

My love of research started even before coming to college. In Science Fair, I got as far as Finalist in the Intel International Science and Engineering Fair with projects ranging from human perception of wind chill and haptic feedback in robotic control systems to logic gate design and the evaluation of mathematical models of language evolution. In addition, I got the opportunity to study at the Villanova Summer Research Institute where I collaborated with others to research and present on advanced topics. These experiences whetted my appetite for exploration and discovery, and I was eager to attain more opportunities by studying at MIT.

Soon after coming to MIT, I got a job at Metacarta, Inc. While I had initially been expecting simply a coding job, I had been hired by the Office of the Chief Technology Officer, charged with doing research into new possibilities for geographic search and data organization, developing prototypes, and determining what options were practical to integrate with Metacarta's offerings. I started working in groups with other OCTO engineers, first investigating different strategies for adding web crawling to Metacarta's databases, later helping to found what became the OpenLayers open-source web mapping project¹, a ground-breaking effort to combine map-data and imagery from a wide variety of sources using a shared, public interface. OpenLayers taught me a lot about about system and interface design, coordination between parties separated both by physical distance and by differing goals, and the importance of incremental stages of design, prototyping, and refinement when building complex systems. I later put these new skills to good use in developing a semi-automated system for extracting map data from scanned raster images, using image processing algorithms to locate labels and other information which could then be efficiently decoded by human operators. This forced me to develop more advanced research skills to understand the proper algorithms to use and also new considerations of interface design for efficiency and throughput without overly sacrificing usability.

In addition to my work at Metacarta, I received an opportunity to do research abroad through the MIT International Science and Technology Initiatives with the Software R&D Group of Ricoh in Tōkyō, Japan. This presented me with a different research environment and a new set of challenges, from overcoming linguistic and cultural barriers and assumptions while collaborating to dealing with a more rigid organizational structure and formal approval process. My main work at Ricoh was developing an automated system for converting legacy parts catalogs, used for the maintenance and repair of Ricoh products, from a static 2D format to a searchable, interactive, and dynamic format that gives technicians more useful information without the huge manual labor it would take to manually re-create all legacy Ricoh parts catalogs in a modern, 3D, searchable format. While I did collaborate with and learn from my Japanese co-workers, the main part of this project was self-directed, and required me to build upon my image processing experience and do additional research and experimentation with various segmentation, data extraction, and OCR techniques from the literature. Making good use of a variety of existing algorithms and tools, I was able to create a system that could automatically recreate the associations between parts in a table and curves in the device diagram and store this information in a searchable, interactive format. While this experience taught me a lot about image processing and data extraction, even more im-

¹<http://openlayers.org/>

portant was what I learned about making use of relevant prior work to avoid recreating the wheel and better make my work useful, both immediately for its practical purposes and long-term for the future development it enables.

As I continued to work with Metacarta, I began to assume more responsibility, working independently to design an interface for the efficient manual tagging of documents using Metacarta's existing geographic parsing models to streamline tagging and provide more accurate training data to improve those same models. In this project, not only was I self-directed, but I also needed to develop the requirements for the project by consulting with engineers that had done similar work, the researchers developing the models in question and using the data produced, and the taggers who would be directly using the interface in their day-to-day work. In order for this project to eventually result in a useful system, I had to undergo several cycles of prototyping, testing, and redesign, learning from each iteration and combining the sometimes-contradictory feedback from various parties within the company to determine the course of further study. In the end, I managed to create a useful tool actively used in Metacarta to improve its geoparsing models, and in doing so, I learned a lot about the nature of real research, where the problem and requirements are not well-defined and there is no one who knows the right answer to help along the way.