

Industry Insights 8: Matt Forshaw

Steve Crawshaw

Developments in AI and Data Skills

Inconsistency in terminology makes it difficult for industry to identify valuable skillsets and unpick trends. Apprenticeships are valuable route into data science. Need to widen access in the field to make more inclusive.

Data skills taskforce

UK faces skills issues - can't be accessed by any single group so DST formed. Representation from across the field to address skills gap. What is the supply and demand question? Untapped potential from data in orgs - how to we realise this?

Understanding the UK AI Labour Market 2020 - report

“Dramatic need for expertise” at lower \ mid end. Demand for 240K roles in data \ AI in UK

Role ambiguity has a direct impact on practitioners.

Venn Diagram

- Maths and Stats
- CS
- Subject \ Domain Expert

“85% of big data \ DS projects fail” (Gartner) (due to culture, organisational readiness)

Newcastle programme focuses on business case development, pitches linking to industry.

Ethics

Opportunities in ethics space. Anaconda report “State of DS 2020”. Findings that only 15% of DS grads believed they had had ethics training in their course.

Need to give confidence to practitioners to uphold ethics standards in practice - maybe through chartered routes.

Algorithmic bias

Pre - trained models may be inherently discriminatory. Need to take steps to check while in development. E.g UK passport photo checker bias on skin colour. Should not have reached production.

notflawless.ai: Video of examples of missclassification of black faces - racial gender injustice.

Equity

Majority of roles at risk of automation are held by women. Office for AI offered degree conversion to AI. Focus on excluded groups through scholarships.

Professionalisation of DS

DS NGO's are coalescing around the DS topic to agree standards for the discipline - accreditation \ certification. **Alliance of DS Professionals** to agree shared definition for values, support for professionals to uphold values.

Questions

Q. How can AI visualise News data?

A. Generate content automatically (narrative in elections). Pandemic may have increased data literacy but can be problematic (e.g. log scale charts). Need to be wary of fake (Fox) news. Good examples of guidelines from BBC and FT (ggplot).

Q. Guidance on DS courses?

A. Review courses to see where they fit on the Venn diagram above.

Q. Professional Accreditation?

A. Possibly by the end of the year (2021)

Q. How is progress on improving access into DS?

A. 13 Universities making good progress in terms of diversity. Should help address bias.

Q. How to access developments for non STEM DS people?

A. Need to devote time to review industry white papers and track latest academic literature. But note potential bias. Need to “responsibly use the black box”. What are the assumptions and constraints?

Q. Is bias more constrained to individuals rather than the system?

A. Possibly but need to embed an approach which doesn’t reinforce or perpetuate this. Lots of work through Turing on explainability to address this. May need to sacrifice performance for explainability.
