Big data ethics

‘“Big Science” is defined as a substantial expansion of scientific collaboration along several axes: geographic (in the concentration of scientific expertise and technological capacities within cities or regions), economic (in the monetary sponsorship of major research endeavors, on the order of several billion dollars), multidisciplinary (in the necessary coordination of teams from previously distinct fields), and multinational (in the coordination of groups with very different research styles and cultures).’ (M. Christen, et al. (2016) Ethical Challenges of Simulation-Driven Big Neuroscience, A JOB Neuroscience, 7:1, 5-17)
Big science ->Big Data = Big knowledge?

Questions Big Data ethics

1. Make sure informed consent (information, material from humans) in projects of Big Data
2. Avoid dependency from IT companies, programs
3. Don’t connect personal and performance data (e.g. metadata to control access to facilities)
4. Reflect acquisition of knowledge using artificial intelligence: modeling, forecast, interpretation
5. Consider pressure by public, politics and economy through big data to research direction and interpretation (e.g. climate debate)
6. Digital sciences: loss of important information by digitalization (analytical vs. holistic approach)
7. Big Variety: can from specific results be made general conclusions?
8. Big Validity: effects of research results to other area (pervasive technologies); significance of metaanalysis, interpretation
9. Big data: difficulties to find the needle in the haystick (signal/noise ratio)
10. Pressure is higher: Risk of naturalistic fallacy (Can implies Ought)

Problems

1. Shared data: allocate scientific credits in publications
2. Waste of data resources: big data potential is underused
3. Data protection IP in collaborative science
4. Exchange of information between social spheres (health, economic, family, education)
5. Informed consent for usage of data: broad consent, stewardship (tracking) contributions
6. Data security and privacy issues across countries
7. Higher expectation, too big promises that undermine research credibility
8. Higher pressure to communicate positive outcome and to meet public expectations
9. Narrowing research directions to marketable application and neglecting subfields and collaborations
10. “Too big to fail or to manage” problem: can’t give up easily a big program without loss of reputation
11. Duty to share data and accountability are higher
12. Effective integration of different scientific cultures: procedures, methods, interpretation
13. Problems in reproducibility and speculative predictions may jeopardize funding for the whole field

Recommendations

1. Ensure ethical competence in big science projects and in departments
2. Scan proposals and publications for ethical issues
3. Establish a whistle-blowing policy
4. Avoid overselling, also by PR
5. Show weakness and strength in simulations
6. Enforce reproducibility and quality of data
7. Define color standards of visualizations
8. Cooperation agreements: include RI guidelines
9. Make clear responsibilities in big research projects
10. Consider privacy, security, surveillance and safety issues in research and applications