

# Open Access Data Sharing

Personal View (not legal advice)

# Background

- Although truly anonymous data is not subject to GDPR,
  - Anonymisation activity itself is probably not fair processing,
  - Re-identification risk.
    - Linkage to other anonymous data
    - Hospital gives data to Google ...
    - Potential harm from scamming/hassle of re-identified subjects.
- UK Biobank and Dementias Platform UK
  - Data access requires manual vetting of application.
  - Clinically relevant information.
- fastMRI.org
  - MR raw data (ISMRMRD) + DICOMs of knees – no other info.
  - Automated sign-up with Data Sharing Agreement.
    - Agree to no re-identification, cite source, ...
- CRUK NCITA: Simon Doran linking 7 sites

# Way Forward (in my view)

- Decide on data release mechanism.
  - Balance safety against effort.
    - Manual screening on institution / email / publications / reputation difficult
    - Robust automated system requires effort – Captcha etc.
  - Suggest ‘quiz’ based – first reconstruct a synthetic image of a passcode?
- Liaise with Data owner (NHS?) who will want Data Sharing Agreement - can be ‘take it or leave it’.
- Explicit consent required.
- Not many subjects needed for exemplar data.
- Restrict to ‘Engineering’ applications – no subject demographics, maximum anonymization, avoid rare or recognizable data.