**Datasheet**

**Purpose**

**-          The problem/context**

**-          What is needed for the company**

**How does it work?**

**-        Technical work (features)**

**-          Maths theory**

**Results and perspective**

**-          Results**

**-          Perspective**

**The product is a piece of software processing natural language analysis.**

**During an election campaign, some candidate fans, or employees from poll institutes go door to door to ask about people's opinion. They gather information they have to handle eventually which represents hundreds of answers. These answers are split in two categories: close-ended questions and open questions. Close ended questions can be easily automatically processed, but open questions are much harder to handle.**

**The Quorum Company is working on a complete solution for people ; such as mayors ; who organise campaign to know about people needs or opinions. The solution should help their users to analyse the questions’ answers automatically.   This is why my project is about natural language processing. The idea is to download raw text data, and to automatically process each question's answers to return words that matter and topics people mentioned. Even though the process won't be as good as a human one, but it will cut the number of people needed to do this tedious work.**

**The project is split in two parts : a human interface and a computational part. The computational part is composed of different types of analysis. A frequency analysis which returns the most used words for each questions. An ontologistic analysis find which topic is mostly probable to be mentioned, according to an onthology I have created. A LDA analysis**