as civil servants, with tenured positions)
- every 2.5 years based on their past activity
- when applying for a promotion or for a bonus
- every 5 years, together with their research unit which is collectively evaluated

CoNRS

CoNRS (France's national committee for scientific research):

- The evaluation and advisory committee of CNRS
- Acting independently of the direction of CNRS
- Majority of elected members (by and among the whole French academic community, not just CNRS); Minority of members appointed by the French Ministry of Higher Education and Research
- Divided into:
  - 11 scientific advisory boards
  - 41 sections covering a specific area of research, including sections 6 (theoretical and core computer science) and 7 (some numerical fields of CS + signal processing, control, robotics)
  - 6 multi-disciplinary commissions

This talk

- I was an appointed member of CoNRS's Section 6  from 2016 to 2021; and have been its elected president since 2021
- The section's 18 academic members spend a significant portion of their time every year evaluating researchers (for hiring, promotion, or just regularly scheduled evaluations)
- We strive at answering these key points:
  - What is expected of academic researchers?
  - How best to evaluate researchers in a qualitative manner, covering all aspects of what is expected of them, taking into account differences across fields
  - How best to make decisions?
  - What should be the outcome of the evaluation?
- All of this taking into account the specifics of computer science!
- This talk: my perspective on these topics, informed by my experience within CoNRS and our practices

Outline

What is an academic researcher expected to do? (1/2)

- Produce innovative, relevant, deep research
- Disseminate this research through publications, software, benchmarks, datasets, etc.
- Supervise students and other junior personnel (especially PhD students) to train the next generation of researchers
- Participate in the animation of research at the national and most importantly international level through program committees, editorial boards, scientific expertise reports, etc.
- Collaborate with other researchers, set up and coordinate collaborative research projects, make bridges to other subfields and scientific communities or through interdisciplinary research
- Transfer the research to leave an impact on the society at large, through industrial partnerships, start-ups, standardization or policy actions, or popularization of science to the general public

What is an academic researcher expected to do? (2/2)

And especially for senior researchers:

- Participate in the management and administration of research at a local level (team or lab leadership, responsibilities within a university, etc.)
- Participate in the management and administration of research at a national and international level (structures federating research in a given area, research evaluation bodies, grant agencies, academic and professional societies, etc.)

Specificity of CNRS: no teaching requirement! (but may be also valuable, especially for advanced-level courses as a way to disseminate the research)

Outline

How to evaluate researchers and research?

Distinguish between:

- Research itself, research output, scientific visibility
- Other activities (supervision, animation, transfer, administration, etc.)

The San Francisco Declaration on Research Assessment (DORA)

- Launched in 2012 to improve research assessment practices
- Opposes the misuse of journal-based metrics (e.g., impact factor)
- Encourages evaluating research on its own merits
- Numerous academic institutions are signatories, including CNRS

Let's change
what we value
in research.

Sign
DORA

What we do not do

- We do not use bibliometrics (number of citations, h-index, impact factors, etc., or even raw number of publications): can be gamed (and often are), impossible to compare across areas of research, unreliably computed, etc. And more importantly: do not accurately measure the quality of research.
- We do not use AI/LLM systems: leakage of private and sometimes sensitive content, somewhat poor reliability, no accountability
- We do not pretend to be fully objective, the decision is that of the committee

What we should theoretically do

- Disregard publication venues (or even publication status) altogether
- Have one or multiple experts read the entire research output of the evaluated individual, and make an informed decision on its quality

Unrealistic in practice:

- Does not scale, too time-consuming
- Sometimes the committee does not include an expert in the specific area of research (or sometimes that expert is in conflict of interest!)

What we actually do

- Ask for a reasonably-sized and custom-written report on past research (and on future plans) and do read it in full
- Sample some of the works of the researcher (e.g., one publication per reviewer) that will be read in depth
- Have technical interviews (for important evaluations) where we can go deeper into one specific contribution
- Still occasionally use the prestige of a conference/journal as a proxy for a publication within (benefiting from the expertise of the reviewers of the paper in this way), but not by doing raw counting
- Awards play a similar role as a proxy for evaluation, amounting to trusting the committee who granted the award
- When available (e.g., some ML or NLP conferences on OpenReview), also read referee reports about publications

Dimensions in qualitative evaluation

- Coherence of research agenda
- Originality and long-term vision
- Autonomy and leadership
- Technical depth and complexity of research contributions
- Diversity of contributions

Important: We take into account:

- Career stage
- Career interruptions or changes
- Disciplinary or subdisciplinary norms
- Constraints of the research environment

This remains subjective! So important to have a diverse committee, with different background, expertise, and sensibilities!

The CRediT Taxonomy of Contributions

CRediT = Contributor Roles Taxonomy

- Conceptualization
- Methodology
- Software
- Validation
- Formal analysis
- Investigation
- Data curation
- Writing – original draft
- Writing – review & editing
- Visualization
- Supervision
- Project administration
- Funding acquisition

Helps attribute diverse roles in collaborative research works, disambiguating what being an author means

Research output is not just publications

Important to also take into account other forms of research outputs which can be used for evaluation:

Open-source software: if made available (e.g., through a Git repository), code can be read, documentation can be checked, software itself can be tested, the weight of each contributor can be assessed; publication artifacts may also have been evaluated at the same time as publications, esp. for repeatability

Licensed software: details of the licensing, its impact on the (e.g., industry) partner, need to be described to be evaluated

Benchmarks, datasets: are they easily reusable? are they reused?

Deployed Web site or platform: who are the users? what is the impact?

Beyond research output

For other activities:

- A somewhat quantitative approach may be appropriate: number of students supervised, number and volume of industry partnerships, workload imposed by admin duties, etc.
- But important to consider the impact of:
  - Research area (there will be more funding available in some "hot" areas; it is expected to have more students in an area where research requires large teams; graph theorists may have fewer opportunities for industry transfer)
  - Local environment: researchers should not be penalized because they have less opportunities in their immediate environment
  - Specific situations (e.g., a handicap preventing taking on certain responsibilities)

Outline

General principles

- Looking for consensus, leave room for everyone to express their opinions
- Avoid raw votes, prefer discussions and asking every member to speak in turn
- The committee is united and takes responsibility for all decisions taken and opinions issued
- Secret of the precise content of discussions
- Transparency whenever possible
- Diversity and representativity are points we care about (parity, thematic diversity, geographic diversity)

Conflicts of interest

- Conflicts of interest unavoidably occur: member of the same research unit, coauthor, etc., or even family member
- Important to define ahead of time what constitutes a conflict of interest in a precise way, and how they are dealt with (being silent about a researcher, withdrawing temporarily or permanently from the committee, etc.) depending on how serious they are

Outline

What should be the outcome of the evaluation?

For hiring, promotions: a partial ranking (often no need and even not appropriate to provide a total order among candidates as long as there are enough slots) accounting for the diversity of the candidates (theme, location, positioning on applied/theoretical axis...)

For regular evaluations: detailed written and hopefully useful feedback sent to the researcher; we write long reports explaining what we have understood of the dossier, strong points, and any recommendations that would help the researcher in his or her career; exceptionally, when the evaluation is negative, it may result in strong recommendations, more frequent requests for evaluation, and possibly in severe cases disciplinary actions

Our responsibility: provide constructive, individualized feedback and recognize excellence in its many forms

Outline

Conclusions

- CNRS offers a unique environment for career-long research
- Having an independent evaluation body is invaluable
- CoNRS emphasizes qualitative, fair, and field-aware evaluation, not based on numbers, on bibliometrics, on raw votes
- We aim to foster diverse and rich research careers in computer science
- We welcome continued reflection and discussion on how to improve evaluation practices

Questions?

Thank you!