

# Research Degrees Examination Decision Form (PhD)

Please complete within three weeks of the oral examination and email to the Research Degrees Team ([researchdegrees@kcl.ac.uk](mailto:researchdegrees@kcl.ac.uk)) together with both preliminary reports, a joint final report and list of amendments (if applicable)

**Name of Student:** Sagar Joglekar

**Names of Examiners:** Prof. Joemon Jose and Prof. Yulan He

- The examiners report that they have examined the thesis submitted by the student and have conducted an oral examination:

Date of oral examination: 6 December 2019

Was the supervisor present: YES ☒ NO ☐

- The examiners further report that they have satisfied themselves that the thesis:  
*Please tick relevant boxes*

a) is genuinely the work of the student	<input checked="" type="checkbox"/>	g) includes discussion of those findings and how they advance the study	<input checked="" type="checkbox"/>
b) forms a distinct contribution to knowledge of the subject	<input checked="" type="checkbox"/>	h) demonstrates a deep and synoptic understanding of the field of study, objectivity and the capacity for judgment in complex situations and autonomous work in that field	<input checked="" type="checkbox"/>
c) affords evidence of originality: 1) by the discovery of new facts <i>and/or</i> 2) by the exercise of independent critical power	<input checked="" type="checkbox"/>	i) is satisfactory as regards literary presentation	<input checked="" type="checkbox"/>
d) is an integrated whole and presents a coherent argument	<input checked="" type="checkbox"/>	j) includes a full bibliography and references	<input checked="" type="checkbox"/>
e) gives critical assessment of the relevant literature	<input checked="" type="checkbox"/>	k) demonstrates research skills relevant to the thesis	<input checked="" type="checkbox"/>
f) describes the method of research and its findings	<input checked="" type="checkbox"/>	l) is of a standard to merit publication in whole, in part or in a revised form	<input checked="" type="checkbox"/>

- Following the oral examination, the examiners have agreed the following result:  
*Please tick one box only*

1)	<b>Pass:</b> The student has satisfied the examiners in the examination for the degree of PhD	
2)	<b>Pass, subject to minor amendments:</b> To satisfy the examiners in the examination for the degree of PhD, the student is required to make specified minor amendments within three months	<input checked="" type="checkbox"/>
3)	<b>Pass, subject to major amendments:</b> To satisfy the examiners in the examination for the degree of PhD, the student is required to make specified major amendments within six months	
4)	<b>Not a pass, however:</b> The student is permitted to re-enter the examination for the PhD degree and to re-present the thesis in a revised form within eighteen months (A further oral examination is at the discretion of the examiners)  A further oral examination is <input type="checkbox"/> required <input checked="" type="checkbox"/> not required	

5)	<b>Pass, subject to one month's amendments:</b> Only to be used if the student has not satisfied minor or major amendments or after the re-presentation of the thesis	
----	--	--

6)	<b>Failed PhD, however:</b> The student has satisfied the criteria for the award of a related lower degree (if available)	
7)	<b>Failed PhD, however:</b> The student is permitted to enter the examination for a related lower degree and to re-present the thesis in a revised form within eighteen months (if available)	
8)	<b>Fail:</b> The student has not satisfied the criteria	

#### 4. Preliminary Reports

4.1 Both examiners should complete their independent preliminary reports prior to the oral examination

4.2 **Please note:** The student will be provided with a copy of both the preliminary reports

#### 5. Final Joint Report

5.1 The final joint report should give the grounds on which your decision is based, and should include the student's name; thesis title; the names of each of the examiners; and the date.

5.2 **Please note:** The student will be provided with a copy of this report.

5.3 Please also refer to supplementary Doctoral Studies Prize form.

#### 6. Minor Amendments (within 1 or 3 months) or Major Amendments (within 6 months)

6.1 Minor and Major amendments may be checked & approved by either or both examiners, **PLEASE NOTE:** The supervisor is no longer permitted to check or approve amendments Please indicate the person(s) responsible for checking the amendments:

Prof. Joemon Jose

(Please note: if no one is indicated, this will default to both the examiners for checking)

6.2 **Please note:** The official start date of the amendments will be when the research degrees team emails them to the student following ratification

#### 7 Amendments

7.1 If the student is required to make amendments to the thesis, please ensure that the person responsible for checking them confirms the final decision by email to [researchdegrees@kcl.ac.uk](mailto:researchdegrees@kcl.ac.uk)

#### Examiners' Names:

.....Prof. Joemon Jose..... Date ...6/12 2019.....

.....Prof. Yulan He..... Date .....6/12/2019.....

## Joint Examiners Report

<b>Student Name:</b>	Sagar Joglekar
<b>Thesis title:</b>	From communities to crowds: quantifying the subjective Data, measurements and models
<b>1<sup>st</sup> Examiner Name</b>	Prof. Joemon Jose
<b>2<sup>nd</sup> Examiner Name</b>	Prof. Yulan He
<b>Date:</b>	6 December 2019

### Comments:

Student gave a power point presentation for 15 minutes. Subsequently we continued with questions. Student answered them competently to demonstrate mastery of knowledge.

The thesis makes strong contributions to science with four published works (and another one under review). More importantly the topic is very timely and important.

We have identified avenues for improvement in the thesis and a separate note will follow (to the supervisor) detailing list of corrections.

## List of Amendments

<b>Student Name:</b>	Sagar Joglekar
<b>1<sup>st</sup> Examiner Name:</b>	Prof. Joemon Jose
<b>2<sup>nd</sup> Examiner Name:</b>	Prof. Yulan He
<b>Date:</b>	6 December 2019

### Amendments:

Streamline research questions in the abstract and the introductions; there are different formulations of main RQ and sub RQs ; use them consistently across the thesis

In general RQs should be motivated first through literature analysis or data analysis, before stating it

Each chapter to follow a standard format; motivate RQs; describe the methodology to answer the RQs; conduct analysis and present data; interpret the results; discuss the results highlighting whether the RQs are answered or not

Figures, tables – always use captions consistently; interpret the results; discuss implications where possible

Introduce a literature survey/background chapter; relate and place the work in context; find relationships between concepts; discuss collective intelligence; discuss network measures; discuss word embeddings; doc2vc embeddings; deep learning models; discuss social/psychological measures etc.

Use proper labelling/numbering for equations throughout the thesis

### Chapter 1

Sharpen DIKW pyramid and ontological description; as discussed in the viva  
Clarify why use are using meso instead micro measures/metrics

### Chapter 2

Discuss “Social support” along with its dimensions; later at the end of the chapter discuss the results with respect to these

Page 20 – methods and frameworks -> define and use consistently across the thesis

RQs need to be motivated considerably and the formulations need to be consistent across the thesis

Fix the Table 2.1 as discussed in the viva; if possible provide some sample data immediately after to highlight the structure and format

Rich user and super users – differentiate and consistently use; same applies to rich-club behaviour

Figure 2.1 degree of node; just introduce a literature review chapter and explain the background material

Largest connected component? describe in the literature review

Explain a Bernoulli process model is appropriate or not for the case under consideration

Language structure/content analysis is not used; explain and state the implications

Figure 2.6 use captions consistently; always interpret figures

Explain whether temporal analysis is possible or not (as discussed in the viva)

Provide Ties and closures definition

Fix the remarks and descriptions

# List of Amendments

Say how you processed data; give the structure of the data;

## Chapter 3

3.2 justify Front page posts are in fact baseline  
 Provide tabular representation of the data  
 Update the figure 3.1 properly  
 Posts structure and length; explain with sample data and if possible, use examples  
 Network characteristics difference or lack of it  
 Equations need to be labelled and consistently used  
 Fix the missing paragraphs in page 52  
 Section 3.6 elaborate considerably; explain triadic motifs; justify ;  
 Explain the methodology for analysis  
 Describe the algorithm on using triadic motifs  
 Describe/Explain the naming conventions using material from HL17  
 describe the figures consistently; explain the significance/impact

## Chapter 4

Fix the use of hypothesis/research questions in first para page 60  
 Describe and label data collection process/naming strategies  
 page 64 justify the Bayesian algorithm  
 Page 66 justify assumption error analysis – give failure figures  
 Provide details of deep learning models; state the hyper parameters  
 Fix the Table 4.1 data

## Chapter 5

Reformulate RQ7  
 Provide details of GAN so that it can be reproduced  
 Describe/summarise Dosovitsy paper  
 Figure 5.2 model should be explained and justified  
 Fix the sudden use of Facelift  
 Explain the methodology to answer research questions  
 Provide enough background details of – for example Inverted-U  
 Table 5.4 creation methodology  
 Provide enough background details of PLACESNET and SEGNET ; use terms consistently  
 Explain how the 120 or so labels in appendix B created  
 Explain surprise findings, for example inverted-u or complexity related  
 Justify and explain logistic regression methodology

## Conclusion

55. Restate/refer RQs and highlight findings

## Preliminary Examiners Report

<b>Student Name:</b>	Sagar Joglekar
<b>Thesis title:</b>	From communities to crowds: quantifying the subjective Data, measurements and models
<b>Examiner Name</b>	Prof. Joemon Jose
<b>Date:</b>	30 November 2019

### Comments:

This thesis investigates the research question, “Can we quantify entities of subjective nature, if the data is large enough, and originates from human communities or crowds?”. I read the thesis with great enthusiasm given the timeliness and importance of the subject. As per contributions to knowledge the thesis reports a number of published works. As per novelty and originality, I haven’t seen any other publication addressing these issues like this. With this generic assessment, I have no reservation in recommending the award of PhD degree subject to satisfactory performance in the viva voce.

Having said this, there exists an element of disappointments on how the document is developed and articulated. There are quite lot of typos, incomplete or hanging sentences in the current thesis. This is much more than I have seen in any thesis under examination and it is clear that the document is developed in a hurry. These can easily be corrected with a careful proofread. It is also the case that the treatment of each chapters is more like reporting case studies. Author failed to put the research in larger context by debating and describing various literature from allied fields. This can also be fixed with a careful drafting of literature and in some cases with some restructuring of the material. Student should make use of consistent referencing styles.

Chapter 1 introduces or reintroduces the research question in slightly different tone and structure. A consistent statement would help rather than related but different formulations. However, author took effort to place the research in rather a larger framework for extracting knowledge and wisdom from data. This “framework” is used consistently in all chapters. This chapter also nicely introduces two parts of the thesis.

Chapter 2 investigates online support to answer two research questions RQ1 and RQ2. I have an issue how the research questions are formed. I would like to have motivated these questions properly and then place the research question. Data collected or crawled and used for analysis. The methodology used is from network science. It would have helped if the author provided literature on these and then placed the used metrics in the context. The larger question still exists are they the appropriate metrics to use or did some other metrics could have been more appropriate. Though a good effort is made to relate to follow-up research, the implications of the findings are not discussed. There is a lack of rigour in the analysis and presentation of the results.

Chapter 3 investigates the signatures of social support and to answer RQs 3 and 4. Data collection process is described. Similar comments I made earlier applied here too. It would be important to justify the metrics used placing them in the larger research context. A proper literature review section will help with this. Similarly, the results need to be discussed rigorously. However, a question exists why the metrics in chapter 2 is not applied in here and vice versa.

Chapter 4 and 5 deals with part 2 of the thesis to answer research questions RQ 5 to 7. Chapter 4 describes data collection and augmentation strategies. Chapter 5 is the exploration of research questions. RQ7 formulation is a bit lousy. Not enough information is given to reproduce the work, example deep learning models. It is commendable that author explained the models in layman’s terms. However, a formal description is needed so that any other person will be able to reproduce the work. All these models require fine tuning of hyper parameters and it is important to treat them carefully. Similarly, I felt a lack of justification of the methodologies used for other investigations in the chapter. As said before results need to be discussed rigorously.

## Preliminary Examiners Report

Chapter 6 gives an interesting description of open problems and planned future work. However, what is lacking is a conclusion section where the research highlights are discussed.

## Preliminary Examiners Report

<b>Student Name:</b>	Sagar Joglekar
<b>Thesis title:</b>	From communities to crowds: quantifying the subjective Data, measurements and models
<b>Examiner Name</b>	Prof. Yulan He
<b>Date:</b>	30 November 2019

### Comments:

The main contributions of this thesis are the proposal of a number of approaches for quantifying subjective signals from communities and crowds. In particular, a framework was developed to abstract out the graphical structure of the interactions of the networked communities where people seek for peer support. In addition, a deep learning based approach for quantifying the perception of “beauty” through the crowd sourced rating of google street view images was proposed. Finally, a generative model of to stimulate the beautification of urban spaces was implemented.

The thesis has given a good introduction of the background of the research. The proposed approaches have been well justified. Experimental evaluations are in general rigorous and the results are convincing.

The candidate has achieved a decent set of publications at different events, most noteworthy of which includes Journal of Medical Internet Research and Social Informatics 2017 conference as the first author, IEEE Computer Graphics and Applications as the second author. He also has two papers, one about the online discussions of mental health in Reddit and another about a deep learning framework to beautify urban scenes, under review.

I would like to clarify the following issues during the oral examination:

#### Chapter 2:

- Page 32, “We consider an expected probability of answering a question by a user as  $\frac{2}{3}$  and the probability of posting a question as  $\frac{1}{3}$ ”, any supporting evidence of this?

#### Chapter 3:

- Page 42, last sentence, repeated “measure”
- Page 47, cosine similarity of two post vectors is used to quantify the semantic alignment of two posts. This metric may not be appropriate if two posts are in different lengths.
- It is perhaps worth to think how content (reply graph) and users’ social networks (user graph) can be combined in this analysis.

#### Chapter 4:

- Page 63, “1.2 pairwise comparison”, missing the unit of measurement?
- Is this a ranking or classification problem?

#### Chapter 5:

- Page 88, explain the “divide by 4 rule”