Email2wiki:
A Wiki-based Email Archival System
Bryan Reinholz
Advisor: Lyle Ungar

Abstract
In the past 20 years, email has risen to become one of the primary methods of communication between individual members of project teams. However email is still an imperfect communication medium within a team environment. While it is largely a private communication between two entities, the topics addressed are often pertinent to the entire team.

The goal of this design project was to refine the process of transferring important information from a private email conversation to an editable wiki visible to the entire project team. As a proxy for a project management suite, this project consists of an adaptable system that uses restaurant reviews as a demonstration.

Technical Approach
One advantage of this senior design project is that it lends itself to easy modular separation. Each module was individually coded and tested in true object oriented fashion (See Module Overview). The language of choice in this project is Python, which has been used to code the majority of the modules, and Perl, needed for Mediawiki operations.

Of all of the modules listed, the first two required the greatest amount of development. The first module takes an email sent through Mailman and parses out important information regarding the email’s state: its sender, recipients, timestamp, subject and body. The second identifies entities referenced in the email body and categorizes the text based on its best interpretation of the entities present. Overall each of the modules have the potential to be upgraded and adapted to different uses.

Module Overview
- Email Sent
- Parse Email
- Entity Resolution and Tagging
- Wiki Search
- Message Segmentation
- Email Summarization
- Wiki Posting

Sample Email 1

Sample Email 2

Sample Wiki Page

Future Applications
One of immediate possible application is the use of the wiki to schedule events, whether it’s for a specific project, or more generally for a research group. The wiki structure allows meeting times and places to be represented in a much cleaner way than constant emails that may pertain to only half of the recipients. Another interesting use of the project could be in a classroom setting.

The most advanced possible use for this project is a project management utility. Although this would require a significant amount of work, the end result would be extremely useful as a tool to organize groups. It would be less cluttered and disorganized than email, while more versatile than most established project management packages.

Conclusion
The final system is able to take an email, parse it, and post it to a wiki page. Although some features were omitted due to unforeseen difficulties, the final result successfully completes my objective at the project’s outset. The email2wiki project is a solid foundation on which others can further build, extending its functionality beyond the few classifications I have concentrated on (mainly restaurants) and incorporate a wider variety of uses for the system.

Senior Poster Day 2007: CIS Department, University of Pennsylvania