Brief Bio and (PR)²: Problems & Pitches – Rants & Raves by Daniel Halsey

I've been developing software for nearly as long as I can remember, and have been doing so professionally for well over a decade now. I've worked almost exclusively for universities, and am proud that the software I've produced and maintained has supported educational and scientific pursuits.

Before moving up to Indiana, I worked for IFAS/CALS at the University of Florida, over a span of 10 years and in a wide range of IT capacities, culminating with my favorite job title so far: "IT Expert". My systems there supported on-campus academic units, the state Extension Service, and the USDA's Sustainable Agriculture Research and Education program.

Indiana became my home in 2008 and I started working at the Kelley School of Business at Indiana University shortly thereafter. My work there included developing and supporting a variety of knowledge delivery tools, classroom support apps, and line-of-business systems for both their online and on-campus programs.

I joined CNS in February of 2014 and I'm looking forward to helping them provide humanity with insights into what it knows. I believe in the advancement of knowledge, and am glad to work with a team organized around the goal of making knowledge accessible.

I’m an old-school computer geek.

http://cns.iu.edu/current_team/bio/daniel_halsey.html

General Questions

1. **Do you consider yourself a developer, user, creator, system evangelist, etc.?**
   I am primarily a developer, with a focus on system architecture and long-term software sustainability.

2. **What are your main interests in attending the workshop?**
   I’m interested in hearing how other groups might be able to use our tools as part of their workflows and tool chains. I’m also very interested in learning more about system architectures other groups are using and how other groups determine the direction their architecture, technologies, techniques, and practices will take.

3. **What would you like to learn / achieve at the workshop?**
   I’d like to learn more about tools that can work in conjunction with ours to build systems that produce visualizations and analyses that provide insight into complex data systems.

4. **What are the tools or services you would like to share at the workshop?**
   I’m eager to share our recent work to standardize our online visualization process and our plans for our systems and frameworks as we move ahead.

5. **Please list three features or functions of your tools or services that are most important for your users.**
   - Interoperability with other systems
   - Unique analyses and visualizations
   - Standards-based and standard-creating open source systems

6. **Software is often a non-static product—libraries and frameworks are evolving continuously. What techniques, practices, and patterns have your team found work best to ensure you're both fulfilling user needs and prepared for the future?**
Our frameworks and systems (like many others) are in constant threat of becoming out-of-date, as underlying Java frameworks get retired, libraries change, and software practices evolve. Fortunately, our architecture allows us to update components individually, and is built specifically for extensibility, so adding functionality is a matter of writing new plugins, rather than continuously refactoring existing code.

7. **What are some underserved user needs that your systems, idealized versions, or an ecosystem that your systems are integral components of could address?**
   
   Our users need and demand a better workflow system in our primary distribution. The current one works…technically, but is not at all easy or straightforward to use.

   Our systems are not always ideally suited to consumption, analysis, and visualization of “big” or real-time/streaming data (though we have produced stream-processing variants of CIShell). Part of this is intentional, and a result of our decision to integrate into analysis and visualization ecosystems, and to provide mid-analysis data sets.

   An ideal version of Sci2 would be able to integrate without as many manual seams into a big data ecosystem, providing, for instance, Hadoop/Storm plugins and WebSockets-based push-updated visualizations.

8. **Are you or your group working on any of these challenges? If yes, please explain.**

   Yes. We're starting work on our new workflow engine and are planning out our preferred methods of big/real-time data analysis.

9. **“Big data” and “cloud computing” get thrown around a lot as terms. How do these concepts and your, your group’s and your users’ interpretation and understanding of them affect your plans for development?**

   Our tools are intended to be used as part of larger analysis and visualization tool chains, rather than end-to-end solutions. The implication of this is that we tend to rely on other tools and systems to feed pre-aggregated, filtered, or partially analyzed sets of truly large data. Likewise, our tools generally operate as single links in this tool chain, and have some parallelizable components which might be suitable for cloud environments, but are largely serial processes by intent. Rather than single parallelized instances providing services to multiple clients, multiple just-in-time instances of our software, providing individual execution sessions for each client, might be a cloud computing route we would want to explore.