



# Semantic analysis of CSTP reports using the IPP vocabulary

OECD Directorate for Science, Technology and  
Innovation

Andrés Barreneche, Alina Deniau, Michael Keenan, Blandine Serve

OECD Library and Archives

Frédéric Abrazian, Mary-Ann Grosset, Jan-Anno Schuur, Thierry Vebr



# Coverage

---

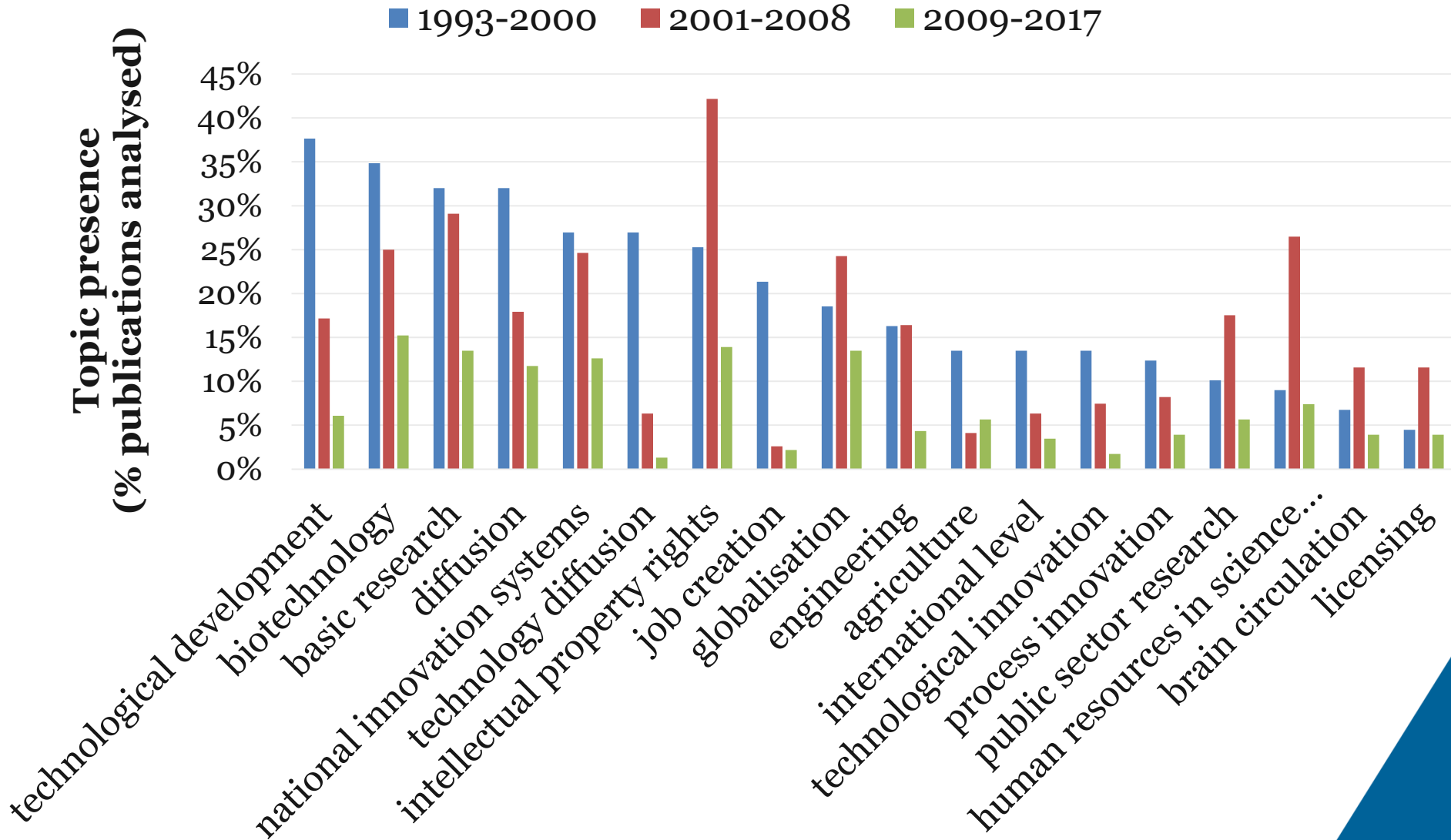
- This is the simplest of the approaches presented here, running the existing IPP vocabulary of 1200 terms against different corpora and presenting the results in simple bar charts and word clouds
- Analysis of the full CSTP corpus of substantive reports
  - 782 **full documents** over 1994-2016
  - Includes a subset of 100 TIP substantive reports
- Analysis of *Research Policy* articles
  - 2527 **title-abstracts** over 1988-2017





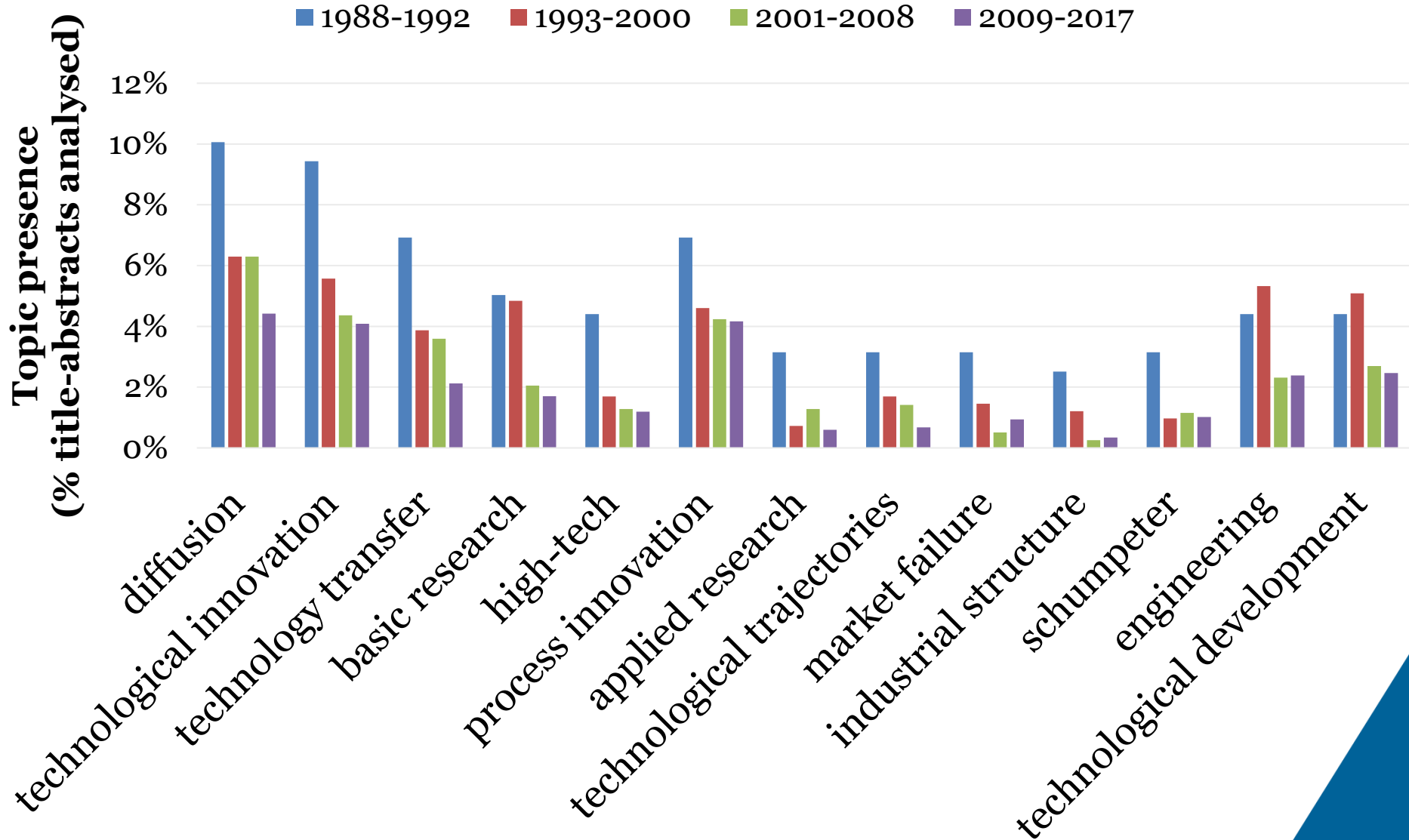


# CSTP: Topics becoming **less** frequent





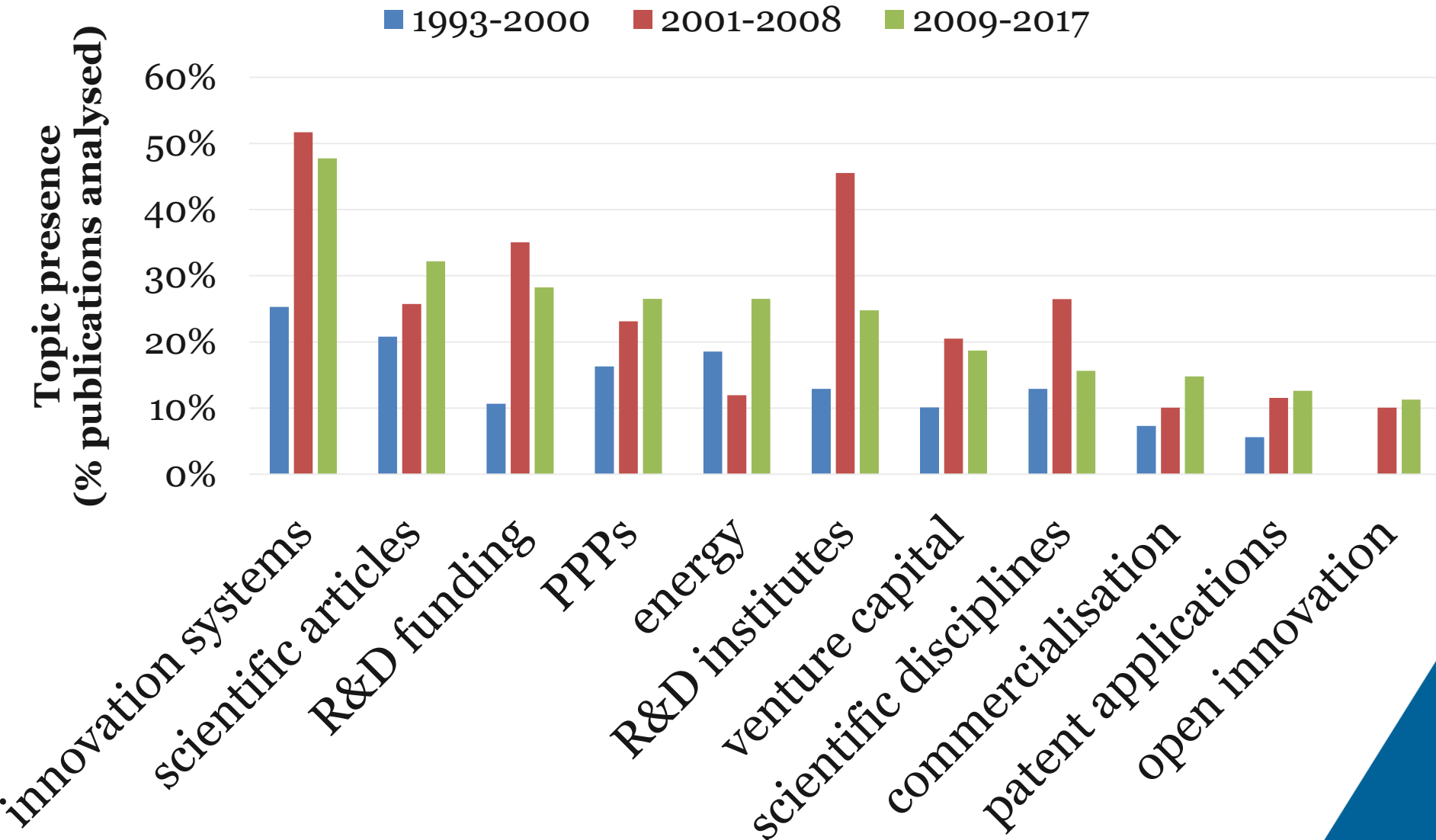
# ResPol: Topics becoming **less** frequent





# CSTP: Topics becoming more frequent

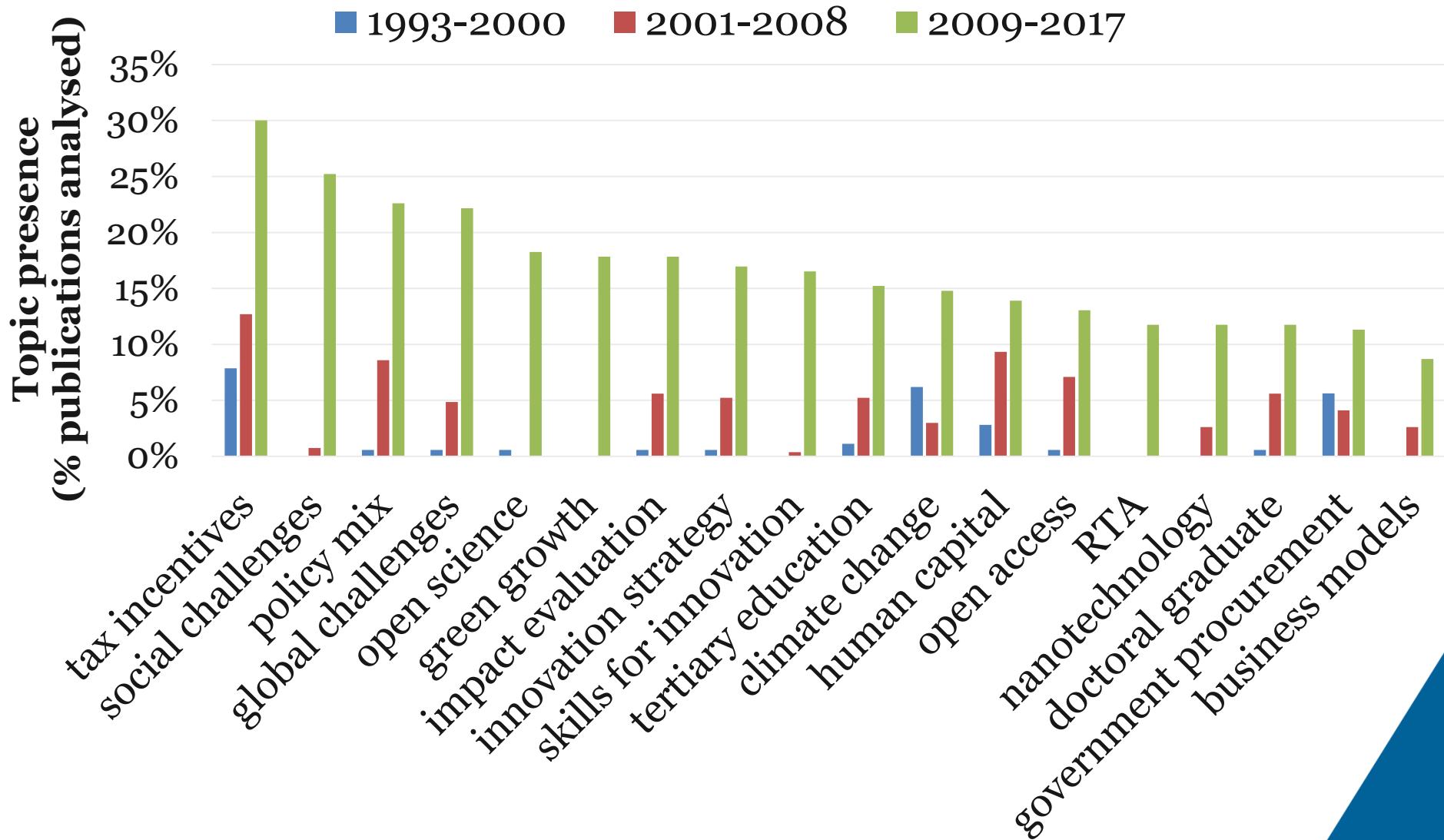
## Traditional core topics





# CSTP: Topics becoming more frequent

## Emerging topics







# ResPol: Topics becoming more frequent

