



## Data Management Plan (DMP)

### Asset Inventory

Answer all questions

- 1. Title of the data asset:** Ting-Chun Chen NLP summer project dataset  
*e.g., the project or study name*

- |                               |  |
|-------------------------------|--|
| <b>2. Data owner:</b>         | Ian Knight   |
| <b>Nominated replacement:</b> | Director of Research   |
| <b>Data steward(s):</b>       | Ting-Chun Chen will collect, process, manage, analyse and report the data. |
| <b>Data users:</b>            | Ian Knight may review the data and reports.                                |
- Note: The details provided in Section 2 may change and need to be updated over the lifetime of the project. Only UoN staff can be data owners, stewards or users; any non-UoN party involved is a third-party and should be described in Section 7.*

- 3. Who or where will the data be obtained from?**  
*Include all persons, organisations or sources (e.g., websites) the data will be obtained from (including parties who obtain data on your behalf); provide evidence (e.g., Terms & Conditions) if data is collected from online sources that supports your proposed use of the data*
- No personal data will be used or obtained. All data is collected from published articles.

- 4. 4.1 Where will the data (digital and / or physical) be stored?**  
*Data should be stored on University servers wherever possible; students should store data on their O365 account if possible (see Section 7 for software repositories, etc.)*

University O365 OneDrive folder dedicated to this project.

- 4.2 How will physical data (e.g., paper forms) be secured?**

N/A

- 4.3. How will you ensure only authorised users have access to the data?**  
*You must notify the ethics committee immediately if a data breach occurs*

MFA is enabled. The folder will not be shared with any other users.

- 4.4 How will access be revoked?**  
*e.g., if a user leaves the University*



OneDrive storage will be deleted automatically after the student completes their degree.

**5. 5.1 Will the asset include data that identifies or could identify people involved in the research, either on its own (e.g., audio or video) or in combination with other data (e.g., location data)?**

☐ Yes ☒ No

If Yes specify the categories of identifiable or potentially identifiable data that will be collected and used in the research:

*e.g., names, postcodes, audio data, video data, IP or MAC addresses, account handles, etc.*

**5.2 Will data that identifies or could identify people be de-identified (anonymised or pseudonymised) before it is used?**

☐ Yes ☐ No ☒ N/A

**5.3 Will data that identifies or could identify people be destroyed after it has been de-identified?**

*e.g., will audio recordings be destroyed after they have been transcribed or textual responses had identifying details removed*

☐ Yes ☐ No ☒ N/A

If No please explain why not:

**Note: It is best practice to de-identify data before use and destroy the original data that identifies or could identify people, if possible and practical to do so**

**6. 6.1 Will the data include special category data?**

*Special category data includes health, genetic and biometric data, and data about people's sexual orientation or sex life, racial or ethnic origin, trade union membership, political opinions, religious or philosophical beliefs.*

☐ Yes ☒ No

If Yes specify which special categories of data will be collected and used in the research:

**6.2 Will large volumes of special category data that identifies or could identify people be collected and used?**

*If you are unsure as to what constitutes a 'large volume' please contact the ethics committee*

**a)** ☐ Yes ☒ No ☐ N/A

If Yes has a data protection impact assessment (DPIA) been conducted?  
*Contact the University's Data Protection Officer for DPIA guidance*



**b)** ☐ Yes ☐ No ☒ N/A

If Yes please provide the DPIA reference and submit the DPIA with the ethics application:

**7. 7.1 Will data collected during the research be shared with third-party recipients (anyone outside the University, e.g., research partners)?**

☐ Yes ☒ No

If Yes please list them:

**7.2 Will data that identifies or could identify people be shared with third-party recipients?**

*i.e., with anyone outside the University*

☐ Yes ☒ No

If Yes please list them:

**7.3 Will data that identifies or could identify people be shared with and / or processed by an online service that is not provided by the University (non-University software)?**

*This includes GitHub and other repositories that store data that may identify people (storage is part of data processing)*

☐ Yes ☒ No

If Yes name and insert a URL to the online service(s) the data will be shared with:

*If using GitHub or other repositories, insert a URL to the project site when this has been created*

**7.4 If data that identifies or could identify people is to be shared with third-party recipients or online services not provided by the University, will the data first be anonymised or pseudonymised?**

☐ Yes ☐ No ☒ N/A

If No please explain why not:

**7.5 If data that identifies or could identify people will be shared with third-party recipients or online services not provided by the University, do they comply with data protection regulation?**

*i.e., are they GDPR compliant*

☐ Yes ☐ No ☒ N/A

If Yes, please describe how they ensure data protection (e.g., copy/paste relevant extracts from their privacy policy, terms and conditions concerning data



usage, or the binding contractual clauses that ensure data protection, include URL's or a University contract number as appropriate):

**8. 8.1 Is it a requirement that the data is open and / or made available to others (e.g., following publication)?**

See SHERPA/JULIET <http://www.sherpa.ac.uk/juliet/index.php>

☐ Yes ☒ No

**8.2 Will data that identifies or could identify people be de-identified before it is made available to others?**

*i.e., will the data be anonymised or pseudonymised*

☐ Yes ☐ No ☒ N/A

If No the data should not be made available to others, unless clearly gathered for public purposes and consent has been obtained

**9. Is the data confidential?**

*e.g., does it involve trade secrets or is will it be protected by a non-disclosure agreement*

☐ Yes ☒ No

If Yes please provide the reference number of the confidentiality agreement signed by the University's Research Operations team:

**10. Declaration**

In signing this form, I confirm that the information provided above is correct.

I confirm that I will not make data that that identifies or could identify people available to others unless it is clearly gathered for public purposes as specified in the ethics forms, and consent has been obtained.

I confirm that I will notify the ethics committee immediately should a data breach occur.

**Signed:** *Ting-Chun Chen*

**Print name:** Ting-Chun Chen

**Date:** 2023-07-01

*The DMP should be submitted to CS REC for audit purposes if ethics approval is required for the research. Submit the DMP along with your ethics application.*