

Symposium on Natural Language Processing (NLP)

Building An Agenda for Research that Is Innovative and Implementable

David Atkins, MD, MPH

Director

Health Services Research and Development

September 9, 2015

Overview of Introduction

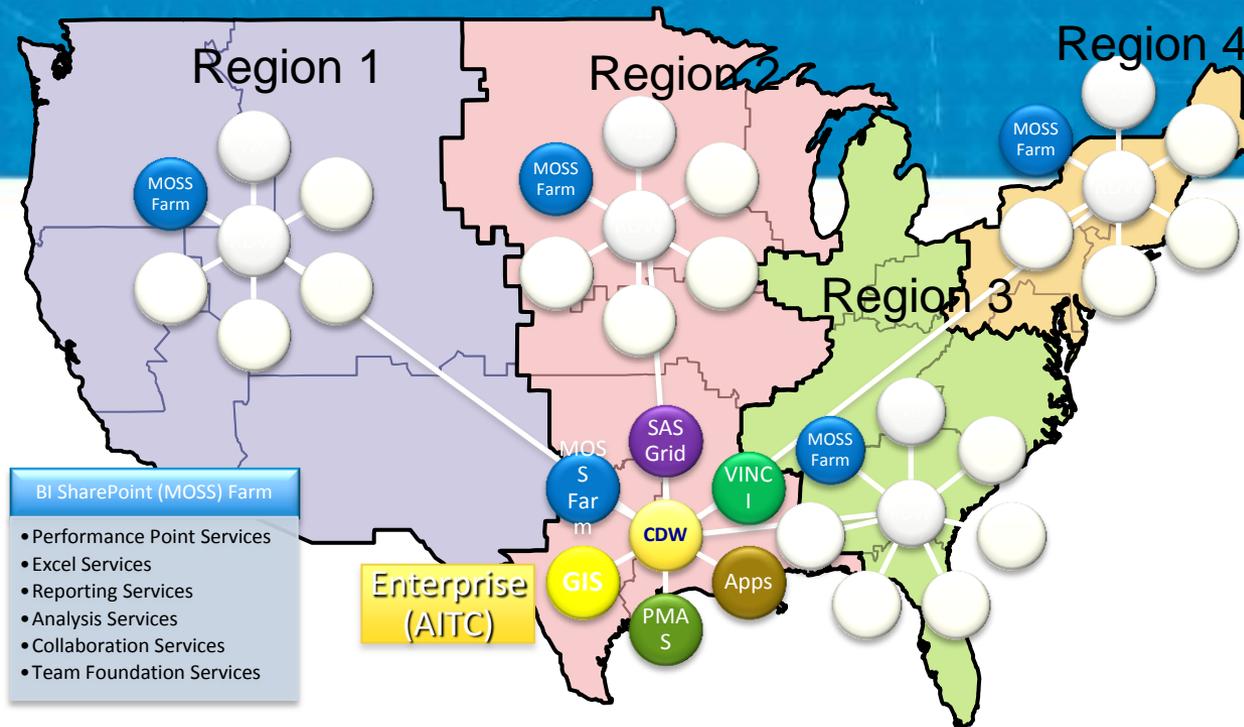
- Orientation to VA research and data
- Goals of HSRD informatics research portfolio
- Goals of this meeting: How can you help us?

Continuum of Health Services Research



- **Foundational research** to improve data sets, validate measures, develop tools for investigation
- Observational research on how **patient outcomes** vary with factors at the level of patient, provider, practice, facility and organization
- **Research on innovations in practice and systems interventions** to reduce variation and improve access, satisfaction, quality or efficiency
- **Implementation research** to improve implementation and spread of effective health service innovations (much of this falls in QUERI)
- **Measurement and evaluation** research to demonstrate that efforts lead to lasting improvements

VA Data



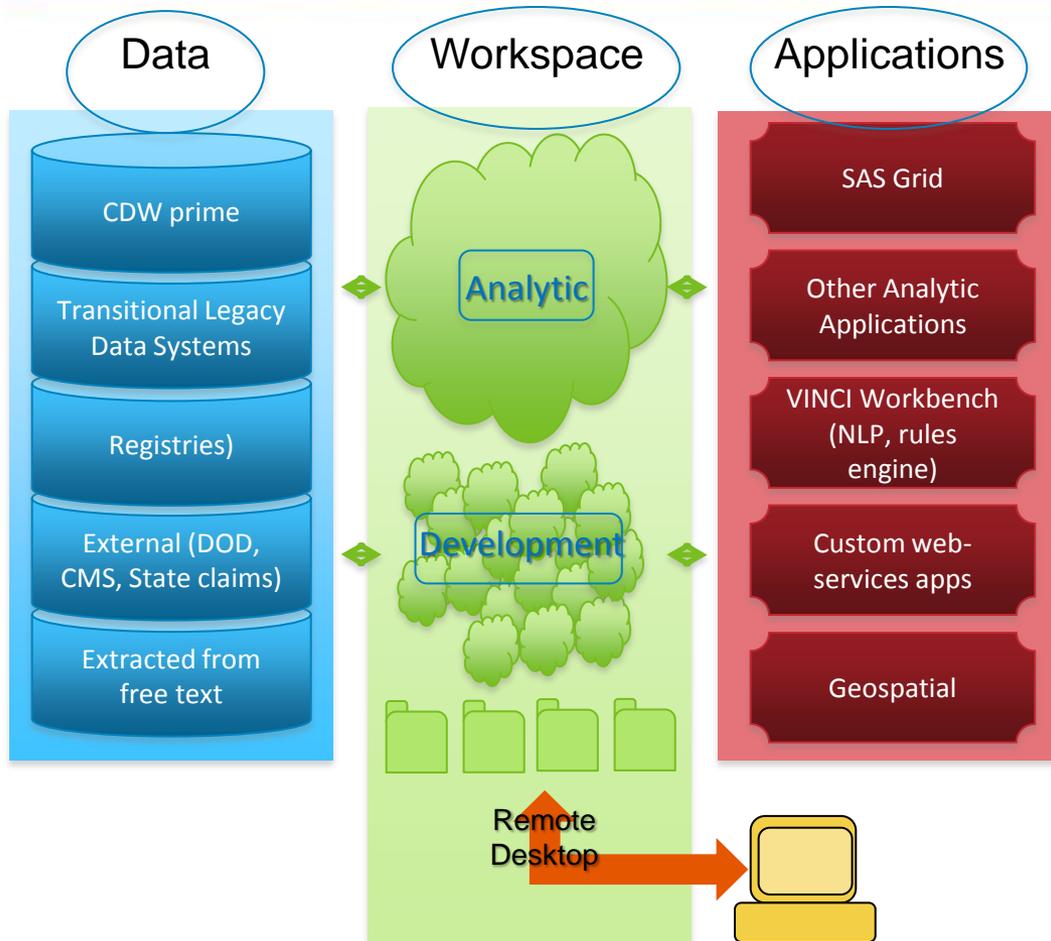
Annually:

- 1.6 billion outpatient encounters, 9 million hospital admissions
- 3.2 billion clinical orders; 5.6 billion lab tests, 1.5 billion prescriptions filled
- 165 million radiology procedures
- 2 billion text notes (unstructured), hugely enriched with a lot of information

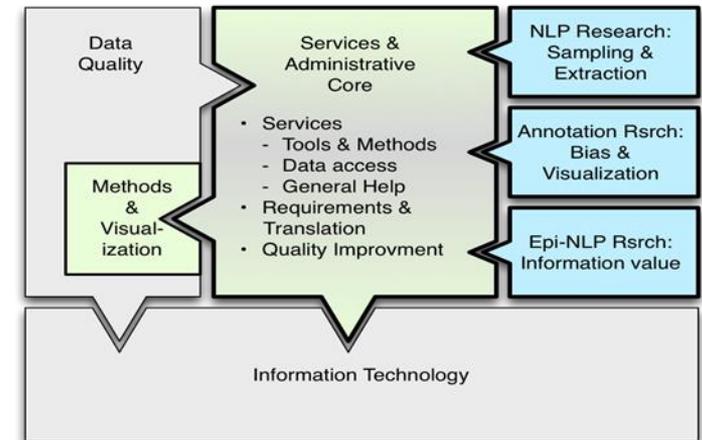
Typical day:

- 420,000 patient encounters, 2.4 million lab results, 553,000 pharmacy fills.

Virtual Computing Environment and Research



VINCI Research



Data in VINCI

- 22.6 million patients (7.4m seen/year, 927k new/year)
- 7.2 billion lab tests (1.6m new/day)
- 4.3 billion orders (949k new/day)
- 3.0 billion procedures (721k new/day)
- 2.8 billion clinical notes (865k new/day)
- 2.1 billion medication fills (405k new/day)
- 2.2 billion outpatient visits (597k new/day)
- 14.7 million inpatient visits (2.2k new/day)

VA Research Priorities that Could Benefit from NLP

- Creating and refining disease phenotypes for genomic research with Million Veterans Program (MVP)
- Improving ability to control for confounders in comparative effectiveness research (e.g., disease severity)
- Improve outcome identification for research on quality and safety (e.g., adverse events)
- Understanding patient experience of care
 - Patient reported outcomes
- Understanding coordination of care (including non-VA care)
- Monitoring for changes in function and disability relevant to Veteran benefits
- Surveillance for new symptom clusters, emerging diseases, etc.

Challenges for VA research on NLP

- We have small but growing pool of researchers with NLP expertise
- Modest budgets for NLP and other informatics research
- We want to move from research that uses NLP as a tool to some a narrow problem to research that is more *innovative* and advances more *generalizable solutions*
- At same time, we want our research to have potential for direct applications in VA care
 - Partnerships with Office of Informatics and Analytics who direct use of CDW data for operational purposes
- **What are innovative new advances in NLP that our researchers should know about?**
- **Where are VA's unique advantages that may allow us to make unique contributions to NLP research?**