**Long-term impact of mining stereotypes on the emotional wellbeing of dwarves in social, community, corporate and governmental settings**

***P. Snow White, School of Interspecies Relations***

1. **Types of Data, formats and expected volume**

The research's purpose is to establish the long-term impact of mining stereotypes on the emotional well-being of dwarves in social, community, corporate, and governmental settings.

**Types of Data:** The research will generate qualitative data from semi-structured interviews and ethnographic field notes, which will be taken from village hall, commons-community, company, and dwarf-union meetings. Audio recordings of interviews will be recorded using an encrypted offline recording device to protect anonymity, except where certain postholders have agreed to be named, e.g. the Mirror on the Wall. Secondary data relating to key topics in human-dwarf relations, e.g. apples, combs, laced bodices, iron shoes, and glass caskets, shall be retrieved through a review of media and corporate archives (both dwarf and human), publicly available litigation records, queendom statistics regarding anti-dwarf crime and economic and social indicators.

**Formats and Software:** Data will be in password-protected text format (.DWARFX while working, .PRINCEDF for final version) for transcripts and field notes. Audio recordings will be in a common audio format (e.g., .MIRRORP3, .PRINCM). Transcription of audio files, containing both Dwarvish- and Commons-language interviews (though principally Dwarvish) will be conducted by the researcher, without use of any transcription software. Data analysis software (like NVivian) may be utilized for qualitative data analysis.

**Volume:** The exact volume of data is not yet determined, although over seven months I anticipate approximately 100 interviews plus participation in several meetings. Each interview will last 2-3 dwarf-hours, so up to 170 human-hours of audio recording may be generated. These will then be transcribed as written notes and saved as .PDF files, first in Dwarvish and then translated to Commons-language.

**Sensitive Information:** The data will include sensitive and confidential information from participants such as occupation, gender, species, and ethnic details, and will require stringent data protection measures.

**Sharing and Long-term Validity:** Formats and software will be chosen to facilitate interspecies sharing (subject to ethical considerations) and ensure long-term validity of data.

1. **Collection and transfer**

Data will be collected during field research in the Dwarvish Region of the Queendom, through interviews and observations over seven months in 2063. Interviews will be recorded with a digital voice recorder that has 256-bit Advanced Encryption Standard (AES) file encryption and device password protection (such as the Grimhilde DwarfSeeker-7000 Digital Voice Recorder). Strong, standardised password requirements will be met. Field notes will be taken during observations on water-resistant Merpaper or on a secure plasma tablet. Before fieldwork, I will conduct pilot interviews to refine the data collection methods.

Pseudonyms will be given to all participants except where specified above (e.g., the Mirror on the Wall) to ensure anonymity. Pseudonyms will be allocated immediately upon request of an interview, regardless of whether an interview is accepted or not, and prior to meeting individuals who agree to participate.

To ensure safe transfer into storage:

* After audio recordings are collected they will be transferred from the secure recording device to my SunDrive account with the University’s VPN on secure wi-fi, and the audio files will be individually encrypted again with 256-AES encryption software. This will provide a robust and high level of security for encrypted files and will ensure that all data remains confidential and protected from unauthorised access during storage and transfer.
* While the exact encryption software package remains to be selected, it will likely be through well-reviewed packages such as AxeCrypt Premium, HuntsmanLock or OrbLocker.
* Written field-notes will be stored on a daily basis in a container such as a key-access waterproof/fire resistant document chest. They will also be scanned and backed up onto my SunDrive account, accessed via the University’s VPN.
* Once physical documents have been duly scanned, encrypted and backed-up, they will be destroyed.
1. **Storage, Backup and Access**

Audio files will be saved in both WOV and MIRRORP3 format. Text files will be saved in DWARFX and PRINCEF, while scans will be saved as both JPGrunhilda and PRINCEDF files to ensure two copies in different format of each document. Again, all files will be encrypted to AES-256 standards, with strong password-protect.

Backups to SunDrive will occur at minimum once per day, and more frequently as access to internet via VPN allows.

I will also establish a secure internet connection at my place of residence in Dwarfheim, and only connect to University networks through the University’s VPN facility, which I will use to upload encrypted documents to the University OneDrive and centrally managed data storage facility.

With the pseudonymization of participants, two key files will be generated: One spreadsheet with a list of different Pseudonyms in one column, and a second with the list of participants and the corresponding row showing their Pseudonym. These files will be kept in a folder separate from the location of the research data, stored on the University’s centrally-managed server, to ensure that should any individual file location be compromised, the identities of participants would not be ascertainable. Due to the key file's sensitivity, it will be kept in a folder that has been provided additional security measures, after discussions with ITS.

Additionally, access to the key files will be strictly controlled to protect participant confidentiality. Only myself and my supervisors can access these password-protected documents.

1. **Documentation**

Briefly describe the documentation produced about the data generated as part of this research that can enable research data to be used by others outside of your own team. This may include documenting the methods used to generate the data, analytical and procedural information, capturing instrument metadata alongside data, documenting provenance of data and their coding, detailed descriptions for variables, records, etc.

Folder and file structure will be as follows:

* /Log
* This folder will contain an Excel log of all files showing their names, contents and relevant notes about the data
	+ /audio records
		- /interviews
		- This folder will contain the original audio recordings of interviews, with files named according to the following the naming structure
			* + [YYYMMDD]-[PSEUDONYM]-[LOCATION]
		- /meetings
			* + [YYYMMDD]-[MEETING TYPE][[2]](https://ukc-word-edit.officeapps.live.com/we/wordeditorframe.aspx?ui=en-us&rs=en-gb&wopisrc=https%3A%2F%2Funiversityofstandrews907.sharepoint.com%2Fsites%2FRDMGeneral%2F_vti_bin%2Fwopi.ashx%2Ffiles%2F5c47f30cf85f462abcff49b3ab3abf49&wdenableroaming=1&wdfr=1&mscc=1&hid=5877bfae-865b-44c7-a0ad-7c0b059072e6.0&uih=teams&uiembed=1&wdlcid=en-us&jsapi=1&jsapiver=v2&corrid=018aacbf-dce0-4980-ad11-a5961b1d8f80&usid=018aacbf-dce0-4980-ad11-a5961b1d8f80&newsession=1&sftc=1&uihit=TeamsModern&muv=v1&accloop=1&sdr=6&scnd=1&sat=1&rat=1&sams=1&mtf=1&sfp=1&halh=1&hch=1&hmh=1&hwfh=1&hsth=1&sih=1&unh=1&onw=1&dchat=1&sc=%7B%22pmo%22%3A%22https%3A%2F%2Fwww.microsoft365.com%22%2C%22pmshare%22%3Atrue%7D&ctp=LeastProtected&rct=Normal&wdhostclicktime=1710332525311&instantedit=1&wopicomplete=1&wdredirectionreason=Unified_SingleFlush#_ftn2)-[LOCATION]
	+ /transcripts
	+ Folder for written transcripts of audio notes with the following naming structure reflecting that of the original audio files
		- /Dwarvish
			* [YYYMMDD]-[PSEUDONYM]-[ROLE]-[LOCATION]-PT-V*nn*
			* Any changes to the file after starting transcription will be record with an incremental V*nn* number (ie V01, V02). Obsolete versions will be deleted, though any original raw copy will be retained.
		- /Commons-language
			* [YYYMMDD]-[PSEUDONYM]-[ROLE]-[LOCATION]-EN-V*nn*
		- Any changes to the file after starting transcription will be record with an incremental V*nn* number (ie V01, V02). Obsolete versions will be deleted, though any original raw copy will be retained.
	+ /Fieldnotes
	+ Folder for with any observational notes taken with the following naming structure
		- [DDMMYY]-Notes-[LOCATION]
1. **Main risks to data security and how these will be mitigated**

Summarise the main risks to confidentiality and security of information and how these risks will be managed. Cover the main processes and security measures for storage and processing of commercial, confidential or personal data, for data access. Describe controls put in place and auditing of user compliance with consent and security conditions.

 The main confidentiality risks are the small size of the interviewee pool and the possibility of identification despite anonymity. To manage these risks, the published version of the transcript data will be redacted as necessary to avoid identification through details mentioned by the participants. The unredacted transcripts will only be available for academic research purposes to the researcher on this project and the supervisors of this project. All other sensitive data cannot be made publicly available and will be managed and stored as described above.

1. **Long-term data retention, preservation strategy and destruction**

Detail the plans in place for long term storage, preservation and retention of the research data, including for how long. Indicate which data should not be retained (if any) and why.

I will also establish a secure internet connection at my place of residence in Dwarfheim, and only connect to University networks through the University’s VPN facility, which I will use to upload encrypted documents to the University OneDrive and centrally managed data storage facility.

1. **Sharing and Publication strategy (including restrictions and delays)**

Describe if, where and in what form the data will be shared. Consider whether all or part of the data should be shared, and why. Researchers should consider institutional, funder and publisher policies before deciding on their approach to sharing data arising from their study. It is crucial that researchers anticipate their potential future data sharing and/or publication requirements.

As described above, redacted transcripts will be shared in an open access repository, embargoed until the release of the associated thesis. All other data is too sensitive to be shared and will be stored on the University’s centrally managed server signed over to the care of the supervisor on the graduation of the researcher from the University. This storage will be reviewed at periods