

# Computing Modular Postgraduate Programmes Dissertation Proposal Form

**Student Number :** 0330840

**Student Name :** Sohrab Khan

**Start Date :** 09/10/2007

**Project Title :** BitDis.com BitTorrent Search Engine

**Supervisor:** \_\_\_\_\_

**This Proposal MUST be agree with your supervisor. Once it is agree please ensure it is SIGNED before submitting to SCOT desk**

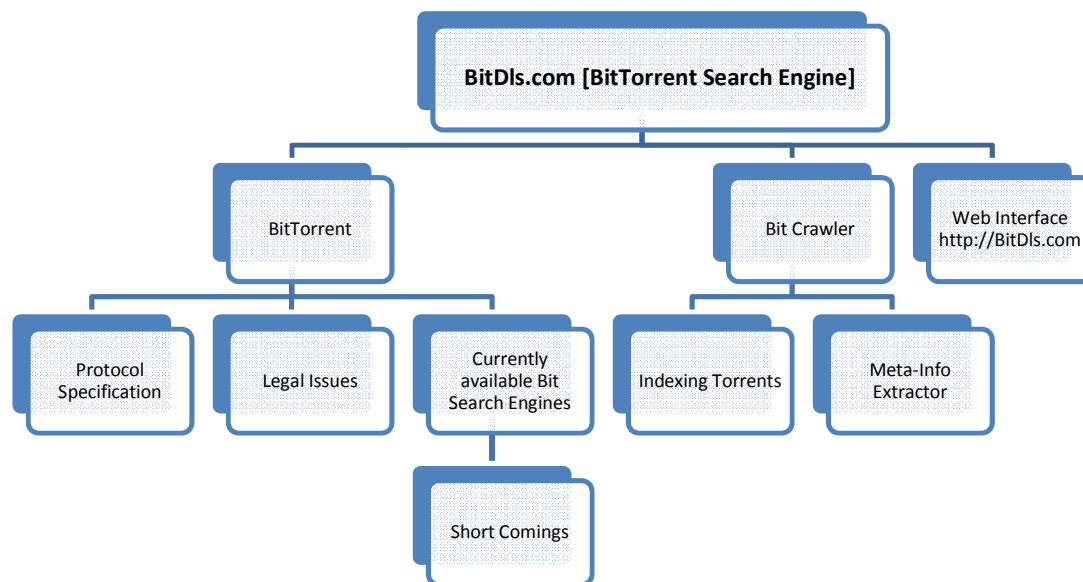
**Signatures:**

**Student:** \_\_\_\_\_ **DATE:** \_\_\_\_\_

**Supervisor:** \_\_\_\_\_ **DATE:** \_\_\_\_\_

1) **Your Main Research (sub)Questions (each can become a Chapter)**

- What is a BitTorrent?
  - How does it work?
  - How widely is it used?
  - Legal Issues.
- Search Engine
  - Crawler
  - Front-end (web interface)
- Shortcomings of currently available search engines
- Solution / Developing the solution
- Evaluating the solution



Dissertation Box

2) **What Research Work has been done previously in this area?**

The BitTorrent protocol is brainchild of Bram Cohen designed in April'2001. Since then a lot of work has been done on the BitTorrent technology. Hundreds of search engines are available on the internet e.g. ISOHunt.com, torrentz.com, torrentleech.org etc.

ISOHunt is considered the most widely used because of its usability and detailed meta-info display.

**3) How will you measure the success of your Dissertation?**

The success of my dissertation can be measured on the following:

- Crawler's white box testing,  
To check if its efficient enough.
- Web Interface black box testing,  
To check if the search does happen plus usability & detailed meta-info availability.

**4) Methodology – what will be your approach to completing this dissertation?**

The dissertation will be done in five phases:

- Data Collection: in which interviews of 5 expert and 5 novice users will be conducted to find out what an ideal BitTorrent Search Engine will look like.
- Torrent Crawler & Meta-info Extractor will be developed.
- Web Interface to the Search Engine will be created. This interface takes input from user and searches the indexed torrent's meta-info.
- Testing & Documentation will be done.
- Finalization of dissertation will be done.

These five phases are in accordance to the Water Fall model of Software Development Life Cycle.

**5) What will be the Practical element?**

Instead of developing a "Google-like" crawler which can take a long time to develop, a basic (.torrent file) crawler will be developed in Java. Plus a front-end website will be developed.

The development of crawler and website makes it 75% practical.

**6) What Data/Examples will you use?**

There are currently hundreds of Bitorrent Search Engines on the web but most of them have incorrect information about a torrent (e.g. The health of torrent). So 5 advanced users and 5 novice users will be interviewed to collect data of what they would expect in a better search engine. After the development they will be given questionnaires to find out if the developed search engine is up to their expectation.

Plus bittorrent.org will be consulted for the protocol specification.

**7) What knowledge and skills do you have that will be most relevant here?**

I have been developing applications & websites since 5 years and have earned Sun Microsystems certifications like Sun Certified Java Programmer (SCJP) & Sun Certified Java Developer (SCJD).

Since the crawler will be Java based so that makes me competent enough to do the job.

8) **Indicative References**

- “BitTorrent Protocol 1.0”, Cohen, B., <http://bittorrent.org> [Oct’2002]  
Explains what BitTorrent is, the requirements & the process etc.
- “Incentives Build Robustness in BitTorrent”, Cohen, B., <http://bittorrent.org> [May’2003]
- “Robustness of the BitTorrent Protocol”, Tamilmani, R., Stony Brook University  
Explains why BitTorrent is the best technique for P2P file sharing.
- “Programming Spiders, Bots, and Aggregators in Java”, Heaton, J., Sybex [Feb’2002]  
Explains how crawlers can be developed in Java.

9) **Time Plan with Main Milestones and estimated hours**

