EPSRC – TANdoRI Grant Application Data Management Plan
PI – Professor Mark Sterling, Head of School of Engineering
University of Birmingham (March 8th, 2016)

The overriding principle of the guidance provided by the *RCUK Common Principles on Data Policy* is that research data produced using public funds is in the common public good and should be made openly available in a timely manner; access should be available with as few restrictions as possible. With this as a central guiding principle, a data management plan is set out below. Best practices are set out by the UK's Digital Curation Centre; this is a "world-leading centre of expertise in digital information curation with a focus on building capacity, capability and skills for research data management across the UK's higher education research community." Ultimately data will be published in appropriate scientific journals that permit free and open access to the general public; this is compliance with RCUK's requirements all partners involved in the project (including Rensselaer Polytechnic Institute) will agree to comply with common practices of data management, storage and management, as outlined in this document.

Data types and storage
Types of research data to be managed from the (especially) Work Packages 1-4 in the following terms: quantitative, qualitative, 3-D flow field, mapped physical simulations, swirl ratios, velocity measurements, pressures, building/vehicle scenarios and measurements, analytical models, kinematic representations of tornado vortex flows, high speed video images, wind tunnel data and generated from tornado simulations.

File formats, software used, number of records, databases, sweeps, repetitions (in terms that are meaningful in your field of research). Do formats and software enable sharing and long-term validity of data?
In the first instance, the collected/generated data will be stored on password protected and encrypted University-networked computers or on the Research Data Store (RDS); access is restricted to the RDS and research group members will have access. For the purpose of data analysis or when working off-site, some of the data may be transferred to University of Birmingham (UoB) laptops, all of which are password-protected and encrypted. All of the data will carry time and date stamps and where appropriate version control software will be used. After data processing or design modification, the files will be uploaded to the RDS; this will act as a central repository of all such data. The file server is routinely backed up daily. Maintaining a curated central repository and ensuring the security of the information with restricted access will avoid the risks of fragmentation, loss and/or tampering of the data; restrictions will also mitigate against obsolescence (prior to later public release), whilst ensuring that from an early stage, the information is available to other researchers involved in the project. Processed or synthesised data in the form graphs, spreadsheets, interim reports and presentations (in commonly used formats such as .docx, .pdf, .txt, .pptx) will be prepared and stored in similar fashions.

These records will be kept in the laboratories in written logs or electronically on the University's networked computers. At Birmingham, the data will be kept primarily on the RDS where it is housed on-campus within a couple of resilient data centres.

Back-up and security of data

The data will be stored on the computers of the PIs, Co-Is and other researchers as required. Further backup of the data will be provided by saving the data through the RDS file store. The data and its analysis can be deposited on RDS and made accessible only to the project team members through the use of passwords. Backup copies of data are taken on a daily basis and data is stored in separate buildings from the live data.

The data will be held on infrastructure which highly resilient and is hosted in two data centres on University of Birmingham campus. If there is a major issue at one data centre, the service will continue to operate from the other.

The UoB has an Information Security document here :
https://intranet.birmingham.ac.uk/it/documents/public/Information-Security-Policy.pdf

The UoB has guidance on file naming here :
https://intranet.birmingham.ac.uk/it/documents/public/Guidelines-File-Naming-Convention.pdf

The RDS has a backup and retention policy on how it looks after the data including archiving of primary data here :
https://intranet.birmingham.ac.uk/it/teams/infrastructure/research/bear/research-data-service/RDS/BackupRetentionPolicy.aspx

Methods for data sharing

For data sharing between project participants, the UoB has a service similar to Dropbox called BEAR DataShare. This is a file synchronisation and sharing service provided by the UoB; this allows UoB research staff and postgraduate research students to securely share files with others (Non-University) and across different desktop computer systems and mobile devices. This facility will be used for the benefit of all partners within the project. The service is free for UoB members, as it is provided centrally.

For the sharing of data with the general public, experimental data that underpins published work will be deposited in trusted digital repositories. To that end, UoB uses PURE which is the UoB's research information and management system; this gathers research outputs and data in a central database. PURE is designed to store and integrate information on research activity in a structured and standardised way. When research data is uploaded to PURE, the associated metadata is created automatically and PURE can act as a repository for smaller data files (an example being text files that contain numerical information). Members of academic staff have PURE accounts and can update and add to their research profiles; supervisors can perform the same function on behalf of their research students.
When a paper is published as a result of this research, the data that underpins it will be stored securely in the research data archive for at least 10 years (if this is not already done so within an external trusted digital repository such as a funders or journal repository).

Discovery of the data will be enabled when the data location is added to the UoB PURE system; this will be done via the research at Birmingham portal (http://rab.bham.ac.uk/),

which in turn is scoured by Google thus making it public. There is a workflow in place on how the data can be accessed.

Proprietary data

It is envisaged that the parties involved will mainly share and use publicly available information. Whenever confidential information is shared, standard confidentiality agreements will be used. In particular, the parties involved will use reasonable endeavours to ensure that any confidential information disclosed or submitted in writing or any other tangible form to one party ("Receiving Party") by the other ("Disclosing Party") shall be treated with the same standard of care and discretion to avoid disclosure as the Receiving Party uses with its own similar information which it does not wish to disclose. Any information disclosed orally or visually that is identified (orally or in writing) by the Disclosing Party as confidential information shall be treated the same as if it had been reduced to writing at the time of disclosure to the Receiving Party. Each party, through their respective Commercial Technology Transfer Offices, have already in place standard operating procedures to record and protect any Background IPR belonging to the Parties. At the same time, ownership of any Foreground IP generated during the project will be apportioned to the Parties on the basis of their contribution to the arising IP.

In accordance with normal academic practice, all employees, students, agents or appointees of the University (including any others who may work on the project) shall be permitted to discuss resulting IP in internal seminars and to give instructions within the University on questions related to such work. All of the parties involved shall be permitted to publish the results of the project in accordance with normal academic practice and each party shall send the other parties a draft of all intended publications in advance of publication; this so that the other Parties can review them for (especially) the possible inclusion of any of its confidential information. The other parties shall have 28 days after the receipt of the draft to request in writing the delay or amendment of such proposed publication on the grounds that there is subject matter which needs patent and/or similar protection or to prevent publication of any confidential information arising from the other Parties. Where in the reasonable opinion of the other parties a proposed publication contains patentable or commercially sensitive subject matter (which needs protection) then a request to refrain from doing so for a maximum of six months can be made; this, to allow for an application for patent protection in the name and at the cost of the owner of the resulting IP.

Timeframes for data release

The UoB pledges to release all experimental data that underpins published work. In compliance with RCUK's policies, all published work will be open access and UoB will release all underpinning experimental data at the point of publication. This will either be at the time a journal publishes a pre-print of a paper online or at which time UoB publishes work itself; otherwise, there is no requirement for the UoB to set a specific embargo period on the release of such data. The UoB will ensure that the experimental data is in a non-proprietary, open format at the point of release. As UoB's primary focus for the research project involves developing physical and numerical models to understand tornado flow fields, the UoB envisages that codification guidelines will be produced prior to suitable release to the general public. However, UoB acknowledges that some intellectual property may be generated during the course of the research project; in this instance, the data may be

subject to the standard procedures for the recording and protection of IPR; in such cases an embargo period will apply prior to the release of this information to third parties and/or the general public.

Format of the final dataset

The UoB will ensure that the final dataset will be in an open format which will be human and machine readable.

Compiled by Aslam Ghumra (IT Services) and Greg Howard (College of EPS)
University of Birmingham
March 8th, 2016