**Information Retrieval Project phase 1**

**TEAM –YAGOO**

**Dinesh Ravi**

**Magizharasu Thirunavukkarasu**

**Manoj Mylapore Chandrasekaran**

**SYSTEM DESIGN DIAGRAM**

**DIRECTORY STRUCTURE**

**STATISTICS:**

**File count : 149118**

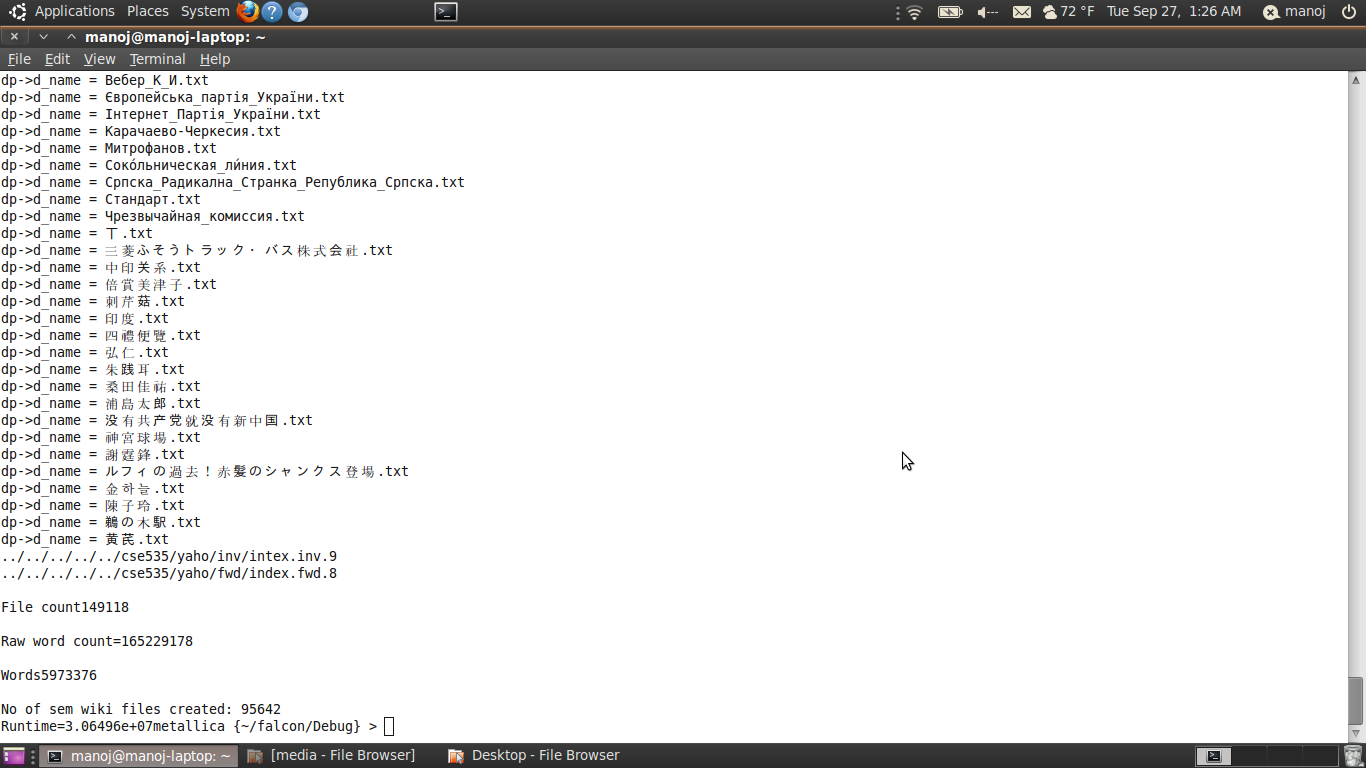
**Raw word count : 165,229,178**

**Word count : 5,973,376**

**Sem wiki files created : 95642**

**Time : 3.06496e^07(units)**

**RUN SNAP SHOT :**

****

**CONTRIBUTION OF EACH MEMBER:**

**Dinesh Ravi :** News files processing , Stemming , term, term count dictionaries

**Magizharasu Thirunavukkarasu :** Wiki parser , Sem wiki creation, Author Dictionary, Category index, Link pair repo.

**Manoj Mylapore Chandrasekaran:** Stopword, Raw text parsing , Index barrels **.**

**Signature :**

**Dinesh Ravi**

**Magizharasu**

**Manoj**