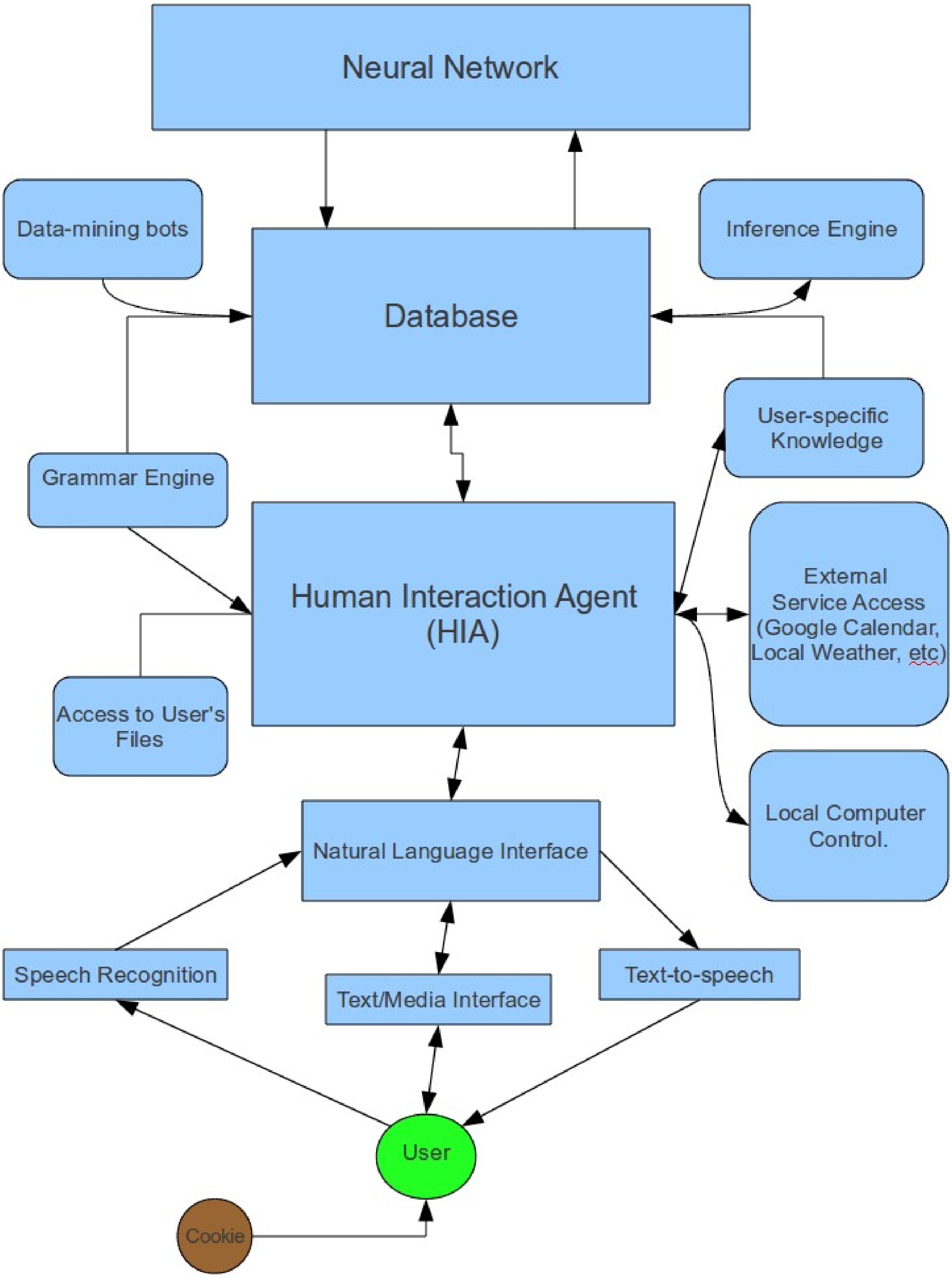
Wintermute


Neural Network
Data-mining bots
Database
Inference Engine
User-specific Knowledge
Grammar Engine
Human Interaction Agent (HIA)
External Service Access (Google Calendar, Local Weather, etc)
Access to User's Files
Local Computer Control.
Natural Language Interface
Speech Recognition
Text/Media Interface
Text-to-speech
User
Cookie

Neural Network: The original design called for building off of VIRGIL: Wintermute's ancestor. However VIRGIL's network may not be able to cope with the amount of data, so VIRGIL will need to be examined and lessons learnt from it. One major problem is that of 'over-training'; feeding it so much data it becomes useless to operate.

With Wintermute, the neural network should (with the assistance of the other components) 'read' data, and be able to generate replies.

Whilst the structure of the neural net is a moot point, I am almost certain that it won't be included in the user-ready releases due to computational requirements (which, currently, point at needing a small server. In the future we may be looking at more advanced computing techniques, such as a Beowulf Cluster system)

Database: The database will be structured in a way not unlike an Ontology (machine-readable dictionary and encyclopedia, basically) It will also store user-specific knowledge (like, for instance, who they are) The design may also allow us to store replies to queries, which the HIA (discussed later) will take and give to the user.

> **Data-mining bots**: Whilst we can gather a significant amount of data from abandoned projects, article dumps of Wikipedia and whatnot, the data will still need to processed. In other situations other websites may need to be scanned and everything processed. This may be one of the more difficult engineering modules.
>
> **Inference Engine:** Whilst the neural network may be able to classify some forms of data and their relationships, the Inference Engine may have to deal with most of the workload. If made particularly advanced then this Engine may be able to 'teach' Wintermute autonomously. This Engine's job is basically link one semantic object with another semantic object.
>
> **Grammar Engine**: This Engine should reduce (if not completely eliminate) multi-lingustic problems Wintermute would face, for instance; in English, we say “A blue coat”. In French, however, they say “A coat blue.” There are other smaller discrepancies; in UK English, we spell 'Colour' in American English, it's spelt 'Color'. The database would store information in such a way that the grammar engine could apply the language’s alterations as and when needed, instead of either having a huge program with an even larger translations team behind it, or having multiple systems for each language, or (at worst) requiring an external translation service.
>
> This design is also extremely unique in that; alongside the database, the 40+ year problem of machine translation of loosing the meaning between two different languages would be solved. A major engineering and scientific achievement in and of itself!

Human Interaction Agent: Arguably the main part of the entire system, the HIA serves as the interface between user and computer. The HIA's conversational ability will be no better then pattern matching between the user and the other components replies and queries to the user. Other, smaller faculties in it may also play a role, such as temporal awareness (IE: aware that it asked the user a question three minutes ago, and has not gotten a response to it, thus leading it to query if the user is there)

Local Computer Control: The HIA will serve as a sort of 'natural language terminal': Let's say the user plugs in a USB memory stick, and says “Please format my memory stick to FAT32” the HIA will accomplish this. Or another case; “Send the funny-kitty video to my phone, please”. The HIA would know from the past that the user has a HTC Desire, and that the user prefers to send files over bluetooth, which the phone is connected to right now. The HIA would then *autonomously* decide to convert the .flv video into .mp4 and transfer it over bluetooth.

Basically put, the user has all the power of the command line using their language (instead of right now, where the user has to basically learn another language entirely to better control their computer)

External Service Access:

For many, computers are useless without an internet connection, so the HIA should be able to do certain activities like:

- Query external services for an answer to a question it does not store
- Provide maps, travel routes (and perhaps even going so far as to help book vacations)
- Provide information about local services (fast food, taxi services, etc)
- Gather information from services like Gmail, Facebook, etc and build up a map of the user's friends, as well as tracking user data, like calandering and to-do services.
- Gather relevent information from the Internet about a given topic.
- Pick up on events and news that may be of interest to the user.

Access to User's Files: The HIA will have access to user data, not entirely unlike KDE's NEPOMUK program; with the advent of the Zeitgeist program, these files could not only be tracked as to when they were created and editied, but whom/what they were sent to, when they were accessed and, with a few patches on our part, able to record what was sent to whom through what!

Natural Language Interface: Simply a module of the HIA; the NLI is the user-facing input/output of the system. It would contain (missing from graph) tools to sanitize the user's input, like a spellchecker and grammar checker.

Speech Recognition: The Speech Recognition system (Julius) would accept input into the system and allow disabled users (or user's who'd prefer a more star-trek way of entering data into their system) to operate Wintermute.

Text/Media Interface: Simple text entry in/output system. Could also display documents, pictures and videos for the user. A more advanced approach would be looking at tying in speech/gesture recognition, to get a *Kinect/Minority Report* way of watching films and dealing with data.

Text-To-Speech: openMary can provide a emotionally capable, high quality voice for Wintermute's use. Emotional algorithms (which can identify emotion) can inform the voice about the proper words to place emotional emphasis on.