

A Study of Clinical Relevant Open Source Systems (The CROSS Study)

Michael Hogarth, MD and Stuart Turner, DVM, MS
UC Davis School of Medicine



The Emergence of Open Source

- What is open source?
 - Not “free” in cost, but “free” for modification
 - Software source code is “open” – license allows modification and unrestricted redistribution
 - Defined by a license, but *more importantly* it involves an open development **process**
 - Benefits of an open source development process:
 - High quality – many ‘eye balls’ looking at the source code
 - Low development cost – cost (time, funds) is spread across many individuals or entities
 - No limitations on distribution/redistribution by commercial entities



Why study clinical open source?

- EMR Market Problems:
 - Est. \$30K-\$50K per physician to implement a computer based patient record system (CBPR)
 - Vendors are highly proprietary in the clinical IT domain
- It is now being touted as “a solution to the EMR cost barrier”
 - American Acad. of Family Practice – open source EMR project
 - VA VISTA – Veteran’s Admin system is now open source
 - NCICB software philosophy is to provide it as open source
 - Will the recent DHS Healthcare IT initiative invest in OSS?
- Metrics are needed to help users make appropriate business decisions when choosing open source software



Specific Aims

- Aim 1: Catalog clinical open source projects and analyze their technical and development characteristics
- Aim 2: Understand clinical OSS developer community
 - Motivations of contributors and users
 - Business models
- Aim 3: Answer the question – “will the open source model work in the clinical software domain?”
 - This is currently assumed, but no research on this yet...
 - Why this assumption could be wrong:
 - Very different marketplace (very proprietary market)
 - Most developing OSS in this domain are not professional engineers/programmers
 - Clinical software requires significant domain knowledge by engineers
 - Is there a funding model?



Methods

- “Find” open source projects in the clinical domain (“mining for open source projects”)
- Analysis of characteristics
 - System - programming language, platform
 - Developers – motivations, attitudes, skill sets
 - Application – type of application, is it in use today?
- Formulate “open source” metrics:
 - Sustainability of the project (“vitality”)
 - Quality metrics for open source projects
 - Metrics for the process
 - Metrics for the software



Data Collection

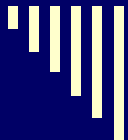
- Inclusion criteria
 - Software had to exist (no conceptual projects)
 - License is Open Source Initiative (OSI) compliant
 - Software had to be “clinically relevant”
 - Defined as: “has application to direct clinical care or common processes in clinical care”
 - Excluded: bioinformatics software, genetic sequencing software, etc..

- “Mining” methods
 - Sourceforge: “medical”, “clinical”, “informatics”, “bioinformatics”
 - Spirit Project database
 - LinuxMedNews project list
 - Google: “open source” and “medicine”, “informatics”

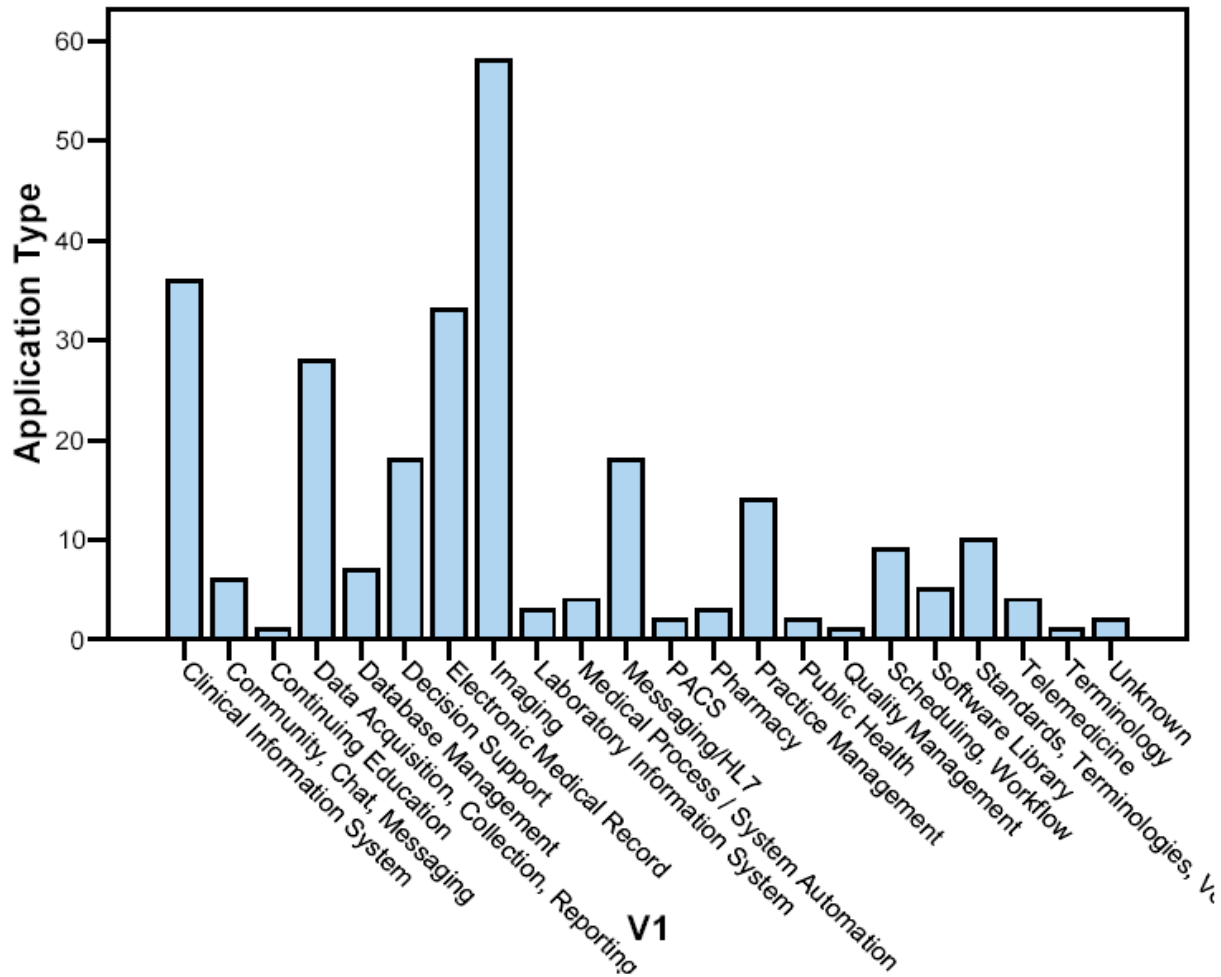


Preliminary Results

- April 2003: 143 projects
- April 2004: 264 projects
 - Only 10% are listed as “stable” (not alpha, beta, early release, etc..)
 - Very low ‘vitality’ metric for the majority of the projects (most are ‘dead projects’)
 - “Vitality” score does reflect the status of the project
 - The source code for the “major” projects was difficult to find
 - 40% of the projects have unknown or undefined licenses
 - 50% are platform independent software
 - Java is the most prevalent language



Application Type





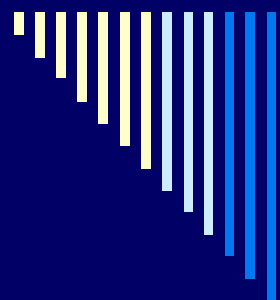
Other observations

- ❑ Very heterogeneous software development practices
- ❑ Poor record of project completion
- ❑ Developers:
 - Most developers are not engineers but technology aware healthcare professionals
 - Fundamentally different than non-clinical domain
- ❑ Very low penetration of corporate involvement/contribution
- ❑ A significant number of projects are from outside the U.S.
 - Europe (France, Germany) – strong OSS clinical community



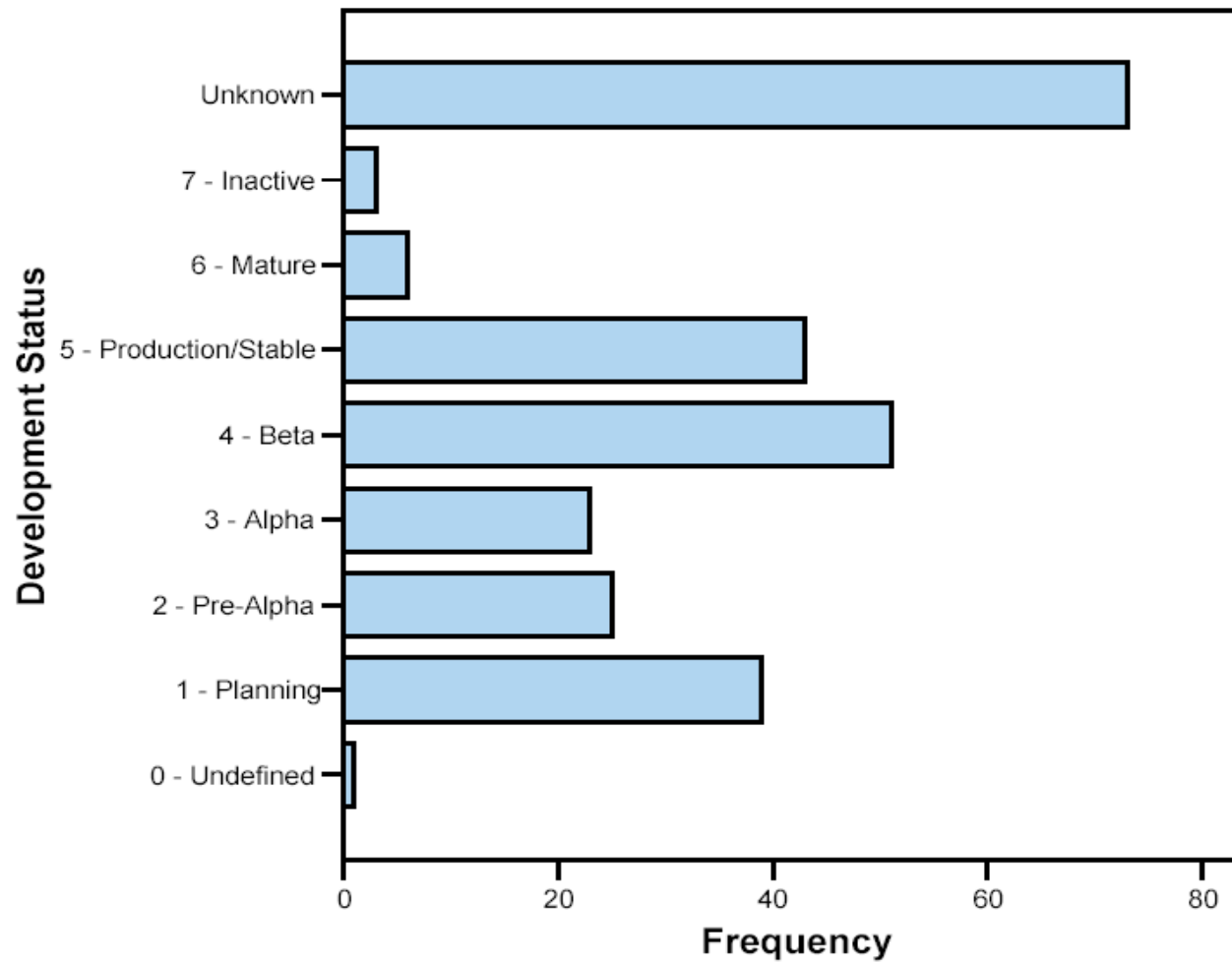
Future work

- Refine metrics for open source software and projects
 - “vitality score”
 - Software quality metrics for OSS
 - What are signs of trouble for the project?
- Understand the contributor community
 - Skill set
 - Motivations
 - Organization
 - Funding
- Can open source in the clinical domain parallel that in the non-clinical domain?

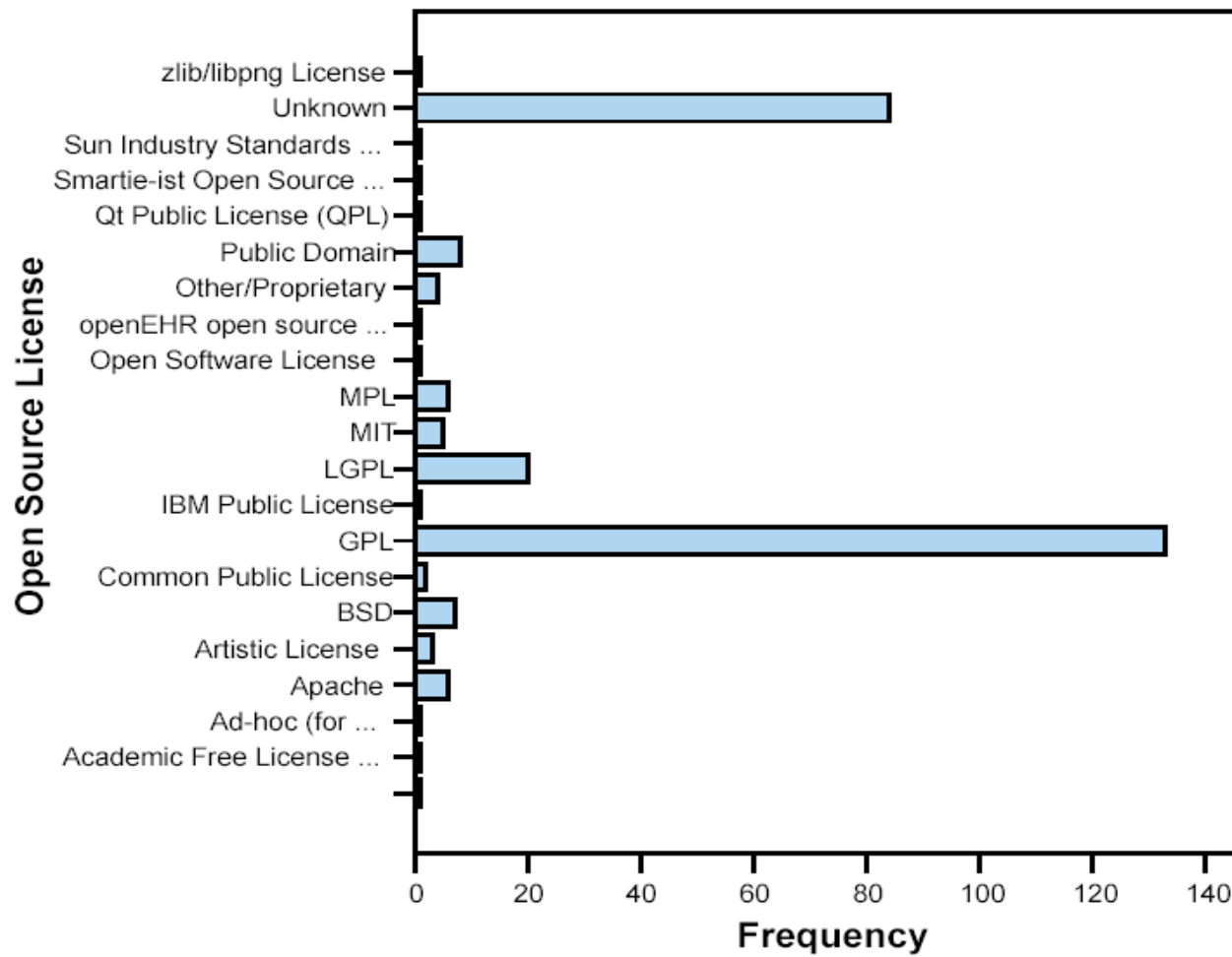


END

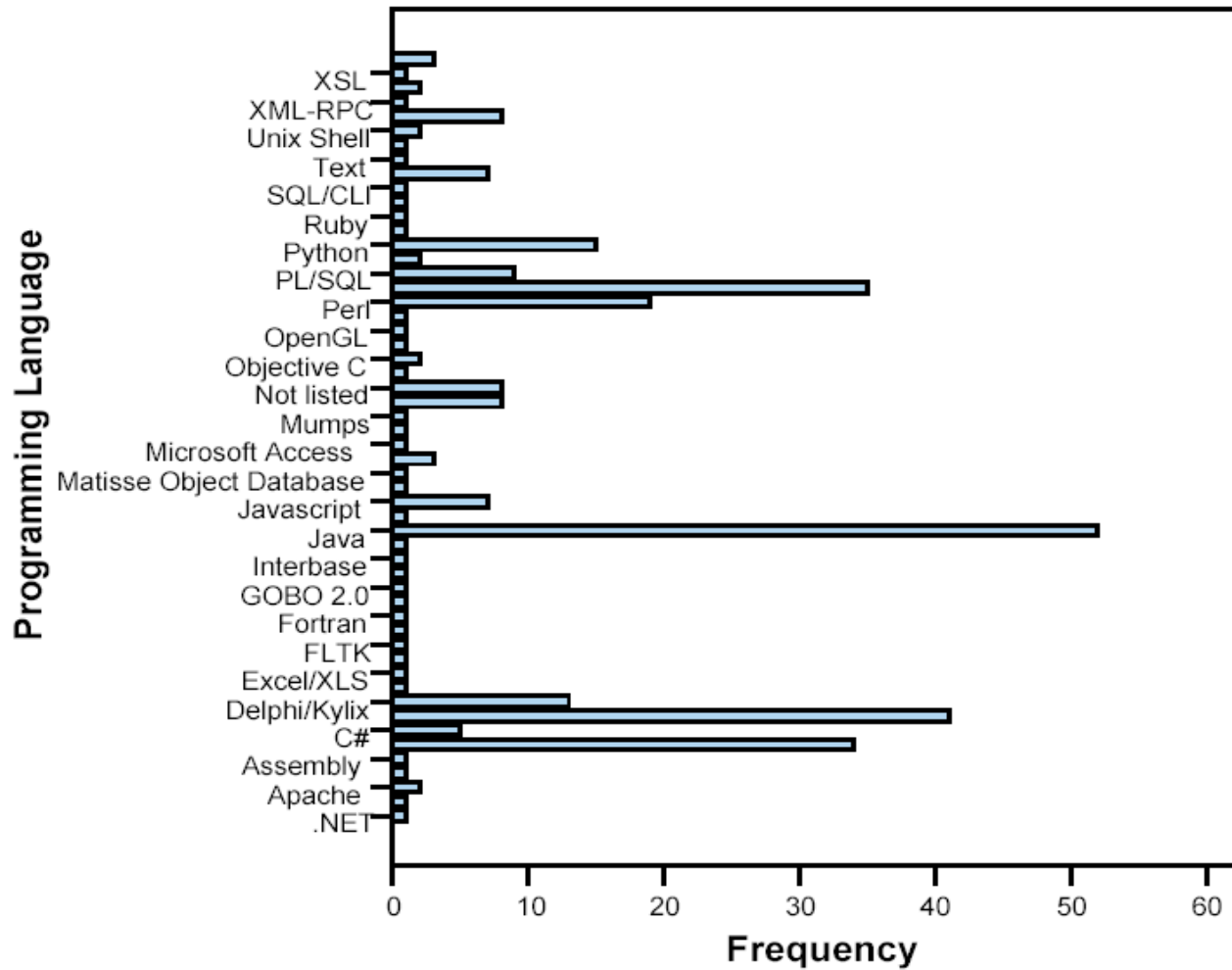
Project Development Status



Open Source License



Programming Language



Operating System

