

# Phase I Research on Research Integrity Compilation

Application Number: ORI2016000135

Application Name: Stichting Katholieke Universiteit Brabant

State: FN City: Tilburg

## Scoring Criteria

### **Criterion 1: Significance**

#### **Strength:**

Page: 1

Successful development of techniques proposed in the application to identify fabricated data would benefit the larger group.

Page: 26

The relevance of the project is reflected in the disconnect between the small number of cases of misconduct that are detected and the substantial amount of fabrication that might reasonably be anticipated based on the number of scientists who bear witness in surveys questionable practices.

Page: 26, 36

The goal of this proposal is to develop and evaluate automatic statistical tools to detect data fabrication which, as described in the proposal, have never been implemented before.

#### **Weakness:**

Page: 26

The application does not provide clear evidence that the type of misconduct to which scientists admit is aligned with the kind of misconduct that the proposed methods would be able to address.

Page: 6

The application did not explain why it only focuses on social and medical sciences.

### **Criterion 2: Investigator(s)**

#### **Strength:**

Page: 23-24

The PI has published quite extensively on scientific statistical methodology and research misconduct. The Co-Investigator and collaboration with Content Mine have relevant expertise to conduct the proposed research.

**Weakness:**

Page: 23-24

The Principal Investigator does not have experience overseeing a large project given that he is still emerging as a scientist in a Ph.D. program.

Page: 23-24

There is no investigator with qualitative analysis experience.

**Criterion 3: Innovation****Strength:**

Page: 28

The applicant provides an innovative approach to collecting fabricated data for Phase 1, incentivizing participants to create false datasets that are difficult to detect. This would hopefully create an adequate dataset that would power a sensitive detection tool.

**Weakness:**

Page: 36

Many literatures have suggested the possibility of using automatic tools to screen for data anomalies, so the idea proposed is not innovative.

**Criterion 4: Approach****Strength:**

Page: 27

The three project aims being proposed are interconnected and provide the basis for one another.

Page: 27

The investigators propose a solid and comprehensive plan to study statistical tools for the detection of data fabrication, both mathematically and behaviorally.

**Weakness:**

Page: 29-33

There are some weaknesses in Project 1, which include: 1) The application does not consider the situation where researchers do not fabricate the entire data but only modify some numbers in real data, 2) Researchers can improve the way they fabricate data if they know what kinds of statistical methods are used for fabrication detection, 3) Twenty participants might not cover all or most of fabrication behaviors, and 4) There is a possibility that none of the methods can distinguish genuine data from fabricated data. Therefore, it is unclear how the applicant will continue to Project 2 and 3.

Page: 33

The application does not make a convincing case that the limited timeframe available to study participants to construct fabricated data is relevant to contexts where actual fabrication takes place.

Page: 34

This application does not clearly demonstrate an inductive methodology.

Page: 36

The application does not provide a convincing basis for drawing sharp inferences regarding the types of fabrication that present the greatest risk given the modest number of available participants.

Page: 37, 39

In Project 3, the investigators plan to use existing software for publication screening before applying their proposed detection method. This screening step might be risky since this software may fail to identify certain types of data anomalies.

Page: 42

It is not clear in the application why material flagged by ContentMine would necessarily be thought of as presenting fabrication risk, which suggests a substantial proportion of false positives.

## **Criterion 5: Environment**

### **Strength:**

Page: 43

Involvement on project team of people having first-hand familiarity with successful investigation of recent high-profile cases of fabrication at University of Tilburg is a strength.

### **Weakness:**

Page: 21

There are no details on the specific resources available to study staff in completion of the proposed research, such as computer capability, storage, security, and software programs. Similarly, there should be separate information about Content Mine, as they are a central resource for the development of proposed Phase 3 tools.