Searching for Jim Gray: A Technical Overview

Joseph M. Hellerstein
David L. Tennenhouse

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2010-142
http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-142.html

December 1, 2010

Searching for Jim Gray: A Technical Overview

Joseph M. Hellerstein and David L. Tennenhouse
writing on behalf of a large team of volunteers

ABSTRACT

When Jim Gray disappeared at sea early in 2007, his friends and colleagues quickly began discussing ways to mobilize their skills and resources to help rescue him. That discussion evolved into an unprecedented civilian search-and-rescue exercise, which—along with the US Coast Guard mission it augmented—was eventually unsuccessful in locating Gray or his vessel.

In May of 2008, search participants and members of the US Coast Guard met face-to-face for the first time to discuss their experience and try to extract constructive lessons. This paper is an effort to distill some of that discussion for the computer science community. We describe the emergent structure of the team and its communication, the "polytechture" of the systems built during the search, and some social and technical challenges that arose from the experience.

1. INTRODUCTION

On Sunday January 28, 2007, noted computer scientist Jim Gray disappeared at sea in his sloop, *Tenacious*. He was sailing single-handed, with plans to scatter his mother's remains near the Farallon Islands some 27 miles outside San Francisco's Golden Gate. As news of his disappearance spread through his social network, Gray's friends and colleagues began discussing ways to mobilize their skills and resources to help authorities locate *Tenacious* and rescue Gray. That discussion evolved over days and weeks into an unprecedented civilian search-and-rescue exercise involving satellites, private planes, automated image analysis, ocean current simulations, and crowdsourced human computing, working in collaboration with the US Coast Guard. The team that emerged included computer scientists, engineers, graduate students, oceanographers, astronomers, business leaders, venture capitalists, and entrepreneurs—many of whom had never met each other before. There was ample access to funds, technology, organizational skills and know-how, and a willingness to work around the clock.

Even with these advantages, the odds of finding *Tenacious* were never good. On February 16, 2007, in consultation with the Coast Guard and Jim Gray's family, the team agreed to call off the search. *Tenacious* remains lost to this day, despite a subsequent extensive underwater search of the San Francisco coastline [9].

Jim Gray was famous for many things, including a determination to transform practical experience and know-how into scientific challenges. As the search for *Tenacious* wound down, a number of us felt that even though the effort was not successful on its own terms, it offered a Jim-Gray-like opportunity for converting the particulars of the experience into higher-level technical observations of more general interest. One of our goals was to encourage efforts to "democratize" the ability for families and friends to use technology to assist in search-and-rescue (SAR)[1], so that people whose social network is not as well-connected as Jim Gray's could undertake analogous efforts. In addition, we hoped to review the techniques we used and ask how to to improve them further to make the next search effort more effective. To that end, in May of 2008 we convened a meeting of search participants and members of the US Coast Guard, the day after a public tribute to Jim Gray at UC Berkeley. This was the first opportunity for the virtual organization that searched for *Tenacious* to meet face-to-face and compare stories and perspectives.

One sober conclusion the group quickly reached is that its specific lessons on maritime search and rescue could only have modest impact, as we detail in Section 2. With that established, we still felt that it would be constructive to cull lessons learned and identify technical challenges. First, maritime search is not a solved problem, and even though the number of lives to be saved is modest, each life saved is precious. Second, history shows that technologies developed in one application setting often have bigger impact in other settings. We were hopeful that lessons we learned searching for Gray could inform efforts during larger life-threatening scenarios including civilian-driven efforts toward Disaster Response and SAR during natural disasters and military conflict. And as part of the meeting, we brainstormed a bit about the challenges of safety and prevention as well.

This paper is an effort to distill some of that discussion for the computer science community, which has become increasingly interested in disaster response (e.g., efforts after the 2007 Kenyan election crisis [2] and the 2010 Haiti earthquake [4]). We document the emergent structure of the team and its communication, the "polytechture" of the systems built during the search, and some of the challenges that arose.

2. BACKGROUND

The amateur effort to find *Tenacious* and its skipper started with optimism, but little context on the task at hand. We did not have any awareness of the practice and technology of SAR, and only a vague sense of the special resources that Jim Gray's friends could bring to bear on the problem.

In this section, with the benefit of hindsight, we provide a backdrop for our discussion of computer science challenges in SAR. We reflect first on the unique character of the search for *Tenacious*, and then describe the basics of maritime SAR as it is practiced today.

2.1 The Tenacious SAR Effort

[1]The acronym SAR also refers to Synthetic Aperture Radar, a remote imaging technology that was employed in searching for *Tenacious*. In this paper, when we use the acronym SAR, we refer exclusively to search-and-rescue.

The search for *Tenacious* was in some ways a very unique effort, and in others a typical case study in volunteer SAR. The uniqueness had its roots in Jim Gray's persona. In addition to being a singular scientist and engineer, Gray was distinctly social, cultivating friendships and collaborations across industries and sciences. The social network he built over decades brought enormous advantages to many aspects of the search, in ways that would be very hard to replicate. First, the team that assembled to find *Tenacious* included leaders in areas like computing, astronomy, oceanography and business management. Second, due to Gray's many contacts in the business and scientific world, the funds and resources available to find were essentially unlimited, including planes, pilots, satellite imagery, and control of well-provisioned computing resources. Finally, the story of a famous-scientist-gone-missing attracted significant media interest, providing public awareness that attracted help with manual image analysis, and tips on debris and wreckage.

On the other hand, a number of the general features that this team wrestled with seem relatively universal to volunteer SAR efforts. First, the search got off to a slow start, as volunteers emerged and got organized to take concrete action. By the time all the expertise and efforts were in place, the odds of finding a survivor or even a boat were significantly diminished. Second, almost nobody involved in the volunteer search for Jim Gray had any experience with SAR. Finally, at every stage of the volunteer search, the supposition was that it would last only a day or two more. As a result, there were disincentives to invest time in improving existing practices and tools, and positive incentives for decentralized and lightweight development of tools and practices that had to be custom-crafted.

If there are lessons to be learned from this effort, they revolve around questions of both the uniqueness of the case, and its universal properties. The first category motivates efforts to democratize techniques used to search for *Tenacious*, some of which need not be as complex or expensive as they were in this instance. The second category motivates efforts to address common technological problems that arise in any volunteer emergency response setting.

2.2 Maritime SAR

Given our experience, the focus of the discussion in this paper is on maritime SAR. As it happens, maritime SAR in the United States is better understood and more professionally conducted than land-based SAR. Maritime SAR is the responsibility of a single federal agency: the US Coast Guard (USCG), a branch of the US Dept. of Homeland Security. By contrast, land-based SAR is managed in an ad hoc manner by local law-enforcement authorities. Our experience with USCG was altogether positive: not only were they eminently good at their jobs, they were technically sophisticated and encouraging of our (often naive) ideas, providing advice and coordination despite their own limited time and resources. In the United States at least, maritime settings are a good incubator for SAR technology development, and the USCG is a promising research partner. As of now it is quite modestly funded, so synergies and advocacy from well-funded computer science projects would likely be welcome.

In hindsight, the clearest lessons for the volunteer search team were (1) the ocean is enormous, and (2) the USCG has a sophisticated and effective maritime SAR program. The meeting in Berkeley opened with a briefing from Arthur Allen, an Oceanographer at the USCG Headquarters Office of Search and Rescue. That office oversees all US Coast Guard searches, with an area of responsibility covering most of the Pacific, half of the Atlantic, and half of the Arctic Oceans. In the remainder of this section, we review some of the main points Allen raised at the Berkeley meeting.

2.2.1 The Phases of SAR

Search and Rescue technology is only needed when people get into trouble. From a public policy perspective, it is cheaper and more effective to invest in preventing people from getting into trouble, than investing in techniques to save them when they get into trouble. Good discussions of boating safety can be found at `http://www.uscgboating.org`, and we cannot overemphasize the importance of safety and prevention in saving lives—this was the consensus high-impact topic at the end of the 2008 meeting in Berkeley (Appendix A).

Even with excellent public safety, SAR efforts are needed to handle the steady stream of low-probability events raised by people who get into trouble. The process of SAR involves four phases: *Notification, Planning, Search*, and *Rescue Recovery*.

The first phase—Notification—plays a key role in shaping the rest of the process. When notification of trouble occurs sufficiently quickly, the Planning and Search phases become trivial, and the SAR task can jump straight to the final Rescue Recovery phase.

SAR gets harder when notification is delayed, as it was in Jim Gray's case. This leads to an iterative process of Planning and Search. Initial Planning is intended to be very quick, and often consists simply of the decision to deploy planes to perform a visual sweep of the area where a boat is expected to be. When an initial "alpha" search is not successful, the Planning phase becomes more deliberate. The second "bravo" search is planned via software that uses statistical methods to model probabilities of the boat's location. The US Coast Guard has developed a software package for this process called SAROPS [3]. SAROPS treats the boat-location task as a probabilistic planning problem that it addresses with Bayesian machine learning techniques. Specifically, it accounts for prior information about weather and ocean conditions and the properties of the missing vessel, as well as the negative information from the alpha search. It uses a Monte Carlo particle filtering approach to infer a distribution of boat locations, and makes suggestions for optimal search patterns. SAROPS is an ongoing effort that is updated with models of various vessels in different states (broken mast, rudder missing, keel missing, etc.) The statistical "training" experiments to parameterize these models are expensive exercises that place vessels underway to track their movement. USCG continues to conduct these experiments on various parameters as funds and time permit. There is currently no equivalent software package or methodology available for land-based SAR.

2.2.2 Statistics and Discussion

Mr. Allen shared some of the US Coast Guard's statistics on Search and Rescue, which we include in Table 1. Most cases occur very close to shore, and many involve land-based vehicles going into the ocean. The opportunities for technologists to assist with maritime SAR are modest. In the US, less than 1,000 lives are lost in boating accidents each year, and roughly 600 people per year remain unaccounted for. Among fatalities, only 200-300 deaths occur after the Coast Guard is notified. Some fraction of the unaccounted-for cases include suicides. Relative to other opportunities to save lives with technology, the margins for improvement in maritime SAR are relatively small. This reality frames the rest of the discussion in this document, with a focus on learning lessons from this experience that apply both to SAR, and hopefully to other important settings as well.

3. COMMUNICATION AND COORDINATION

As in many situations involving groups of people sharing a common goal, communication and coordination became major efforts in the volunteer search for Jim Gray. The efforts to organize these

Year	Cases	Lives Saved	Lives Lost Before CG Notified	After CG Notified	Total	Unaccounted for
2003	31,562	5,104	409	246	655	481
2004	32,517	5,555	502	277	779	676
2005	29,780	5,648	521	320	841	606
2006	28.316	5,260	476	310	786	667

Table 1: US Coast Guard Search and Rescue statistics.

"back-office" tasks were ad hoc and evolving. In retrospect, some interesting patterns emerged around various themes related to Social Computing, including organizational development, brokering of volunteers and know-how, and communication with the media and general public. Many of these issues could be improved by better software.

3.1 Experience

The volunteer effort began via overlapping email threads among Gray's colleagues and friends in the hours and days following his disappearance. Various people exchanged ideas about getting access to satellite imagery, hiring planes, putting up missing persons posters, and so on. Many of these tasks involved reaching out in a coordinated and thoughtful manner to third parties. But it was unclear who was hearing what information, and who might be contacting third parties. To solve that problem a blog called *Tenacious Search* was set up to allow a broadcast style of communication among the initial group of participants [11]. Initially, authorship rights on the blog were left wide open. This simple "blog-as-bulletin-board" worked well for a day or two to coordinate the initial group involved in the search, loosely documenting various people's questions, efforts, skills and interests in a way that helped define the group's efforts and organization.

Within a few days the story of Jim Gray's disappearance was widely known, and the blog transitioned from being an in-group communication medium to a widely-read publishing venue for status reports on the search effort. It served this role for the remainder of the volunteer search. This function was quickly taken seriously, so authorship on the blog was closed to additional members, and a separate "Friends of Jim" mailing list was set up for internal team communications. This transition led to an increased sense of organizational and social structure within the core group of volunteers.

Over the next few days, various individuals stepped into unofficial central roles for reasons of expediency, unique skills, or both. The blog administrator evolved into a general "communications coordinator", handling messages sent to a public email box for tips, brokering skill-matching for volunteers, and serving as a point of contact with outside parties. Another volunteer emerged as a "aircraft coordinator", managing efforts to find, pilot and route private planes and boats to search for Gray. A third volunteer took on the role of "analysis coordinator", organizing various teams working on image analysis and ocean drift modeling at a variety of organizations around the United States. A fourth person was chosen by Gray's family to serve as a central "media coordinator", the sole contact for press and public relations. These coordinator roles were identified in retrospect, and the role names were coined for this paper simply to clarify discussion. Individuals with management experience in the business world provided guidance along the way, but much of the organizational development happened in an organic, "bottom-up" mode.

On the communications front, an important task that quickly emerged was to do *task brokering* between skilled or well-resourced volunteers, and people who could take advantage of those assets. This started happening in an ad-hoc broadcast mode: someone would make a suggestion to the whole team, and a volunteer would broadcast an intention to follow up. For example, one member of the team volunteered to raise funds to pay for pilots and search planes if someone else would coordinate the operational aspects. But as the search progressed publicly, offers of help came in from various unexpected quarters, and coordination and task brokering became more complex. Volunteers with science and military backgrounds emerged with offers of specific technical expertise and suggestions for acquiring and analyzing particular satellite imagery. Other volunteers offered to search in private planes and boats, sometimes at serious danger to their own lives. (Many of these were dangerous, and discouraged by the team and the USCG.) Many volunteers offered to post "Missing Sailor" posters at marinas; this required coordination. There were a number of offers of psychic assistance. Each offer took time from the communications coordinator to effectively and diplomatically pursue and route onward, or deflect. As subteams emerged within the organization, this became easier: the communications coordinator could skim an inbound message and route it to one of the other volunteer coordinators for follow-up.

Similar information brokering challenges arose in handling thousands of messages from the general public, who were encouraged by the media to keep their eyes open for boats and debris, and report to a public email address. This included a variety of information along with many prayers and thoughts from well-wishers. Some of the tips were clearly useful, some clearly bogus or irrelevant. Many were somewhere in between, and—given the sense of urgency at the time—it was often hard to decide whether to bring them to the attention of busy people: the Coast Guard, police, Jim Gray's family, and technical experts in image analysis and oceanography. In some cases tipsters got in contact repeatedly, and it became necessary to assemble conversations over the period of days to establish a tipster's credibility. This became burdensome as the volume of email grew.

The choice to involve the media in the effort generated work as well. The goal of the media outreach was to keep the story of the disappearance in the news, in hopes of maintaining public vigilance. But not all media outlets were equally respectful of this goal, and some attempted to "dig for dirt" among members of the team. After a short time, Jim Gray's family asked that all media contact be handled through a single volunteer, with some advice from a professional in public relations. The team was quite disciplined in respecting this request, particularly since it came from the family.

3.2 Discussion

On reflection, the organization's evolution was one of its most interesting aspects. Leadership roles emerged fairly organically, and subgroups formed with little discussion or contention over the process or outcomes. Some of the people had certain baseline competencies: e.g., the aircraft coordinator was a recreational pilot, and the analysis coordinator had both management experience and contacts with image processing experts in industry and government. In general, though, leadership developed by one person stepping up to take on a responsibility, and others stepping back to let them do the job, and then jumping in to help as needed. The grace with which this happened is a bit surprising, given the vigorously ambitious people that surrounded Jim Gray, and the fact that the organization evolved largely over email. The evolution of the team seems worthy of a case study in ad-hoc organizational development during crisis.

It became clear that better software is needed to facilitate group communication and coordination during crises. By the end of the search for *Tenacious*, a variety of standard communication methods were in use: point-to-point email and telephony, broadcast via blogs and web pages, and multicast via conference calls, wikis and mailing lists. This mix of technologies was natural and expedient in the moment, but made communication and coordination challenging. It was hard to work with the information being exchanged, which was represented in natural language text and stored in multiple separate repositories. As a matter of expediency in the early days, the communications cordinator relied upon mental models of basic information like who knew what information, and who was working on what tasks. The emphasis on mental note-taking made sense in the short term, but limited the coordinator's ability to share responsibility with others as the "crisis watch" extended from hours to days to weeks. Much of the coordination effort was probably similar to "command and control" challenges in military or emergency response teams, with the key difference that it was being staffed by amateurs who were learning the process on the fly, using whatever tools were at hand.

Various pieces of this problem are addressable with well-known information management techniques. But using current communication software and online services, it remains difficult to manage an evolving discussion that includes individuals, restricted groups, and public announcements, especially in a quickly-changing "crisis mode" of operation. Identifying people and their relationships is challenging across multiple communication tools and multiple recipients' endpoints. Standard search and visualization metaphors—folders, tags, threads—are not well-matched to group coordination. The problem of brokering volunteers and tasks brings up further problems. In Appendix D.1 we list some specific challenges in this domain in more detail. In any software approach to these problems, one issue is critical: in an emergency, people do not reach for new software tools. So it is important to attack these challenges in a way that augments popular tools, rather than seeking to replace them.

4. IMAGERY ACQUISITION

When the volunteer search began, our hope was to use our unique skills and resources to augment the search efforts of the USCG with satellite imagery and private planes. As we learned, real-time search for boats at sea is not as simple as getting a satellite feed from a mapping service, or borrowing a private jet.

4.1 Experience

The day after *Tenacious* went missing, Jim Gray's friends and colleagues began efforts to access satellite imagery and planes. One of the first connections made was to colleagues in Earth Science who had expertise in remote sensing. In an email message in the first few days, one of the Earth scientists explained the difficulty of using satellite imagery to find *Tenacious*: "The problem is that the kind of sensors that can see a 40ft (12m) boat have a correspondingly narrow field of view, i.e. they can't see too far either side of straight down... So if they don't just happen to be overhead when you need them, you may have a long wait before they show up again. ... [A]t this resolution, it's strictly target-of-opportunity."

Undeterred, the team pursued multiple avenues to acquire remote imagery via connections at NASA and other government agencies, and at various commercial satellite imagery providers. The satellite data teams at both Google and Microsoft directed us to their commercial provider, Digital Globe. Table 2 outlines the data sources that were considered during the search for Jim Gray. Figure 5 is taken from the website that was put together at the time of the search to geo-position and catalog the imagery that was eventually made available [1]. It shows the geographic boundaries of both the data captures in the imagery, and the Coast Guard's airborne search.

As we discovered, distribution of satellite data is governed by national and international law. We attempted from the very beginning of our discussion to get data from the SPOT 5 satellite, but this effort was halted by the US State Department, which invoked the International Charter on Space and Major Disasters to claim exclusive access to the data over the study area, retroactive to the day before our request. We also learned, when getting data from Digital Globe's QuickBird satellite, that full-resolution imagery is only available after a government-mandated 24-hour delay; before that time Digital Globe could only provide reduced-resolution images.

The first data acquired from Digital Globe's QuickBird satellite was focused well south of San Francisco near Catalina Island. There were very small odds that *Tenacious* would be found in that region. On the other hand, it seemed important to begin experimenting with real data and seeing how the team could process it. This early start turned out to be critical for getting the various pieces of the image processing pipeline in place and tested. Even data acquisition was challenging, and Digital Globe was extremely generous with their time and resources, very aggressively producing and sharing their data. As the search progressed, Digital Globe was able to acquire imagery solidly within the primary search area, and the image captures provided to our team formed some of the biggest data products Digital Globe had ever generated. Even so, the areas covered by the satellite captures were dwarfed by the airborne search conducted by the Coast Guard immediately after Gray went missing (Figure 5).

We were able to establish contacts at NASA regarding planned flights of the ER-2 "flying laboratory" aircraft over the California coast. The ER-2 is typically booked on scientific missions, and of course requires resources to launch under any circumstances (fuel, airport time, staffing, wear-and-tear, etc.) As it happened, the ER-2 was scheduled for training flights in the area where *Tenacious* disappeared. Our contacts were able to arrange flight plans to pass over specific areas of interest and record various forms of digital imagery: a combination of fortunate circumstance and a well-connected social network. Unfortunately, a camera failure early in the ER-2 flight limited the data collection.

In addition to these relatively rare imaging resources, we chartered private planes to fly over the ocean, enabling volunteer spotters to look for *Tenacious* with naked eyes, and record digital imagery. This effort ended up being more limited than we originally expected.

As we learned, one cannot simply charter or borrow a private jet and fly it out over the ocean. Light planes are not safe to fly far offshore. Very few individuals maintain planes equipped for deep-sea search, and flights over deep sea can only be undertaken by pilots with appropriate maritime survival training. Finally, aircraft of any size require a flightplan to be filed and approved with a US Flight

RADARSAT-1	A commercial Earth Observing Satellite (EOS) from Canada, whose products are distributed by MDA Geospatial Services. NASA has access to RADARSAT-1 Data, in exchange for having provided a rocket to launch the satellite. `http://en.wikipedia.org/wiki/RADARSAT-1`
Ikonos	A commercial EOS operated by GeoEye (USA). `http://en.wikipedia.org/wiki/IKONOS`
QuickBird	A commercial EOS owned and operated by Digital Globe (USA). It was in use at the time by Google Earth and MS Virtual Earth. `http://en.wikipedia.org/wiki/QuickBird`
ER-2	A high-altitude aircraft operated by NASA similar to the US Air Force's U2-S reconnaissance platform. `http://www.nasa.gov/centers/dryden/research/AirSci/ER-2/index.html`
SPOT-5	A commercial EOS operated by SPOT Image (France). `http://en.wikipedia.org/wiki/SPOT\_(satellites)`
Envisat	A commercial EOS launched by the European Space Agency. Data products are distributed by the SARCOM consortium, created and led by SPOT Image. `http://en.wikipedia.org/wiki/Envisat`

Table 2: Some of the remote imagery sources considered during the search for Jim Gray. An extensive list of Earth Observing Satellites is kept at `http://en.wikipedia.org/wiki/List_of_Earth_observation_satellites`.

Service Station in order to cross the US Air Defense Identification Zone (ADIZ) that begins a few miles offshore. As a result of these limitations and many days of bad weather, the number of private overflights we arranged was small, and all but one were close to shore.

Another source of imagery that was considered was land-based video cameras. These could have more accurately established a time of departure for *Tenacious*, beyond what we knew from Gray's mobile phone calls to family members on his way out. The Coast Guard operates a camera on the San Francisco waterfront pointed out toward the Golden Gate and the ocean, but much of the tape for that day was in a state of "white-out" rather than useful imagery, perhaps due to weather.

4.2 Discussion

The volunteer search effort was predicated on quick access to satellite imagery. This was surprisingly successful — over 87 Gigapixels of satellite imagery were acquired from Digital Globe alone within about four days of being captured. Yet in retrospect we would have liked to get much more data, with fewer delays. In Appendix refsec:imageacq we review some of the limitations we encountered, and some ideas and challenges in improving the ability to acquire imagery in life-threatening emergencies.

Policy concerns come up naturally when discussing large volumes of remote imagery. Various members of the amateur team voiced concerns about personal privacy during the process. In addition, national security issues arose, including our inability to access SPOT-5 data. (While we were not given reasons for limitation on access, we speculate that certain obvious maritime features like military fleet movements are classified for good reason.) Meanwhile, popular media-sharing websites already provide widespread access to crowdsourced imagery, and recent work has demonstrated the power of aggregating individual images into richer data products like panoramas and 3-d views. While aggregation applications to date have largely confined themselves to benign settings like tourism and ornithology, maritime SAR applications (e.g., monitoring marinas and shipping lanes) seem closer to pure surveillance. This raises understandable concerns, and the policy issues are not simple. Perhaps our main observation on this front was the need for a *contextual* treatment of policy [8], balancing general-case social concerns against specific circumstances for using the data—in our case, for trying to rescue a friend. While the search for *Tenacious* was uniquely urgent for us, similar life-and-death scenarios occur on a national scale with some frequency. So it seems natural to think about technical solutions that can both aggressively harvest and process imagery, while provably respecting policies that limit image release based on context.

5. FROM IMAGERY TO COORDINATES

In this section, we discuss the processing pipeline(s) and coordination mechanisms used to reduce the raw image data to *qualified* search coordinates—locations to which planes were dispatched for a closer look. This aspect of the search was largely data-driven, and involved significant technical expertise. It required much more structured and tool-intensive processes than those described in Section 3. On the other hand, since time was short and the relevant expertise was quite specialized, it also led to simple interfaces between teams and their software. The resulting amalgam of software was not the result of a specific architecture, in the usual sense of the word (*archi-* "chief" + *techton* "builder"). A more apt term for the software and workflow described here might be a *polytechture*—the kind of system that emerges from the design efforts of many independent actors. Though our experience is limited, certain "polytechtural" aspects of the effort appear interesting in retrospect, and we describe them with the hope that the processes that emerged in our effort may inform work in other contexts.

5.1 Overview

Figure 1 is a simplified illustration of the critical-path data and control flow that emerged. This diagram and the ensuing discussion in this section depicts the "ad-hoc" pipeline that was developed for Digital Globe's satellite imagery. In Sections 5.3.1 and 5.3.2 we describe the Mechanical Turk pipeline that was primarily used for NASA ER-2 overflight imagery. RadarSat data was handled in a much less structured way: a few members of the imagery team looked at it and quickly decided it would not be helpful.

Before diving into the details, it is instructive to work "upstream" through the pipeline, from the final qualified targets back to the initial imagery. The objective of the exercise was to identify one or more sets of qualified search coordinates to which aircraft could be dispatched (lower right of Figure 1). To do so, it was not sufficient to simply identify the coordinates of qualified targets on the imagery—we had to apply a mapping function to those coordinates to compensate for drift of the target from the time of image capture to the flight time. In our case, this mapping function was provided by two independent "drift teams" of volunteer oceanographers: one based at the Monterey Bay Aquarium Institute and Naval Research Lab (MBARI/NRL), and another at NASA Ames ("Ocean Drift Modeling" in Figure 1, described in more detail in Section 5.5.)

The careful qualification of search coordinates was particularly important. It was quickly realized that many of the potential search coordinates would be far out at sea; as mentioned in Section 4,

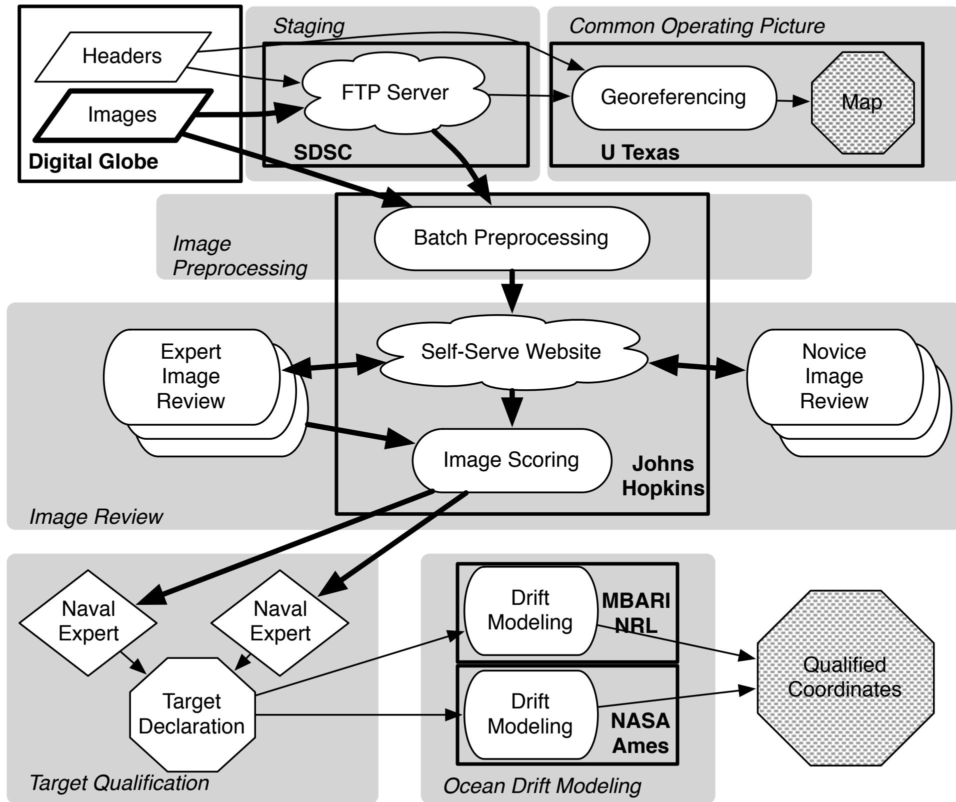

Figure 1: The rough dataflow for image processing. Bold arrows represent images; other arrows represent metadata.

they would require special aircraft and crews. Furthermore, flying low altitude search patterns offshore and in single-engine aircraft implied a degree of risk to the search team. Thus, it was incumbent on the analysis team to carefully weigh this risk before declaring a target to be qualified. A key step in the process was the review of targets by naval experts prior to their final qualification ("Target Qualification" in Figure 1, Section 5.4).

Prior to target qualification, an enormous set of images had to be reviewed and winnowed down to a small set of candidates that appeared to contain boats. To our surprise and disappointment, there were no computer vision algorithms at hand that were well-suited to this task, so this was done manually. At first, the image analysis effort was achieved by using the Mechanical Turk infrastructure at Amazon to coordinate volunteers from around the world — mostly novices in the task of image analysis. Subsequently, a distributed team of volunteer experts in image analysis (including astronomers, radiologists, geologists, machine vision researchers and former military personnel) performed the review function ("Image Review" in Figure 1, Section 5.3).

Shifting to the start of the pipeline, each image dataset required a degree of pre-processing prior to the human analysis of the imagery; this step was performed by astronomers at Johns Hopkins. At the same time, a separate team at the University of Texas took the image file headers and georeferenced them onto a map, which was included in a web interface for tracking the progress of image analysis ("Image Preprocessing", "Common Operating Picture" and "Staging" in Figure 1, Section 5.2).

In sum, the eventual workflow was a distributed, multi-party process. The pieces of this workflow were designed and built individually, "bottom-up", by independent volunteer teams at various institutions. The teams also had to quickly craft interfaces to stitch together the end-to-end workflow with minimal friction. An interesting variety of design styles emerged depending on a variety of factors. In the following subsections, we describe the above components in greater detail, this time from start to finish. Since our processing pipeline increased in sophistication during the course of the search, we typically only describe the state of the evolved pipeline stages, unless their evolution itself is of interest.

5.2 Pre-processing

Once the image providers had data and the clearance to send it, they typically sent notification of the availability via email to the image analysis coordinator, together with an FTP address and the *header* file that described the collected imagery ("the collection").

Immediately upon notification, the pre-processing team at Johns Hopkins began copying the data to their cluster. Meanwhile, the common storage repository at the San Diego Supercomputer Center (SDSC) began ftp-ing the data to ensure its availability, and a copy of the header was passed to a separate geo-coordination team at the University of Texas, which mapped the location covered by the collection and added it to a website. The Texas website provided the overall shared picture of imagery collected and analyses completed, and was used by many parts of the search team to track progress and solicit further collections.

5.3 Analysis Tasking and Result Processing

During the course of the search two approaches to the parallel processing of the tiled images were used. In both of these approaches small image tiles (or smaller sub-tiles) had to be farmed out to human analysts, and the results of their analysis collated and further filtered in order to avoid a deluge of false positives.

5.3.1 Tasking Novices with Mechanical Turk

The initial tasking approach was to use Amazon's Mechanical Turk service [6] to solicit and task a very large pool of anonymous reviewers whose credentials and expertise were not known to us.

Mechanical Turk is a "crowdsourcing marketplace" for coordinating the efforts of humans in performing simple tasks from their own computers. Given that the connectivity and display quality available to these users was unknown, the Mechanical Turk was configured to supply users with *HITs* (work items) that consisted of a few 300 × 300-pixel image sub-tiles each. Using a template image we provided of what we were looking for, the volunteer was asked to score each sub-tile for evidence of similar features, and provide comments on artifacts of interest. This was an exceptionally slow process because of the number of HITs required to process a collection.

In addition to handling the partitioning of the imagery across volunteers, Mechanical Turk bookkeeping was used to ensure that each sub-tile was redundantly viewed by multiple volunteers prior to declaring the pipeline "complete". At completion, and at checkpoints along the way, the system also generated reports aggregating the results received concerning each sub-tile.

False positives were a significant concern even in the early stages of processing. So a virtual team of individuals who identified themselves as having some degree of familiarity with image analysis (though typically astronomical or medical imagery rather than satellite imagery) was assembled to perform this filtering. In order to distribute the high-scoring sub-tiles to them, the image analysis team configured an iterative application of Mechanical Turk accessible only to this sub-team, with the high scoring sub-titles from the first pipeline fed into it. The reports generated by this second pipeline were then used by the coordinator to drive the target qualification process. This design pattern of an "expertise hierarchy" seems likely to have uses in other crowdsourcing settings.

5.3.2 Tasking Experienced Users

A significant cluster of our image reviewers were co-located at the astronomy research center at Johns Hopkins. These volunteers—who had ample expertise, bandwidth, high-quality displays and a sense of personal urgency—realized they could process the imagery much faster than novices scheduled by Mechanical Turk could. This led to two modifications in the granularity of tasking.

These individuals were accustomed to looking for anomalies in astronomical imagery, and were typically able to rapidly display, scan and discard sub-tiles that were 3-4× larger than those presented to amateurs. This led to an individual processing rate of approximately one (larger) sub-tile every 4 seconds, including the tiles that required detailed examination and the entry of commentary, as compared to the 20-30 second turnaround for each of the smaller sub-tiles in the Mechanical Turk HITs. The overall improvement in productivity over Mechanical Turk was in fact considerably better than these numbers indicate, because the analysts' experience reduced the overhead of redundant analysis, and their physical proximity facilitated communication and cross-training.

A further improvement was that the 256 sub-tiles within each 8k×8k pixel tile were packaged into a single zip file. Volunteers could then use their favorite image browsing tools to page from one sub-tile to the next with a single mouse click. To automate the tasking and results collection, this team used scripting tools to create a web-based visual interface through which they (and similarly equipped volunteers worldwide) could visually identify individual tiles that required work, download them and then submit their reports.

In this interface, tiles were super-imposed on a low resolution graphic of the collection that was, in turn, georeferenced and superimposed on a map. This allowed participants to prioritize their time by working on the most promising tiles first, e.g., those that were not heavily obscured by cloud cover.

The filtering component of this pipeline also operated somewhat more collaboratively and synchronously. The expert analysts worked in "shifts", and at the end of each shift the sub-team leaders gathered together and scored the most promising targets. Although the scoring of extremely promising targets was performed immediately, this periodic and collective review promoted discussion among the analysts, allowing them to learn from each other and adjust their individual standards of reporting.

5.4 Target Qualification

The analysis coordinator examined reports from the analysis pipelines to identify targets for submission to the qualification step. In the Mechanical Turk case this involved spending a few hours sifting through the output of the second Mechanical Turk stage. Once the expert pipeline was in place, the coordinator was examining only a few filtered and scored targets per shift.

Promising targets were then submitted to a panel of two reviewers, each of whom had expertise in identifying engineered artifacts in marine imagery. The analysis coordinator isolated these reviewers from each other, in part to avoid cross-contamination, but also to isolate them from individually carrying the weight of a potentially risky decision to initiate a search mission—this would avoid overly biasing them in a negative direction. Having discussed their findings with each of the reviewers, the coordinator made the final decision to designate a target as qualified.

Given the dangers of deep-sea flights, this review step included an intentional bias: it imposed less rigorous constraints on targets that had likely drifted close to shore than on those further at sea.

5.5 Drift Modeling

Relatively early in the analysis process, an individual with marine expertise recognized that, should a target be qualified, it would be necessary to estimate its movement since the time of image capture. They formed a drift modeling team that ultimately consisted of two sub-teams of oceanographers with access to two alternative drift models. As the image processing proceeded, these sub-teams worked in the background to parameterize their models on an ongoing basis with weather and ocean surface data during the course of the search. Thus, once targets were identified, the sub-teams were able to rapidly estimate the likely drift patterns.

The drift models utilized a particle filtering approach of virtual buoys that could be *released* at an arbitrary time and location, and for which the model would then produce a projected track and likely endpoint at a specified end-time. In practice, one releases a string of adjacent virtual buoys to account for the uncertainty in the initial location and the models' sensitivity to local effects that can have fairly large impacts on buoy dispersion. Figure 2 illustrates the results of one such model. The availability of two independent models, with multiple virtual buoys per model, greatly increased our confidence in the prediction of regions to search.

It is worth noting that although these drift models were developed by leading scientists in the field, the results often had significant uncertainty. This was particularly true in the early part of the search, when drift modeling was used to provide a "search box" for Gray's boat, and had to account for many scenarios including whether the boat was under sail or with engines running. These scenarios had very large uncertainty and led to large search boxes. By the time the image processing and weather allowed for target qualification, the plausible scenario was reduced to a boat adrift from

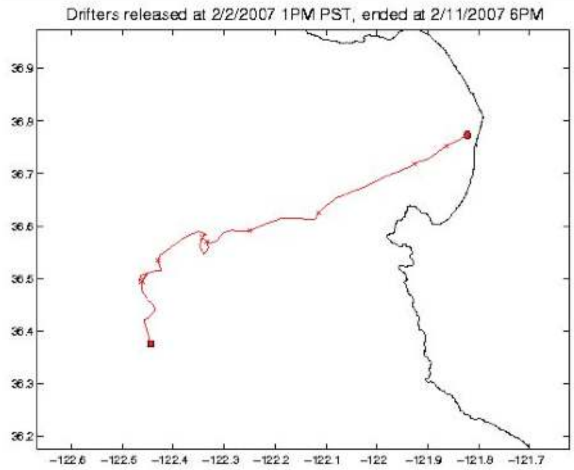

Figure 2: A drift model output showing an eventual landing point in Monterey Bay.

Figure 3: Cropped version of an ER-2 satellite image flagged by an inexperienced Mechanical Turk volunteer.

a relatively recent starting point. Our colleagues in Oceanography and the USCG tell us that the problem of ocean drift modeling merits more research and funding; it would seem to be a good area for collaboration with Computer Science.

The drift modeling team developed its own wiki-based workflow interface for their tasks. The analysis coordinator was given a website where he could enter a request to release virtual "drifters" near a particular geolocation at a particular time. Requests were processed by the two trajectory modeling teams, and the resulting analysis, including maps of the likely drift patterns, were posted back to the coordinator via the website. Geolocations in latitude/longitude are hard to transcribe accurately over the phone, so this helped ensure correct inputs to the modeling process.

5.6 Analysis Results

The goal of the analysis team was to identify qualified search coordinates. During the course of the entire search, numerous targets were identified but only two were qualified. One was in ER-2 flyover imagery near Monterey, and was originally flagged by Mechanical Turk volunteers (Figure 3). The other was in Digital Globe imagery near the Farallon Islands, and was identified by a member of the more experienced image processing team (Figure 4.) Although the low number might suggest our filtering of targets was overly-aggressive, we have no reason to believe that potential targets were missed. Our conclusion is simply that the ocean surface is not only very large, but also very empty.

Once qualified, these two targets were then drift-modeled so that coordinates for search boxes could be identified. For the first of these targets, the drift models indicated that the target should have washed up on shore in Monterey Bay. Because this was a region very close to shore, it was relatively easy to send a private plane to the region, which we did.

The second target was initially not far from the Farallon Island and both models predicted it to have remained relatively close by. Given our knowledge of Gray's intended course for the day, this was a very promising target, so we arranged a private off-shore search-and-rescue flight. Although we did not find *Tenacious*, we did observe a few fishing vessels of approximately *Tenacious*' size in the area. It is possible that the target we identified was one of those vessels. Though the goal of the search was not met, at a technical level this provided some validation of the targeting process.

5.7 Discussion

The image processing effort was the most structured and technical aspect of the volunteer search. In trying to cull lessons from it, we highlight three rough topics: the "polytechtural" design style, networked approaches to search, and the need for civilian computer vision research targeted at disaster response applications. In Appendix C.2 we also discuss organizational issues that arose in this more structured aspect of the search.

5.7.1 Polytechture

The software development and deployment process that emerged was based on groups of experts working independently. Some of the more sophisticated software depended on pre-existing expertise and componentry (e.g., parallelized image processing pipelines, and sophisticated drift modeling software). By contrast, some software was ginned up for the occasion, building on now-standard web tools like wikis, scripting languages, and public geocoding interfaces. Not too many years earlier, these tools were typically used for low-throughput "rapid prototypes". They are now increasingly used for production systems, and, in our case, for a hybrid "rapid production system": a system that had to be built quickly, and also get a real job done. It is encouraging to see how much was enabled by these lightweight tools.

Redundancy was an important theme in sensitive steps of the process. Redundant ftp sites ensured availability, redundant drift modeling teams increased confidence in predictions, and redundant target qualification by experts provided both increased confidence and limits on "responsibility bias".

Perhaps the most interesting aspect of this process was the variety of interfaces that emerged to couple these independent components. A number of these are described above: the cascaded Mechanical Turk interface for hierarchical expertise in image analysis, the ftp/email scheme for data transfer and staging, the web-based "Common Operating Picture" for geolocation and coarse-grained task tracking, the self-service "checkin/checkout" interface for expert image analysis, the decoupling of image file access from image browsing software, and the transactional workflow interface for drift modeling. The variations in these interfaces seemed to arise from both the tasks at hand, and the people involved. Sometimes a

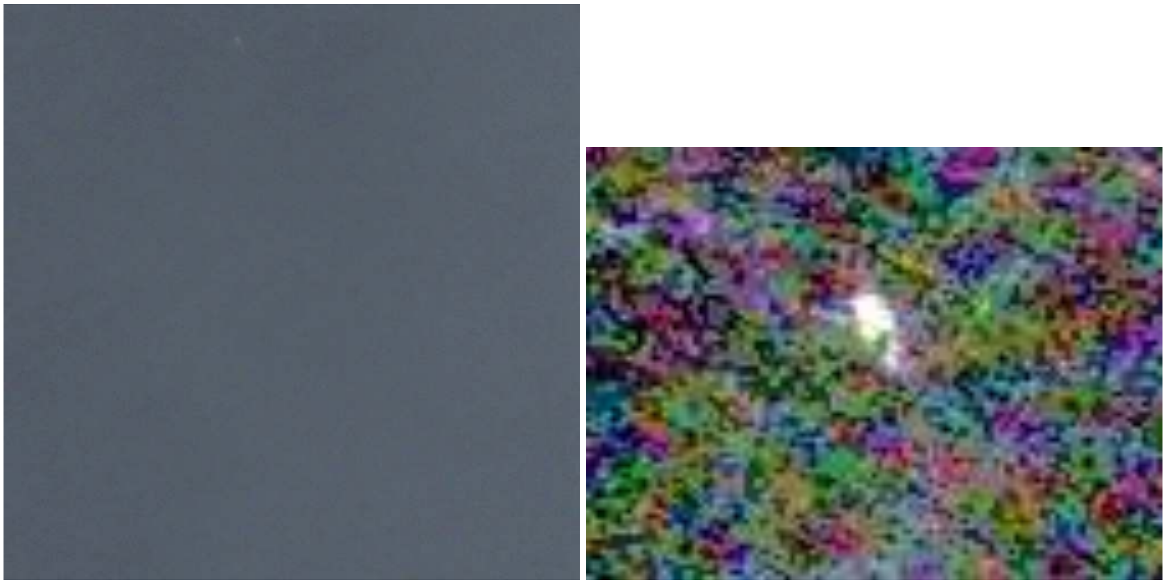

Figure 4: A Digital Globe satellite image flagged by an experienced volunteer. On the left it is in its original format—gray with clouds and hard to reproduce on paper. On the right is an enhanced, cropped version of the upper center of left image. The white spot was confirmed by naval imagery experts to very likely be a boat, of a size consistent with *Tenacious*.

structured process seemed necessary (e.g. the Mechanical Turk interface for tasking amateurs), and sometimes it was desired by one party (e.g. the web interface to the drift team). Other tasks evolved ad hoc (e.g. data staging, and the bulk of the initial communication and coordination tasks.)

The evolution of the web over the last decade enabled this polytechtural design. One key aspect of this is the standardization of software tools and design patterns: the commonality of web-based interfaces to remote systems and services, the availability of easy-to-use scripting languages, and the widespread standardization of image formats and viewers. Perhaps most remarkable were the interactions between public data and global communications. The manufacturer's specifications for *Tenacious* were found on the web, aerial images of *Tenacious* in its berth were found in publicly available sources (Google Earth, Microsoft Virtual Earth), and a former owner of *Tenacious* discovered the blog in the early days and provided additional photos of *Tenacious* under sail. These details were helpful for parameterizing drift models, and for providing "template" pictures of what analysts should look for in their imagery. Despite its inefficiencies, the use of Mechanical Turk volunteers to bootstrap the image analysis process was remarkable, particularly the benefit of having many people redundantly doing data analysis. Beyond the Turk pipeline, an interesting and important data cleaning anecdote occurred on the blog, in building the template picture for *Tenacious*. Initially, a family member identified *Tenacious* in the satellite image geographically: by identifying Gray's boat slip in the marina. In subsequent discussion, one of the analysts noticed that the boat in that image did not match *Tenacious*' online specifications, and after some reflection the family member confirmed that Gray had swapped boat slips some years earlier, and the online satellite image predated the swap. Few if any of these activities would have been possible 10-15 years ago—not because of the march of technology per se, but because of the enormous volume and variety of information that is placed online, and the rapidly increasingly subset of the population that is habituated to using that information.

5.7.2 Networked Search

It is useful to reflect on the relative efficacy of this componentized polytechtural design approach, compared to more traditional and deliberate strategies. The amateur effort was forced to rely on loosely coupled resources and management, operating asynchronously at a distance. By contrast, the Coast Guard operates in much more prepared and tightly-coupled manner, performing nearly all the search steps at once, in real time: after a planning phase maps out the maximum radius that a boat can travel, trained officers fly planes in carefully-plotted flight patterns over the relevant area, using real-time imaging equipment and their naked eyes to search for targets. This allows the Coast Guard searches to be nimble and focused. But a componentized approach may offer certain advantages in scaling and evolution, since it does not rely upon tightly-integrated and relatively scarce human and equipment resources.

A compromise architecture is what might be called *Networked Search Infrastructure*, in which the relevant components of the search process are decoupled in a manner akin to our volunteer search, but more patiently architected, evolved, and integrated. As a simple example, Coast Guard imagery experts need not be available to board search planes nationwide: a remote image analysis team could examine streaming (and archived) footage from multiple planes in different locales. In fact, weather hazards and other issues suggest removing the need for people to board planes entirely: imagery could be acquired via satellites and unmanned aerial vehicles (UAVs), which are improving rapidly. A componentized approach takes advantage of the independent evolution of technologies, and the ability to train domain experts on each component more quickly: image analysis tools can improve separately from imaging equipment, which can evolve separately from the devices flying the equipment. This kind of networked componentry and expertise is becoming relatively common in military settings, and in medical imaging in the public sector. It seems useful to explore these ideas further for other civilian settings like SAR, especially when considering applications to widespread disaster response.

5.7.3 Automated Image Analysis

The volunteer search team included experts in image processing for astronomy, and consulted with experts in computer vision as well. The consensus early on was that off-the-shelf image recognition software would not be sufficiently accurate for the urgent task of identifying boats in satellite imagery of the ocean. During the course of the search a number of machine vision experts examined the available datasets and concluded that it was not of sufficient quality for automated processing, though that may have been because we did not have access to the "raw bits" obtained by the satelite-based sensors. Although some experts attempted a simple form of automated screening by looking for clusters of adjacent pixels that stood out from the background, even those efforts were relatively unsuccessful.

It would be good to know if this problem is inherently hard, or simply requires more focused attention from computer vision researchers. The problem of using remote imagery for search-and-rescue is a topic where computer vision would seem to have a lot to offer, especially at sea where obstructions are few.

6. REFLECTION

Having described the amateur SAR processes cobbled together to find *Tenacious*, we can return to some of the issues we set out to discuss initially when we the group of searchers met in Berkeley in 2008.

Some of the lessons learned along the way are mentioned above; more specific challenge problems are presented in Appendix D. As to whether this kind of effort could be democratized, there is reason to be optimistic on the computational front. Computer hardware has continued to shrink in price since early 2007, and cloud services are commoditizing access to large computational clusters: it is is now affordable to get rapid access to enormous computing resources without any social connections or up-front costs. By contrast, custom software pipelines to perform tasks like image processing, drift modeling, and command-and-control coordination are not widely available. This is not an inherent problem—it is an area where small teams of open source developers and software researchers could have significant impact. The key barrier to democratization of these efforts may be access to data. It is unclear whether data providers like satellite imagery companies or plane owners can support large-scale, near-real-time feeds of public safety imagery. And it is also not clear, from a policy perspective, whether this is an agreed-upon social good. This topic deserves more public discussion and technical investigation. This discussion can be accelerated by low-fidelity open-source prototypes that make the best of the data that is publicly available: e.g., by aggregating volunteer webcams as we outline in Appendix D. Sometimes the best way to democratize access to resources is to build disruptive low-fidelity prototypes.

The volunteer search team's experience reinforces the need and opportunity for technical advances in Social Computing. In the end, the team exploited technology for many uses, not just the high-profile task of locating *Tenacious* in images from space. Modern networked technologies enabled a group of acquaintances and strangers to quickly self-organize, coordinate, build complex working systems, and attack problems in a data-driven manner. Along the way, we found plenty of room for technological improvement, including the limitations of standard email and blogging tools, and design patterns in architecting workflows to coordinate volunteers with differing skills. More work is needed here.

The efforts documented here are not the whole story of the search for *Tenacious* and its skipper: in addition to incredible work by the USCG, there were other, quieter efforts among Jim Gray's colleagues and family that happened outside the public eye. Although we were frustrated in achieving our goal, the work that was done along the way was remarkable in many ways, and the tools and systems that were developed so quickly by the amateur team actually worked startlingly well. This was due in part to the incredible show of heart and hard work from the volunteers, for which many people will always be grateful. It is also due to the quickly maturing convergence of people, communication, computation and sensing on the Internet. Jim Gray was a shrewd observer of technology trends, and what they suggest for the next important steps in research. We hope that the search for *Tenacious* sheds some light in those directions as well.

7. REFERENCES

[1] Center for Space Research, University of Texas, Austin. Emergency response, 2007. `http://magic.csr.utexas.edu/tenacious.htm`).

[2] Joshua Goldstein and Juliana Rotich. Digitally networked technology in Kenya's 2007-2008 post-election crisis. Technical Report 2008-09, Berkman Center for Internet and Society at Harvard University, September 2008.

[3] U.S. Coast Guard. Search and Rescue Optimal Planning System (SAROPS), 2009. `http://www.uscg.mil/acquisition/international/sarops.asp`.

[4] Jessica Heinzelman and Carol Waters. Crowdsourcing crisis information in disaster-affected Haiti. Technical Report Special report 252, United States Institute of Peace, October 2010.

[5] KDD Cup Center, 2009. `http://www.sigkdd.org/kddcup/index.php`.

[6] Mechanical Turk, 2007. `http://www.mturk.com`.

[7] Claire Cain Miller and Jenna Wortham. Technology aside, most people still decline to be located. *The New York Times*, August 29, 2010.

[8] Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford Law Books, 2009.

[9] E. Saade. Search survey for S/V Tenacious: Gulf of Farallones and approaches to San Francisco Bay. *ACM SIGMOD Record*, 37(2), June 2008.

[10] D. Song and K. Goldberg. Networked robotic cameras for collaborative observation of natural environments. In *International Symposium of Robotics Research (ISRR)*, October 2005.

[11] Tenacious Search blog, 2007. `http://www.openphi.net/tenacious/`.

APPENDIX

A. PREVENTION

In Section 2.2 we noted that the best way to improve the chances of Rescue and Recovery are to provide Notification of distress early, taking the Search and Planning phases out of SAR. In addition to improving the ability to locate individuals, a major focus of the US Coast Guard is to shrink the time to notification, which can radically decrease the complexity of Planning, Search and Rescue Recovery. In this section we discuss some standard practices in boating safety, and ways that these can be improved via both policy changes and technology.

A.1 Background and Experience

Radio beaconing devices such as the Emergency Positioning Radio Beacon (EPIRB) and Personal Locator Beacon (PLB) are available to sailors, producing signals on satellite channels monitored by agencies including the US Coast Guard. Many of these devices are activated when a hydrostatic switch hits the water: when mounted properly, they will automatically send out a radio beacon that is detected via a geo-positioning satellite known as SARSAT (Search and Rescue Satellite-Aided Tracking). Cheaper models have a manual switch. In the United States, EPIRBs are legally required of most commercial vessels. They are not mandatory for recreational boats, though they are strongly encouraged. Many recreational boaters do not own EPIRBs, or if they do own one it may be mounted incorrectly or not kept in working order.

In addition to emergency communication devices, most boaters today carry cell phones as a matter of course, and keep them in working condition. Cell phones are useful not only for communication, but also for geolocation by third parties tracking their signals. Also, many smartphones connect periodically to Internet services like mail- and location-servers, signalling that the device is still functioning.

Tenacious was outfitted with an EPIRB, but no signals were received from it the day that Jim Gray went missing. There are many possible explanations for this failure. It may simply have malfunctioned, or it may not have been properly maintained. In addition, the location of the EPIRB on *Tenacious* was not fail-safe: it could possibly have become trapped in the hold as the boat went under. By contrast, Gray's smartphone continued to connect to email servers at his employer for a few hours after he was last in verbal contact with family and friends.

We do not know whether Gray's cellular telephony provider had information useful for constraining the possible location of *Tenacious*: they were unable or unwilling to share the required information with private parties, likely because of legal guidelines about customer privacy. In a late phase of the search, the volunteer search team discovered that there are "mobile cell tower" devices that can be placed into planes and flown over a region to listen for signals from cellular handsets. Apparently such devices were used in the aftermath of Hurricane Katrina. Extraordinary efforts were made to get one of these devices deployed over the region where *Tenacious* had disappeared, but the devices are apparently quite rare, and the amateur team did not succeed in getting access to one in time, despite the involvement of highly influential people. Cellular telephony was one of the only technology industries where even Jim Gray's circle of colleagues was unable to pull sufficient strings to make things happen in time.

A.2 Discussion

EPIRBs and similar technology can make it easy to search for missing boats; where they often fail is in proper deployment. This raises a natural policy direction, which was one of the clearest lessons that came out of the 2008 Berkeley meeting of the volunteer search team: mandate the installation of maritime safety technology in a failsafe way, as we have done with other technologies like automobile airbags. It is both possible and inexpensive (relative to the cost of a boat) to require EPIRB-like technology to be integrated into boat construction. Better solutions could also be devised to enforce effective EPIRB deployment for pre-existing boats, and to enforce the use of PLBs in small planes, backcountry hiking registration and related settings. In the US in particular there are objections in some quarters to mandatory tracking technologies, focusing on individual liberties and electronic privacy issues, so this issue will need significant advocacy. It is, however, the most cost-effective approach to improve marine safety by a large margin.

Another major frustration during the search for *Tenacious* was that Gray's cellphone signal was available to his cellular telephone provider, but that information was not available to the volunteer search team. Now, just a few years after Gray disappeared, there are many services for sharing or acquiring the geolocation of GPS-enabled smartphones (location sharing services like Foursquare and Google Latitude, or on-demand phone-location services like Apple's "Find My Iphone" tool.) However, these are still relatively primitive in their access control policies, which would discourage reliable use. We return to this issue in Appendix D.3.

B. FURTHER DETAIL ON IMAGERY

The satellite data we received from Digital Globe had already been partially processed into visual form, i.e., it was not raw sensor data. However, it was processed by a pipeline tuned for land-based imagery, and was not easily usable by humans due to the lack of contrast in marine imagery. Thus, the pre-processing team applied processing steps to "stretch" the contrast within the pixels across a wider range of color values.

The imagery also had to be divided up into manageable chunks. Each Digital Globe collection was delivered to us as a longitudinal strip (sometimes divided into sub-strips) that was divided into 8k$\times$8K tiles. Each tile was, in turn, divided into either 256 subtiles of 512$\times$512 pixels (for processing by experienced users as described in Section 5.3.2) or a larger number of 300$\times$300 pixel sub-tiles (for crowdsourced processing using Mechanical Turk as in Section 5.3.1). These collections were delivered as either 0.82 m/pixel panchromatic images or 0.6 m/pixel pan-sharpened natural color images (the latter only after a 24 hour hold). The pixel depth was either 8 or 16 bits/pixel, but the useful depth was typically only a few bits per pixel due to the low contrast of over-the-water imagery (commercial satelite image processing is typically optimized for land-based imagery).

The Nasa ER-2 collection was captured using a Cirrus Digital camera system and consisted of 34 tiles, of 4500$\times$4500 pixels (8-8-8 bit RGB), which included a 10% overlap between adjacent images. These were divided into sub-tiles of 300$\times$300 pixels for crowdsourced processing using Mechanical Turk. The nominal resolution was 2.2m, though the actual resolution was substantially better due to a difference between the aircraft's actual altitude and the nominal altitude cited for the sensor.

Ten days after Gray's disappearance, SPOT Image approached us with the possibility of our purchasing Envisat data of the California coast for relevant dates. This was not deemed worthwhile by that time.

C. FURTHER DISCUSSION

C.1 Image Acquisition

During our acquisition of imagery, delays arose as the result of three main factors: scheduling latencies, image processing overheads, and policy challenges. In addition, it was clear that more nimble techniques could be developed for remote sensing during emergencies. We briefly summarize these data acquisition issues, and then discuss some lessons and challenges for future efforts similar to ours.

C.1.1 Scheduling

As we noted, satellites and planes have specific hardware and workload constraints that introduce inherent latencies in "positioning" their sensors. Another factor is weather; as it happened, this was the biggest source of delay in searching for *Tenacious*, since foul weather grounded search planes for many days, and cloud cover obscured much of the satellite imagery. In addition, remote imaging hardware tends to get very high utilization, since the large fixed cost of the equipment is typically amortized across many tasks. In our case, scarcity of resources did not impede our progress substantially, thanks to generous donations of time and resources from organizations like Digital Globe and a number of volunteer flight crews. However, this is a case where democratization of our experience would require new approaches: even with the best of intentions, Digital Globe could not provide this level of effort on a regular basis.

Latency and workload management are familiar topics in computer science, and the natural solution is to enable more informed global scheduling to assign imaging assets to worthy tasks. Given that various entities own such assets, some kind of clearinghouse—perhaps akin to travel reservation services—would be useful to provide better access to remote imaging equipment. Realistically, the market for this service may currently be too small to motivate participation by satellite providers. A crowdsourced bootstrapping effort might be more fruitful, as we outline in Appendix D.

C.1.2 Image Processing Overhead

The second source of delay—image processing—was a software issue that could be easily improved, if not substantively avoided. As we discuss in Appendix B, satellite image providers typically perform significant image processing before they release data. In our case, this had multiple negative impacts. First, computational bottlenecks at the provider introduced delays that could have been avoided by using the much larger computational facilities available to the volunteer team. In the last three years, cloud computing services have made access to such resources much easier, so custom image processing should be an increasingly affordable service decoupled from data acquisition. Second, the provider's effort to convert the data to human readable imagery was tuned to land imagery, so the volunteer team had to reprocess the images to improve the contrast within the color range applicable to "at sea" images. Finally, much of the underlying contrast range had been lost in the initial conversion at the provider; access to the "raw bits" that would have facilitated machine vision processing were lost in the provider's image processing. This was clearly a case where a very well-intentioned effort to make things easier for the volunteer team went astray due to the lack of a prior working relationship amongst the parties, i.e., the provider and the volunteers lacked knowledge of each other's working assumptions and capabilities. This would be alleviated by the availability of widely-known open-source software.

C.1.3 More Sensing

As Figure 5 demonstrates, the coverage of the remote imagery acquired by the volunteer team did not compare in scope to the overflight area covered by the Coast Guard. This fact only became clear to the amateur team towards the very end of the process, and was both heartening and humbling: we were happy to learn more about the effectiveness of the Coast Guard, even as we realized that our high tech effort covered a smaller area. On the other hand, the Coast Guard was encouraging of our efforts to leverage remote sensing data, and there is general agreement that there are lessons to be learned from the experience.

One obvious challenge raised by our effort is to enable more—and more targeted—image capture by making more hardware available. Planes are cheaper, nimbler and more plentiful than satellites, and they can often fly below obscuring cloud cover, so they seem like a promising avenue for improving this picture. Of course the Coast Guard maintains planes and crews for maritime SAR. So as a matter of national policy, one desirable outcome would be to better equip the national air fleet for SAR. This may require lobbying on behalf of USCG or other agencies.

Regardless of the capabilities of the government, it is hard for private citizens to get access to government-operated remote imaging assets. Most such assets are managed by the military, and even in a high profile case like this one it was not possible to get direct civilian contact with the military on this topic[2]. Meanwhile, we expect that many of the military's key assets are not traditionally pointed at domestic targets, nor at empty ocean, so in the *Tenacious* case the military equipment may not have been easy to utilize. Some government imaging assets are managed by non-strategic agencies like NASA and the National Guard. In these cases, it is possible to take advantage of the assets if the costs are sufficiently low. For example, we "got lucky" in our effort to access an ER-2 due to the training mission mentioned in Section 4. It would be nice to see this kind of opportunity systematized for the public good in a less ad hoc manner.

Another option to consider is the use of unmanned aerial vehicles (UAVs) including planes, helicopters, and blimps, which may be cheaper and safer to operate at a large scale than human-piloted planes. This seems particularly useful for fine-grained sensing in time and space during widespread disaster response. This raises challenges in robotics and computer vision, in addition to the scheduling and policy issues mentioned above. According to contacts at the USCG, the use of UAVs near shore in densely populated areas is highly problematic due to air traffic, but there are other settings (e.g., in the Pacific Islands) where they could be very useful both for SAR and for other USCG tasks like fisheries enforcement.

Independent of whether the sensors are carried on UAVs or a larger number of smaller aircraft, there is the opportunity to decouple the sensors themselves from the processing of the sensor-derived information as discussed in Section 5.7.2.

C.2 Organization of the Analysis Teams

While the entire volunteer search effort was a case study in rapid organizational development, the technical efforts employed by the imagery and drift modeling teams were somewhat more structured due to the technical nature of their tasks. The team structure emerged relatively organically bottom up, via domain expertise. In particular, Jim Gray's Astronomy colleagues had expertise, equipment and software experience with bulk image processing and organized themselves; Gray's Oceanography colleagues had expertise and software experience in drift modeling and again formed their own organization. Other parties that were contacted to help exhibited extreme generosity. The staff at Digital Globe and Amazon Mechanical Turk were remarkably giving of their time and resources, and trusting of the volunteers' efforts.[3].

The teams were very loosely guided by the analysis coordinator, who undertook a number of tasks. He coordinated the flow of information between teams (e.g., satelite imagery provider, preprocessing team at JHU, analysis teams at JHU and elsewhere, georeferencing and drift modeling), provided a common interface to the individuals supporting "higher layer" functions of the search

[2]In the search for *Tenacious*, efforts were made to directly connect the relevant Coast Guard officers with military contacts known to the volunteer team. It remains unknown whether those contacts occurred, and if so whether they affected the case.

[3]We stress that this list is not exhaustive. Many other groups offered help, but were not called upon to the same degree. And many individuals at institutions not listed here were critical to the effort.

described earlier (e.g., organization of search aircraft, management of the public web site, communication with the Coast Guard, etc.), and attempted to anticipate future needs, e.g., the recruiting of experts with naval imagery experience who could provide a final review of the results. Due to urgency of the crisis and the loose organization, the search coordinator used a relatively "hands-off", non-prescriptive approach to team management and tasking. This seemed to work well, in part due to a diversity and excess of human capital. When a volunteer team announced to the coordinator that they were pursuing a particular task, there was little reason to discourage their efforts—many such activities were afoot in the early days, and the ones described here gelled from a subset of those. A key example of this was the image geocoding process at Texas. Although the relative value of that effort was not clear to the coordinator at first, it quickly became essential to coordinating all the imaging efforts, and was very helpful in bringing new volunteers up to speed.

C.3 A Note on Uncertainty

Much of the data in the search process was noisy or uncertain. In some cases this was due to properties of sensing equipment. In others it was due to the use of statistical methods like drift modeling to predict or bound unknown information. And as is to be expected, there were also cases of human error in managing and interpreting the available data – some of which were caught and rectified, some of which were managed as part of the design (e.g., the redundancy in crowdsourced image analysis), and very likely some of which were unaccounted for. In retrospect it seems natural that the inputs and processes in a volunteer crisis response setting will have a high degree of noise, which should be modeled and mitigated as much as possible.

D. TECHNICAL CHALLENGES

D.1 Social Computing

Three specific technical challenges distill many of the issues that arose for us in communication and coordination. Solutions to these would have broad applicability in SAR and emergency response, and in general use as well.

- *Ad-hoc integration of communications software*. Develop tools that break down barriers between popular software packages for email, mailing lists, bulletin boards, and blogs, to provide features useful for crisis management and group coordination. Desired features include search, entity extraction and resolution (people, organizations, tasks), timeline and information flow analyses, and social network features like centrality and reputation assessment. Ideally it should be possible to quickly integrate multiple endpoints as well—e.g., multiple remote email inboxes—with contextual controls over privacy for each endpoint.

- *Bootstrapping task brokers from communications.* Provide a rich "swap board" website to match tasks and volunteers in an evolving crisis. In the absence of a pre-defined lexicon of tasks and expertise, matching must be done from contextual information, e.g., text in communications and external data sources, or communication patterns in a social network history with other team members. Even shallow analysis of text can help assess general topics, the level of technical discourse, bona fides mentioned by volunteers, etc. The dynamics of this information over time is also important, as roles and efforts shift—sometimes quickly.

- *Visualization and search in communications:* Develop visualization and search interfaces appropriate to crisis management. Standard interfaces for visualization and search in communications media (email, blogs, mailing lists, etc.) are a poor match for the activities described above. Rather than focusing on folders, tags, and message threads, group coordination calls for a focus on principals (people, organizations, roles), topics and events. In the Gray SAR search, common query patterns included "mnemonic expansions" like "who was that person who sent the email with all the complex satellite terminology?", aggregate analyses of messages, e.g., "who is at the center of the discussion about putting up posters?", temporal analyses like "when was the last message discussing red dye?", reputation analyses like "has this person demonstrated expertise on this topic previously" and so on. Interfaces need to be very efficient: in our experience, the effort to construct longer sequences of low-level queries was stymied by the latency of interaction with modern mail and web tools (often painful seconds per click!).

D.2 Image Acquisition and Analysis

The specific technical challenges we encountered in image acquisition and analysis can be combined into a larger-scale challenge: the development of a publicly-available system of similar capabilities, targeted for use in both personal and large-scale crises by citizens of modest means. From a software engineering perspective, it would be helpful to compare a patiently architected networked SAR infrastructure like this to the rapidly-developed polytechture used to search for *Tenacious*.

In full, the goal is to develop a cost-effective and reliable end-to-end solution that includes all the steps of our target qualification pipeline: image acquisition via a variety of manned and unmanned equipment; staging and processing of image streams; interfaces for human analysis, markup and scoring; dashboards to monitor workflow progress and outstanding tasks. To ensure rapid, cost-effective "spin-up" of the service in times of crisis, it should be deployed in a hosted environment, and be able to quickly scale up. In times when it is not needed, it should maintain an operational standby mode that regularly drives data through the entire pipeline, ensuring that all the components are in a constant state of readiness. This includes not only technical readiness, but also regular checks on potential organizational barriers like bureaucratic authorization to access relevant information and resources.

The San Francisco Marina from which *Tenacious* embarked is an example environment where a system of this sort could be bootstrapped. The marina is surrounded by affluent, tech-savvy homeowners with views of boat traffic, who would likely be willing to participate and contribute in a crowdsourced public safety effort. Such a community of volunteers could contribute to and host an array of motorized webcams monitoring the marina, keeping track of boats leaving and returning to their slips. Data feeds would likely be sent to a cloud storage utility, where they could be analyzed in batch or even in near-real time.

Some of the component research challenges in building such a system could include the following:

- *Crowdsourced Robotic Arrays of Webcams.* Foster communities of boaters and shoreline homeowners to provide Internet access to control motorized (pan/tilt/zoom) web cameras, akin to what is being done for ornithology [10]. Provide an online clearinghouse to register such cameras, and design mechanisms to schedule coverage of regions of interest for public safety applications.

- *Crowdsourced UAV Image Harvesting:* Develop a toolkit to enable robotic radio-controlled UAVs to participate in crowdsourced image harvesting, and investigate the practical and legal issues involved in keeping such devices in circulation.

- *Image Processing and Archiving (with Secure Proxies).* Develop a hosted image processing pipeline and corresponding data staging/archiving service for crowdsourced public safety imagery. As an additional challenge, in order to facilitate processing prior to legal release of imagery, devise provably secure access controls that can accommodate various release policies.
- *Image recognition for Maritime SAR.* Develop high-precision algorithms for recognizing boats and their activities in remote imagery. Multiple imaging technologies should be considered, including redundant coverage of the same targets with different technologies. Image recognition techniques can also inform imaging technology, and the scheduling of disparate imaging resources for particular tasks. This specific challenge is sufficiently focused that it might benefit from a competition akin to KDD Cup [5].

D.3 Prevention

Early notification is the best way to improve chances of rescue and recovery. Although cell providers are not allowed to divulge geolocation of phones, the widespread use of Internet-enabled devices with GPS should make geolocation and tracking relatively easy to deploy on a large scale. Unfortunately this is not well-supported by standard practice today, for various socio-technical reasons. On-demand location tools like “Find My iPhone” are not well suited to many SAR situations, since the device must be functioning when the location request is submitted. Push-based location-publishing services like Foursquare provide relatively coarse controls over publishing. Recent statistics suggest that these services are currently used by a narrow demographic group of young males [7], perhaps because other users are unwilling to configure or trust mechanisms to control publishing. Tools and policies are needed to make these services dependable in emergencies, while also providing privacy controls that are simple and attractive to the majority of users.

- *Personal tracking services with contextual access control.* Develop a web service and corresponding mobile applications that provide safety-oriented geolocation tracking. Provide acceptable contextual publishing policies and enforcement mechanisms that provide failsafe activation of tracking in risky situations, and enable contextually-appropriate release of information during emergencies.

Figure 5: Map of search areas, including Coast Guard overflights, satellite imagery, and privately-funded searches. Missing from this map are private searches along the coast North and South of San Francisco.