An Ultra-Large-Scale Systems Approach to National-Scale Health Information Systems

Kevin Sullivan
Computer Science
151 Engineer's Way
University of Virginia
Charlottesville, VA 22904 USA
sullivan@virginia.edu

William Knaus
Public Health Sciences
Health System West Complex
University of Virginia
Charlottesville VA 22904 USA
wak4b@virginia.edu

Richard Marks
Patient Command
6004 Balsam Drive
Maclean, VA, 22101 USA
richardmarks@earthlink.net

ABSTRACT

The U.S. urgently needs a major initiative to develop software and systems engineering foundations for a national-scale health information network (NHIN). The NHIN will be an *ultra-large-scale (ULS) system.* An *ULS systems perspective* therefore must guide these activities. There are enormous opportunities in this area for software engineering and computing systems research—to advance knowledge and improve our society.

Categories and Subject Descriptors

D.2.11 [**Software Engineering**]: Software Architectures

General Terms

Design, Economics

Keywords

Ultra-large-scale, national health information network

1. INTRODUCTION

The U.S. healthcare industry lags far behind other industries in adoption and effective use of information technology. We urgently need to find ways to use IT to improve individual and population health, rein in healthcare costs, increase quality and accessibility, and provide breakthroughs in services, biomedical research, public health and public policy.

The most urgent need is for software-intensive systems analysis, requirements setting, and architecture for an effective and sustainable national-scale health information network (NHIN). At a minimum, a NHIN will provide availability of electronic health records anywhere for all individuals in the U.S. Ideally a NHIN will provide a secure, privacy-protected national trove of medical and other health data (from sensors, research, grocery stores, etc.) that will revolutionize health, our health systems, and the economy.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
FoSER 2010, November 7–8, 2010, Santa Fe, New Mexico, USA.
Copyright 2010 ACM 978-1-4503-0427-6/10/11 ...$10.00.

The software engineering research community has a vital role to play in devising a workable approach to creating a NHIN. However, traditional software engineering thinking is not ideally matched to problems at this scale. The NHIN will be *ultra-large-scale (ULS) system [2].* A national research and innovation initiative is needed to re-conceive the NHIN, guided by an *ULS systems perspective.* Collaborations among software researchers and experts in health and medical informatics and law and policy will be essential.

2. PROBLEM

The healthcare industry incurs over two trillion dollars in expenditures annually in the U.S. alone; yet it remains a decade or more behind in its adoption and effective use of information technology. One reason is that healthcare is a complex and human-intensive activity for which simplistic IT solutions have often failed. A second reason is that there is a long-standing disconnect between the health and medical informatics (HMI) and computer science and engineering (CSE) research communities. As a consequence, HMI is now operating with an inadequate approach to computing systems, and CSE is operating with an inadequate understanding of this enormously rich and important domain.

The nation can no longer afford this situation. The U.S. has already started to spend tens of billions of dollars in an attempt to create a NHIN in the next few years, but it is doing so absent basic engineering foundations in systems analysis, requirements and architecture.

The urgency of the NHIN undertaking stems from a combination of economic, political and legal factors. In economic terms, it is vital that we rein in spiraling healthcare costs while extending health insurance to many more citizens. In political terms, the President has promised an electronic health record (EHR) for each individual by 2014. Yet in 2008, only 17% of physicians used any EHR in their offices, and only 4% used a comprehensive EHR [1]. Nor does the U.S. have a national health data sharing system today other than by facsimile machine. In legal terms, the Health Information Technology for Economic and Clinical Health Act of 2009 (HITECH)—part of the American Recovery and Reinvestment Act of 2009 (ARRA)—provides incentive payments to healthcare providers who adopt and show *meaningful use* of certified EHR software systems, and eventual penalties for those who do not. Meaningful use is defined in phases, with interoperability being a key requirement in Phase II, starting in 2013. These incentives and

penalties are the main mechanisms by which policy-makers hope to catalyze the development of a NHIN in the required time-frame.

The prevailing statutory and policy framework effectively requires a transition, in just a few years, from a backward state of affairs to one with a nationally integrated system of healthcare systems covering all citizens. This would be an ambitious undertaking in any circumstances. What makes it particularly unlikely to succeed *on the present course* is that basic foundations have not yet been put in place in systems analysis, requirements and architecture.

There has been an intensive focus on data interchange standards, which are a necessary but insufficient basis for a national-scale information infrastructure. There is also a loosely defined architecture for the NHIN. In this design, a patient's records are scattered across EHR systems in institutions where the patient has been seen. When a complete record is needed, client-server interactions occur with other EHR systems, mediated by a *record locator service* (RLS). The RLS finds records for a given patient using *probabilistic demographic matching.* Using this approach, a client EHR queries for patient by name, age and other demographic information. The RLS then finds and returns records from other sites that "probably" belong to the given patient.

To date, no serious engineering analysis has been conducted, and no real case has been made to support the proposition that such an approach will work at a national scale, or that the resulting system will be sustainable, or that it will support many essential uses for health information now being envisioned. There have been prototyping efforts, but not at scale, and their findings are mixed at best. Evidence so far thus seems to point to the opposite conclusion. There are serious remaining risks in such areas as economic viability with respect to reengineering of installed legacy hospital EHR systems, system performance, data reliability and availability, privacy and security, and a lack of growth options to accommodate important uses of health-related data. To spend tens of billions of dollars on a critical national infrastructure without validated software and systems engineering foundations is not in the public interest.

3. APPROACH

What is needed now is an urgent national effort to bring together the software engineering research community, computer science and systems engineering more broadly, health and medical informatics research, and experts in relevant areas of law and policy, to address this situation. The goal of such an undertaking is to enable progress in the short term while ensuring that requirements and architectural commitments are made in a manner that creates a viable path the kind of system that the nation needs in the long term.

While much of what needs to be done is routine modernization and standardization, such an undertaking nevertheless requires leadership from the research community, because success requires more than just the application of existing knowledge. Many traditional software engineering assumptions—about life-cycle models, centralized control over development, detailed knowledge of implementations, even the ability to see the system as a whole—are not operative in systems at this scale. What we face is not an ordinary system-of-systems problem, but an *ultra-large-scale (ULS) systems* problem. What we need, then, is a national initiative that brings together academic research, industry, and government in an effort to understand and address the NHIN as a ULS systems problem. This section of this paper elaborates this perspective.

A web server is designed and engineered. The internet protocol, IP, and HTTP were designed and engineered. The World Wide Web (WWW) was not. Rather, it emerged from the decentralized and locally autonomous actions of many independent actors acting within the framework of the WWW architecture. A building is designed and engineered. A city is planned and governed, but it emerges largely outside the direct control of a designer or engineer. A garden plan is designed, but the garden emerges, without either the control or the need for actions by the gardener.

Traditional computing systems and systems-of-systems are designed and engineered. The *architectures* of ultra-large-scale systems are designed and engineered, but the systems themselves have to emerge from innumerable autonomous decisions and actions by self-interested entities (organizations, individuals, automated entities) within the system. Carefully designed and engineered technical architectures and infrastructures support and constraint the emergence of a flourishing socio-technical ecosystem. Thinking backwards, one sees that the nature and the needs of the socio-technical ecosystem must be understood and considered in designing and engineering the supporting architecture.

By any measure, the American health care system is an ultra-large-scale socio-technical ecosystem. A NHIN that integrates all American health care providers, patients, payers and others would be an ULS information system supporting this socio-technical ecosystem. To select its requirements and architecture requires serious consideration of the nature and needs of the health care ecosystem, including its players and their interests. Such an ULS system cannot readily be designed using traditional software engineering lifecycle models. It has to be designed from different perspective. The influential 2006 report, *Ultra-Large-Scale Systems: The Software Challenge of the Future* [2], provides such a perspective.

An ULS system is a software-reliant system that includes people, organizations, policy, economics, politics, and that, in technical scale and overall complexity exceeds the grasp of traditional software and systems engineering methods. Dimensions of scale include but are not limited to lines and kinds of code, intellectual investment, capital investment, data volume, number and kinds of hardware elements, variety of purposes and uses, number of processes, emergent behaviors, politial and institutional boundaries crossed, number of people and roles. New challenges arise in such systems: e.g., involving system dependability, system health monitoring, orchestration and governance, evolution, etc.

At such a scale many assumptions have to change: there can be no centralized control; no one can ever know all requirements; requirements will inevitably conflict; the system will always be both in production use and under construction; the distinction between system and user no longer makes sense: people are part of the system; localized failure is commonplace and not systemically harmful, yet there are critical points of systemic vulnerability and risk that must be treated with utmost care; ULS systems will inevitably be constructed on top of, or out of, already vastly complex legacy computing systems and institutions; traditional approaches to design and acquisition do not work.

To originate a ULS system requires a different kind of approach. Software is of the essence, but the process has to factor in many other issues, and the overall system will incorporate these issues as integral parts: economics, industry structure, legal and regulatory issues, certification and enforcement, incentive schemes, change, devolution of authority to the edge of the system, harnessing the value propositions of participants to foster bottom-up emergence, rather than through top-down control. A NHIN will be a ULS system. If it is not treated as such—if requirements and architecture are formed under the assumption that what has worked at smaller scale will work at the ULS systems scale—then it will in all likelihood either fail outright or impose a long-term burden on the country.

The ultimate product of such an ULS development effort effort is technical—requirements, architecture, standards, interfaces—but to get these right we have to start by considering the needs and capabilities of the ULS sociotechnical health ecosystem as it exists today, and understand how to catalyze significant change with clever, low-cost, rapidly deployable, and value-producing (not value-destroying) innovations. A truly successful result will lead the ecosystem to re-architect itself in pursuit of new sources of value created by architectural innovations, in much the same was as HTTP catalyzed a massive transformation of the internet and its use.

The ULS systems perspective can be understood as providing a set of principled heuristics that guide the search for effective solutions. Here are some of the principles that we believe are relevant to healthcare.

1. A national health information system for a democracy such as the U.S. should empower individuals with the full knowledge needed to understand, manage, improve and control their own health, health information, and healthcare (so a system cannot be just about exchange of medical records among healthcare providers).

2. A national health information system will have to support a broad diversity of data sources and processes extending far beyond mere interchange of medical records generated in clinical settings.

3. A ULS system architecture must support and organize an entire industry; institutions are basic units of modularity and composition in such an architecture, enabling rapid autonomous evolution within institutions and new compositions of institutions at the overall industry scale.

4. Trading rich but demanding integration requirements for rapid progress to a minimally functional but flexible system that works at scale can help bootstrap a self-catalyzing process of ULS systems development and evolution.

5. It is essential to fold a deep understanding of of long-term sustainability issues into short-term architectural decisions, to create options for long-term progress.

6. The research ecosystem should be meaningfully coupled, albeit somewhat loosely, to the larger industrial ecosystem and to the application domain.

7. National-scale information systems in vital sectors of the economy become critical national information infrastructure systems, and should be considered as such from the outset.

4. CONCLUSION

The authors are organizing software engineering and health informatics researchers and leaders to develop the vision sketched in this position paper. They invite expressions of interest in this initiative. It is time for these communities to confront the societal grand challenge problem of health and affordable, quality health care, through major innovations in software and information technology. The research opportunities for the engineering, computer science, and health informatics communities are enormous, especially, today, around the question of national health information systems.

5. ACKNOWLEDGMENTS

This work was supported in part by grants from the National Science Foundation. Acknowledgements go also to all those who contributed to an early proposal in this area, including groups at the University of California San Diego, Carnegie Mellon University, University of California Irvine, University of Massachusetts, California Institute of Technology and Emory University. We benefitted greatly from our interactions with these groups.

6. REFERENCES

[1] D. Catherine M. DesRoches, P. Eric G. Campbell, P. Sowmya R. Rao, S. Karen Donelan, M. Timothy G. Ferris, M.D., M. Ashish Jha, M.D., M. Rainu Kaushal, M.D., P. Douglas E. Levy, J. Sara Rosenbaum, P. Alexandra E. Shields, and M. David Blumenthal, M.D. Electronic health records in ambulatory care Ů a national survey of physicians. *New England Journal of Medicine*, 359(1):50–60, July 2008.

[2] P. Feiler, R. P. Gabriel, J. Goodenough, R. Linger, T. Longstaff, R. Kazman, M. Klein, L. Northrop, D. Schmidt, K. Sullivan, and K. Wallnau. *Ultra-Large-Scale Systems: The Software Challenge of the Future, Study Report.* Software Engineering Institute, 2006.