**Alan Dix - citation based analysis on REF UOA 11 investigating latent bias between sub-areas**

1. Citation based
   1. Assume citation-metrics can be used as a valid measure of quality between large enough units in computing
      1. Broad agreement
      2. HEFCE are using them to help ensure diff between subject sub-panels are defensible
   2. In sub-panel 11, data is particularly rich
      1. sub-area score table by Sloman more able to reflect on output-by-output results
   3. Citation data
      1. Raw Scopus citation data
      2. Normalised Scopus citation data
      3. Google scholar citations
2. Sub-areas
   1. The ACM Computing Classification System (CCS) is a subject classification system for computing devised by the Association for Computing Machinery (ACM). The system is comparable to the Mathematics Subject Classification (MSC) in scope, aims, and structure, being used by the various ACM journals to organise subjects by area.
   2. Divided computer science into several major fields - standardised
   3. 4\*/3\*/2\*/1\* profile available for ACM sub areas by Morris Sloman
   4. Trend in CS: more theoretical areas top the table with over 30% 4\* while more applied areas tend towards the bottom with often below 15% 4\*
3. Analysis
   1. Seven analyses:
      1. Scopus all years
      2. Scopus 2008-2011
      3. Normalised Scopus 2008-2011
      4. Google all years with no citations as missing value
      5. Google 2008- 2011 with no citations as missing value 2008-2011
      6. Google all years with no citations as present and zero
      7. Google 2008-2011 with no citations as present and zero
   2. Process
      1. For all paper in UOA 11, find the top 25% quartile of papers based on citation number data in each year
      2. For each sub-area, what is the % of papers in the top 25%, 25-50%, 50-75% and last 25% quartiles?
      3. Variants?
      4. Compare to 4\*/3\*/2\*/1\* data by Sloman
         1. Rank order of sub-areas in citation vs rank order in 4\*
4. Results
   1. Citation
      1. All work across different sub-areas should be evenly cited
   2. REF CS sub-area score
      1. More 4\* work in formal/theoretical areas (2-3 x citation prediction)
      2. Less 4\* in applied and human-centric areas (2/3 or 1/3)
   3. Correlation
      1. No discernible correlation or pattern
5. Conclusions
   1. Inter-area bias emerged
      1. Causes
         1. Due to distortions in hiring /strategic decisions within institutions
         2. Due to halo effects of good institutions
      2. Consequences
         1. Areas of more 4\* work will definitely get more funding
         2. Post-1992 universities have higher possibility in losing team than pre-1992 universities
            1. Old – post-1992 – theoretical areas – more funding
            2. New – pre-1992 – applied areas – less funding
   2. Effect of inter-area bias on institutional assessment
   3. Impact of gender bias between sub-areas
      1. Different M/F participation