THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT
GREATER CHICAGO STUDY CENTER

Open Source Software

Laying the foundation

Warren A. Kibbe, PhD
Northwestern University Health and Biomedical Informatics
24 August 2011

Jane Holl, MD
Principal Investigator
Feinberg School of Medicine
Northwestern University

Greater Chicago Study Center

Clinical Research Informatics Challenges

- Make protocols computable
- Make software only as complex (or as simple) as it needs to be
- Change only what needs to change, but change it everywhere. DRY (Don't Repeat Yourself)
- Identify and enforce a single source of truth

Everything should be kept as simple as possible, but not simpler. Albert Einstein

About open source

- Open source is an approach to delivery software solutions based on community input, resources, and needs
- Done right, open source projects offer a lightweight framework for enabling flexible collaboration and support ‘microunits’ of contribution
- Open source does not mean lack of control. A control board typically oversees standards, training, review and hands out the ability to ‘commit’ (submit code and documents) to the project

About agile

- Agile is about finding the ‘right’ balance of communication, planning, change, and high quality deliverables
- Agile focuses on tight communication with domain experts and stakeholders
- Agile delivers production quality software in short iterations, typically two weeks
- Short iterations drive technology requirements for development, like version control, automated builds, automated testing, continuous integration

Supporting Research with Software
What is the problem?

- Protocols are exception driven and hard to make computable
- Flexibility in design usually makes complexity in operation
- Change is good, necessary, but hard to manage without computable representation

Providing a robust model for clinical research with the ‘right’ level of flexibility is difficult

Supporting Research with Software
What is the problem?

- Scalable change processes are hard to find
- How to couple stakeholders, review, management (control boards), communication so that iterations can be short and useful
- Data definitions are famously short lived

Agile and Open Source can address many of these issue. The NCS has strongly endorsed this approach.

NCS, Open source, and MDES

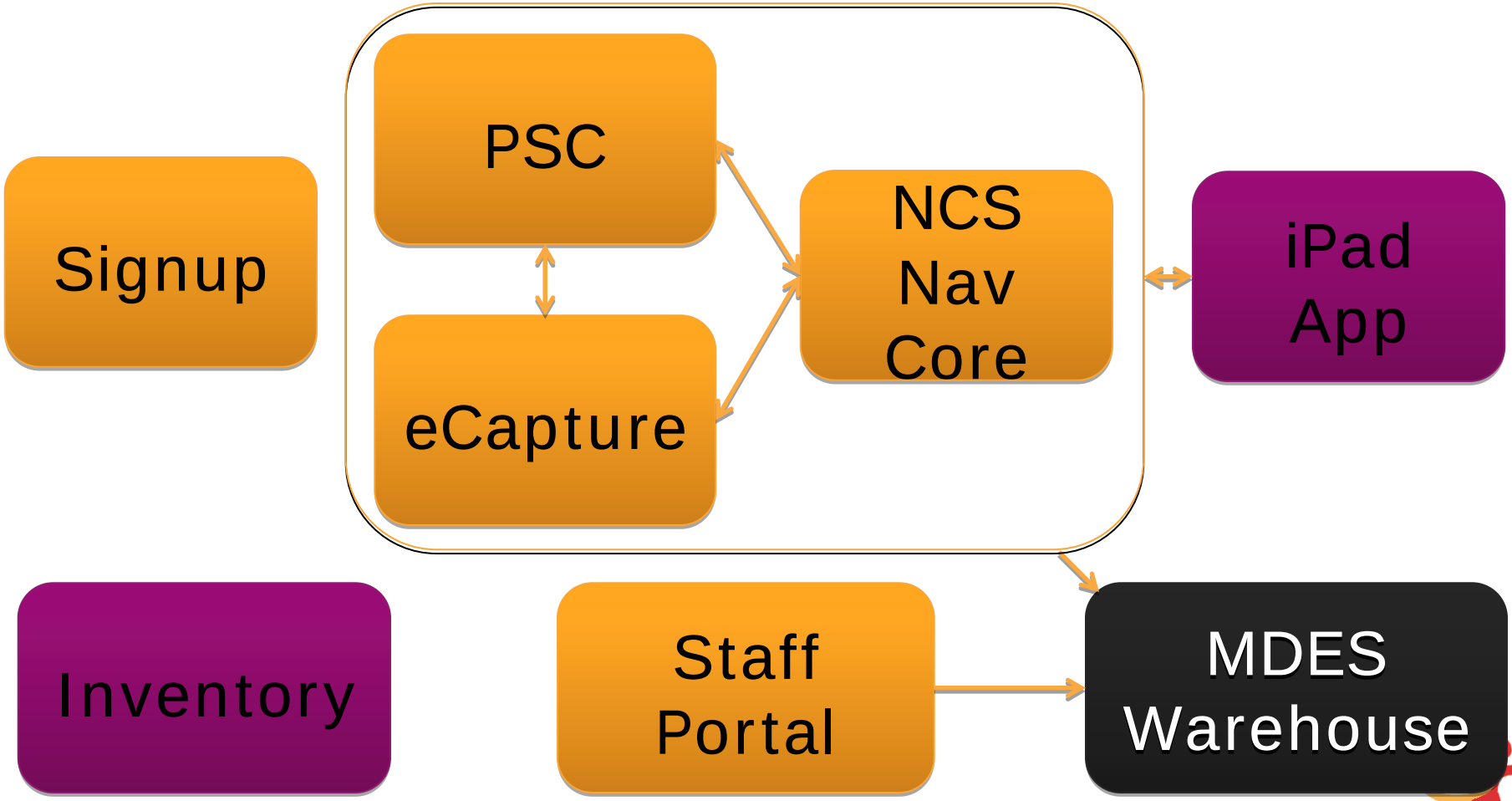
- The NCS has come out strongly in support of open standards, interoperable systems, and open source
- The drive to have all NCS data in open systems enables interoperability and simplifies software migration
- NCS uses “facilitated decentralization” with central data repository and multiple flexible case management and data acquisition platforms that comply with central specifications
- The NCS MDES (Master Data Element Specification) is an excellent step toward having a single computable source of truth for data definitions and rules for data collection

NCS Navigator

- An open source set of tools for managing the NCS study
- Uses modern software development practices
- The NCS Navigator Consortium comprises seven study sites located at seven institutions. The institutions are:

Northwestern University, Emory University, Johns Hopkins University, Maine Medical Center, University of Colorado Denver, University of Minnesota, University of Pittsburgh

NCS Navigator



NCS Navigator and the MDES

- A key component of the NCS Navigator is the NCS Navigator MDES Warehouse. MDES is the Master Data Element Specification. It is a computable representation of the data definitions, rules and collection instruments available in the MDES.
- The MDES Warehouse versions and accessions submissions based on the then current version of the MDES and migrates existing data to newer MDES schemas with as little human intervention as possible

Next Steps

- Provide an open source, collaborative framework for collection instrument life cycle management
- Provide an open source, collaborative framework for versioning and sharing changes to the data elements, rules and definitions in the MDES
- Provide tools for automating the creation of validated data collection instruments from the MDES definitions

The goal is on demand central data submissions, data exchange between sites, and 'future proofing' the data repository

THE NATIONAL CHILDREN'S STUDY

GREATER CHICAGO STUDY CENTER

HEALTH GROWTH ENVIRONMENT

Thank you!

wakibbe@northwestern.edu

Eunice Kennedy Shriver NICHD

National Institute of Child Health & Human Development