OS^{3}A and the Potential for Open Collaboration in Clinical Research

Ann Marie Martin, Dante diTommaso,
Paul OldenKamp, Paul D Hamilton.

The Future is in Open Collaboration

2 April 2009 1

1977 - 1996

- WHO IN THIS AUDIENCE IS BORN:
  - BEFORE 1977?
  - AFTER 1977?
- IF AFTER 1977 YOU GREW UP WITH THE INTERNET

The Future is in Open Collaboration

OS SAS® APPs

2 April 2009

2

Outline

- Open Collaboration & FOSS
- **OS^{3}A: O**pen **S**ource **S**AS® **S**oftware **A**pps
- Framework For Achieving an Industry-Driven Open Source Clinical Reporting System
- Good Programming Practice for Clinical Trials
- What's in an Open Collaboration?

The Future is in Open Collaboration

Open Collaboration & FOSS

- Wikipedia, YouTube, Facebook, LinkedIn, sasCommunity.org, Redscope, Orphanet
- Innocentive, GoldCorp
- LuLu, PLos, PLosMedicine, OpenCourseWare
- Linux, GNU, SubVersioN, CVS, SpikeSource, FUTS
- FireBird, OpenCDISC, OpenClinica, CRIX, caBIG®, CAISIS, openCDMS
- ClinicalTrials.gov, ClinicalStudyResults.org
- Tylers' Hope for a Dystonia Cure and others

The Future is in Open Collaboration

OS SAS® APPs 2 April 2009 4

•Redscope web2.0 site of the German SAS programming communiity posts 12 libraries of SAS programs/macros ranging from utilities to programs suitable for genome analysis. Propensity score based on logistic regression

•sasCommunity.org wiki for the SAS community includes draft of the Good Programming Practice for Clinical Trials

•Orphanet web2.0 application with information on orphan drug diseases, including info on patient support groups

•Innocentive web2.0 application for posting scientific projects, award winning project catches the prize

•GoldCorp the industry example. Goldmine where internal geologists could not locate the gold, CEO posted the project and background material on the web and invited contributions for an award, scientists from other disciplines such as mathematicians helped to locate the gold.

•LuLu award winning web2.0 application for publishing own books

•PLos Public Library of Science

•PLosMedicine A peer reviewed Open Access Journal

•OpenCourseWare is an initiative of MIT to provide coursecontent on the internet

•Linux, GNU FOSS operating systems

•SVN SubVersioN widely used versioning software available both FOSS and commercially, x support groups, y references on Google

•CVS Concurrent Versioning System

•FUTS open source SAS unit test tool, FOSS application provided by Thotwave Technologies

•FireBird Federal Global investigator registry for commonly used/referenced data. Replaces 1572 form. Memo understanding FDA, NCI, CRIX International to develop the next generation FireBird.

•OpenCDISC FOSS testing of compliance versus SDTM, SDTM+, SDTM-, including WEDSDM, JANUS tests.

•OpenClinica Open Source Electronic Data Capture system created within 2 years of inception.

•CRIX Clinical Research International eXchange collaboration between various companies to provide common IT applications for facilitating clinical research

•caBIG® Cancer Biomedical Informatics Grid

•CAISIS Integrated Clinic and Research Management System

•openCDMS Clinical Data Management System

•ClinicalTrials.gov obligatory publishing of clinical trials to be conducted

•Clinicaltrials results.gov obligatory posting of the results for marketed products Maine law and FDA. FDAA act

•Patient advocacy groups

Open Source: The Web Changes Everything!

- Open Source technology and policy (Deek & McHugh)
- Wealth of networks (Benkler)
- Wikinomics (Tapscott)
- Open Source paradigm shift (O'Reilly)
- Intellectual Property and Open Source (Van Lindberg)

The Future is in Open Collaboration

OS SAS® APPs 2 April 2009 5

Open Source technology and policy (Deek & McHugh)

•WW attempt @realizing a vision: keep scientific advances created by software development openly available for everyone to understand and improve upon

•The peer review process is even more open than that of traditional science

Wealth of networks (Benkler)

•Networked economy opportunity for anyone who values human welfare, development and freedom

Wikinomics (Tapscott)

Web 2.0 turning into a business trend

1977-1996 largest generation now already economically active that grew up with the internet, bigger than the baby boom

peer production communities

•elements of hierarchy and self organization

•you do benefit from voluntary participation through development opportunities and making a name

•are more efficient than firms or markets at allocating time and attention for certain tasks

Open Source paradigm shift (O'Reilly)

•Software commoditization: driven by standards, large drop in market revenue but greater social value. Software as a service (SpikeSource, Thotwave)

•Network enables collaboration: recognize role of Usenet in establishing the internet standard

•Software customizability: architecture is key to allow adaptability

Intellectual Property and Open Source (Van Lindberg)

•OS Legal construct for cooperation and trade in IP

•In cooperative models the owners and consumers are the same conf cooperative and private banks

•Cooperative leads to different incentives for involvement and investment

•Participants receive full benefit of the software product while only paying a fraction of the cost

•Many classes of software are becoming commodity goods e.g. antivirus and firewall

•Lowers the cost of contributions by an individual in comparison with the value of the product produced

Open Source: Licenses as Economic Solution

- Requirement for a license to be approved by the OpenSource Initiative as a OS license:
  - Provide incentives for collaboration
  - Increase cost of defecting
- Reciprocal: GNU General Public [GPL]
- Non-reciprocal: Eclipse Public
- GNU Library or Lesser GPL [LPLG]
- GNU Free Documentation License [FDL]

The Future is in Open Collaboration

OS SAS® APPs 2 April 2009 6

License to be approved by the Open Source Initiative impose conditions to increase the cost of defecting (taking the software over and claiming ownership is not possible) and provides incentives for collaboration

Reciprocal: GNU General Public [GPL]

create derivative work then distribute under license with same rights

Non-reciprocal: Eclipse Public

encourages adoption of code in proprietary systems

GNU Library or Lesser GPL [LPLG]

allows call of code from proprietary <u>and</u> non-proprietary systems

for SAS macro libraries

Open Source: FOSS - a How-to

- Technology
- Social and political infrastructure
- Money
- Communications
- Daily development, packaging and releasing
- Managing volunteers
- Licenses, copyrights and patents

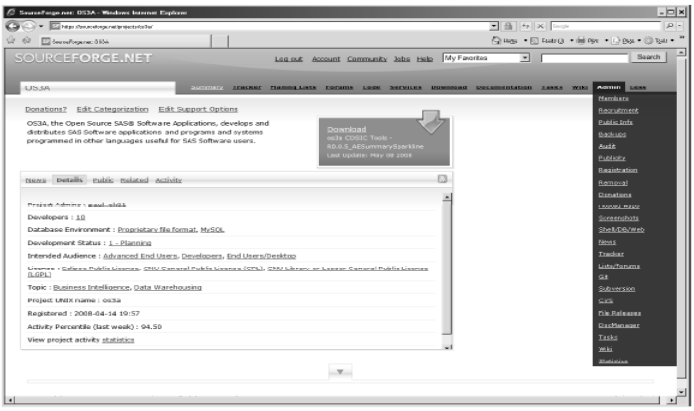
The Future is in Open Collaboration

OS SAS® APPs 2 April 2009 7

Technology:
•Website
•wiki
•mailing lists
•version control (code, documentation etc)
•bug tracking (CVS Concurrent Version Systems or SubVersioN SVN)
•real-time chat
Social and political infrastructure:
•flexibility and cooperation among the project team members to try to avoid “FORKS”
•fork is ability of anyone at any time to take all the work to date copy it to a competing project, and go off in a new direction
•Not a place to play celebrity death match!
Money
•Project members voluntary or paying day job
•Share the burden
•Augmenting services Thotwave, SpikeSource
•Support sales of hardware
•Competition with other companies, get market entry
•Marketing increasing brand value
•Dual licensing OS can be free but also included in proprietary soft
•Donations: transparency is key
Communications
•high quality communication speaks volumes and produces great leverage
•close attention to writing style, readability of program code, emotional tone
Daily development, packaging and releasing
•Same SDLC, same major minor release cycle including validation
•multiple threads in parallel, release branches
Managing volunteers
•what really motivates programmers
Licenses copyrights and patents

OS^{3}A **O**pen **S**ource **S**AS® **S**oftware **A**pps: Sourceforge.net Development Env.

OS^{3}A @ http://sourceforge.net/projects/os3a/



The Future is in Open Collaboration

OS SAS® APPs

2 April 2009

8

OS^{3}A Project

- 4 license options (see above)
- 1 windows SAS license provided by SAS Institute
- Individual & corporate contributors sign license to contribute to the project designed to protect contributors and OS3A
- Success factor is an attractive **initial objective**
- Potential projects
  - innovative statistical methods
  - graphic display methods & templates
  - ➡ clinical reporting system
  - ... *ideas from the community!*

The Future is in Open Collaboration

OS SAS® APPs

2 April 2009

9

•Code organization: a way to locate all SAS Code available on the internet.

•Metadata application programming: mapping to e.g. CDISC

•Code Design Standards: use the project to demonstrate how to use the consensus GPP or write a program to measure compliance with the standard.

•ODS templates and Methods: share ‘standard’ and sophisticated templates.

•SAS/IML Studio Applications: library of innovative statistical methods and reporting examples.

•Clinical reporting: develop a clinical trials reporting system to save cost and effort in reporting of clinical trials.

OS^{3}A Activities

- Weekly TC's:
  - Administrator Paul OldenKamp, Seattle Children's Research Institute
  - Paul D. Hamilton, Amgen
  - Dante diTommaso, Roche
  - Ann Marie Martin, UCB Pharma
- Presentations and Discussion forums PharmaSUG, PhUSE
- Support from PhUSE!!!

The Future is in Open Collaboration

OS SAS® APPs

2 April 2009

10

Framework for Achieving an Open-Source Clinical Reporting System

- Every conference…
- Open-source library build on standards
  - MSSO MedDRA, medical terms
  - WHO ATC, chemical names
  - SI Units, laboratory measures
  - CDISC, data model
- Modularized, validated system creating efficiencies and business value
- Meta-data driven (customizable) algorithms e.g. specific disease, lab abnormal values

The Future is in Open Collaboration

OS SAS® APPs 2 April 2009 11

Introduction

- Each year presentations on reporting systems, each with individual good ideas
- Traditional objective
- Effort repeated throughout the industry
- Remarkably similar redundant applications

Framework for Achieving an Open-Source Clinical Reporting System

- Rationale for participation
  - Industry, Regulatory, Programming community
  - ➡ Patients
- Framework
  - Technical framework & Project scope
  - Development methodology & standards
  - Testing and Validation
  - Support and Maintenance
- Staged development plan

The Future is in Open Collaboration

OS SAS® APPs

2 April 2009 12

Rationale for participation

Industry : opportunity to build an application based on emerging standards and share costs; enforce standards through software

Regulatory: Regulators want standard data displays to facilitate review

Programming community: Allows significant development opportunities for programmers

Framework

•Broad participation, support and sponsorship from the industry is desirable

Technical framework and project scope: technical challenges and decisions, by definition we cannot answer them here, must be collaborative

•What hardware and software version, avoid interactions with the directory tree?

•Stay with the old methods of listings and summaries or use graphics?

•Adoption of open source tools depends on establishing a central framework based on transparent and rigorous software development principles, ease of implementation

Development methodolgy & standards:

•Procedures for requirements, design, develop, test, validate, document, deliver, train and maintain

•Basic design considerations separation customizable business rules and fixed algorithms

Testing and validation

•Certain syntax and routines depend on the local environment (hardware and software)

Support and maintenance plan

•Some commitment of the industry for maintenance seems necessary

Staged development plan

•Lot to learn & new ways to work, therefore a staged plan

•Initially stand-alone utility macros, easy-to-read template programs for creation of analysis datasets, summaries and listings, to give experience with new tools and an automated testing framework

•Once comfortable with the basics, we could deliver more complex systems

Good Programming Practice for Clinical Trials

- Goal: consensus Guidance by Oct 2009
- Editorial Board, timelines
- PharmaSUG, PhUSE
- http://www.sascommunity.org/wiki/Good_Programming_Practice_for_Clinical_Trials
- 12,700+ views, watchlists
- Points to develop, need contributions

The Future is in Open Collaboration

OS SAS® APPs 2 April 2009 13

Points to Consider (see web):

•version of SAS®

•Defensive coding to avoid floating point error so disturbing when flagging for example laboratory values outside a certain range.

•Length of variable names and labels, character variables. Currently the format for submitting datasets to the regulatory authorities is still the v5 xpt format not allowing variable names with more than 8 characters etc..

•Rules to observe when making imputations.

•Is hard coding of information in a program permissible or not? If permissible should it be governed by an SOP?

•Validation of programs, how do we interpret the 21CFRPart 11 regulation in the context of clinical trials?

•What is the function of the SASLOG as opposed to the SAS code? A record of all the operations executed on the clinical trial data?

•Techniques for automating the define document required by the FDA.

•Do we need a section on the use of other statistical programming packages than SAS? Admittedly SAS is probably the most widely used programming language for clinical trial reporting however other software is used for specific purposes: advanced statistical techniques, sample size calculations etc...

Business Case (a theoretical attempt)

The Future is in Open Collaboration

	year 1	year 2	year 3	year 4	year 5	Assumptions
Cost develoment	x					
Cost maintenance	0	y	y	y	y	
Individual organization	x	y	y	y	y	x + 4y
Total cost over 10 organizations	**10x**	**10y**	**10y**	**10y**	**10y**	**10 (x+4y)**
Build and maintain consortium	z	z	z	z	z	5z
Cost develoment	x/10					assume 10 organizations
Cost maintenance		Y/10	Y/10	Y/10	y/10	1/10 * (x+4y)
Efficiencies through commoditization	0	5%	5%	5%	5%	5% each year
Total cost consortium	**x/10**	**Y/10*0.05**	**Y/10*0.05**	**Y/10*0.05**	**Y/10*0.05**	**1/10*0.05(x+4Y)+5Z**

OS SAS APPs

2 April 2009

14

Business Case (a theoretical attempt)

The Future is in Open Collaboration

	year 1	year 2	year 3	year 4	year 5	Assumptions
Cost develoment	100,000					x=40 business weeks or 100000€
Cost maintenance	0	10,000	10,000	10,000	10,000	y= 10% of X
Total cost per organization	**100,000**	**10,000**	**10,000**	**10,000**	**10,000**	**140,000**
Total cost over 10 organizations	**1,000,000**	**100,000**	**100,000**	**100,000**	**100,000**	**1,400,000**
Build and maintain consortium	20,000	20,000	20,000	20,000	20,000	z=20% of X
Cost develoment	100,000					x=40 business weeks or 100000€
Cost maintenance		10,000	10,000	10,000	10,000	y= 10% of X
Efficiencies through commoditization	0%	5%	5%	5%	5%	5% each year
Total cost consortium	**120,000**	**28,500**	**28,500**	**28,500**	**28,500**	**234,000 so 2X**
Total cost per organization	**12,000**	**2,850**	**2,850**	**2,850**	**2,850**	**23,400 so 1/5**

OS SAS® APPs

2 April 2009

15

So what's in an Open Collaboration?

- Economic rationale for open-source
- Licenses selection, A FOSS "How-to"
- OS^{3}A and suggestions of projects
  - utilities → modules → systems
- Role of a consensus GPP
- Why? We should and we can
  - Medical needs (e.g. generalized dystonia)
  - Industry standards, Technical foundation
  - Industry expertise in programming community

The Future is in Open Collaboration

OS SAS® APPs 2 April 2009 16

Many diseases with no cure. Patients are trying to raise awareness through eg YouTube:

•Epilepsy, 5930 results

•Dystonia 322 results

•Rheumatoid Arthritis 558 results

•Mycosis fungoides 50% of a malignancy called cutaneous T-cell lymphone 1 result

Useful links

- www.OpenCDisc.org
- www.sourceforge.net
- www.sascommunity.org/wiki/Good_Programming_Practice_for_Clinical_Trials
- http://sourceforge.net/projects/os3a/
- www.redscope.org
- www.lulu.com
- www.wikibooks.org
- www.CRIX.org
- www.clinicaltrials.gov
- www.clinicalstudyresults.org
- https://cabig.nci.nih.gov
- www.caisis.org
- www.openCDMS.org
- ….

The Future is in Open Collaboration

OS SAS® APPs

2 April 2009

17

References

- Developing a clinical reporting system collaboratively under an Open Source software license, Ann Marie Martin paper RA05, PhUSE 2008.
- Open Source SAS® Software Applications: the OS3A Program for Community programming, Paul OldenKamp et al. paper PO06, PharmaSUG 2009, to be published.
- A Framework For Achieving an Industry-Driven Open-Source Clinical Reporting System, Dante diTommaso et al. paper PO016, PharmaSUG 2009, to be published.
- Need for a Good Programming Practice for Clinical Trials, Ann Marie Martin et al., paper PO18, PharmaSUG 2009, to be published.
- For references Deek&McHugh, Benkler, O'Reilly, Van Lindberg see reference paper PO06 PharmaSUg 2009.

Glossary

- FOSS, FLOSS, F/OSS = Free (Libre) and Open Source Software
- GNU = GNU's Not Unix

The Future is in Open Collaboration