

## NEW CHALLENGES FOR AI IN MAINTAINING QUALITY AND REDUCING COSTS IN HEALTH CARE

**Bruce G. Buchanan**  
University Professor of Computer Science Emeritus  
(Formerly also Professor of Philosophy, Medicine, and Intelligent Systems)  
University of Pittsburgh

B.Buchanan 2/21/12 1

### MAIN POINTS OF THIS TALK


- I. Small apps on mobile devices are exploiting prior research in AI and the plethora of information on the ever-expanding web.
- II. Follow the money to find opportunities
- III. A new challenge for AI is helping people avoid both the volume of relevant information and the number of relevant apps.

B.Buchanan 2/21/12 2

### \$10m X-Prize

- The \$10 million bounty will go to the first team that can create “a mobile platform that most accurately diagnoses a set of 15 diseases across 30 [patients] in three days”

Old Style Medical Decision Making Program



B.Buchanan 2/21/12 3

### Centralizing: Electronic Health Records

- Nearly 40 percent of American primary care doctors and about 25 percent of hospitals use electronic patient records.
- Suppliers of health data systems include Epic, Alscripts, Meditech, Cerner and units of diversified giants like I.B.M., McKesson, Siemens and GE Healthcare.
- “All of those companies’ varying systems have raised concerns about usability, especially when different systems must share information. *Some wonder if the government should play a bigger role in creating uniform technical standards and designs across systems.*”

B.Buchanan 2/21/12 4

### I. What’s Different Now?


- Interconnectedness
  - Social Computing
  - Crowdsourcing
  - “Internet of Things”
- Mobile Devices
  - Ubiquity
  - Voice + Vision
- Volume of Accessible Data
  - Reminding / Alerting Systems
  - NLP -- Semantics + Statistics
  - Data Mining
- Mature AI Technology
  - Natural Language
  - Vision
  - Speech

Some New Apps

B.Buchanan 2/21/12 5

### Reference: From PC to Phone

- Micromedex offers iPhone and iPad apps for its drug reference guide and medication interaction checker.



Looks like a book

B.Buchanan 2/21/12 6

### Reference: From PC to Phone

Medscape from WebMD

B.Buchanan 2/21/12 7

### Social Networking

- Sermo -- free app "allows physicians to share clinical images, ask questions, and discuss and share ideas on clinical cases, health policy, practice management, and more."

B.Buchanan 2/21/12 8

### Crowdsourcing

- **Foldit** by Seth Cooper et al., University of Washington a game designed to tackle the problem of protein folding.
- "Foldit showed that it is possible to effectively "crowdsource" human problem solving to solve very hard scientific problems."

B.Buchanan 2/21/12 9

### Foldit

B.Buchanan 2/21/12 10

### Today, the Internet -- tomorrow, the Internet of Things?

- Anything with intelligence (including machines, roads and buildings) will have an online presence, generating data that could be put to uses currently unimagined.
- In healthcare there should be 1.5 million devices at the end of the year, with a growth rate of 20% to 25%. -- usually worn by a patient to monitor a chronic condition, such as a device that advises a heart disease patient when to take medicine

[ComputerWorld, Nov. 9, 2011]

B.Buchanan 2/21/12 11


### Data Collection

"One experimental device fits snugly around an animal heart, collecting information and stimulating the heart muscles with electric current like a nearly invisible pacemaker. Other circuits can be attached to the brain or skin like a temporary tattoo."

B.Buchanan 2/21/12 11

### Advances in Vision

- **MIM Mobile** – “a remote diagnostic imaging tool for the iPad, iPhone, and iPod touch.”




B.Buchanan 2/21/12 13

### Mobile Devices: Ubiquity

“Four out of five practicing physicians use smartphones, computer tablets, various mobile devices, and numerous apps in their customary medical practices”

[ InformationWeek October 21, 2011 ]


Apple sells more iPhones per day (402k) than people born in the World per day (300k)  
[<http://cnet.co/zNhPz6j>]



B.Buchanan 2/21/12 14

### Interconnection: Alerting

- Patient monitors are connected to a network and every alarm they produce is captured by AlertLink. A small percentage of those alarms are considered life critical and can be forwarded to a caregiver’s mobile device, such as a phone.



B.Buchanan 2/21/12 15

### Petabytes of Online Data = Thousands of Terabytes, Millions of Gigabytes

AMAZON.COM – 59 million active customers; Personal information (phone number address, etc), receipts, wishlists, and virtually any sort of data the website can extract from its users while they are logged on. Plus 250K books online (full text)

GOOGLE.COM  
 91 million searches per day (half of all Internet searches)  
 33 trillion entries stored in databases per year  
 ETC.

B.Buchanan 2/21/12 16

### Recognizing the Value of Large Amounts of Data: FDA Initiative

- FDA currently houses the largest known repository of de-identified clinical data -- safety, efficacy, and performance + postmarket safety surveillance data.
- allow us to address fundamental scientific questions about how different types of patients respond to therapy.

B.Buchanan 2/21/12 17

### Advances in NLP: Medical Data in Text

- “Approximately 1.2 billion clinical documents are produced in the United States each year. These documents comprise around 60% of clinical information...”

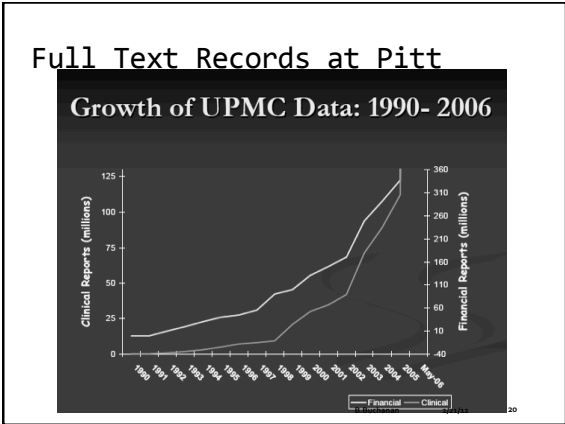
HealthStory Project  
<http://www.healthstory.com/about/about.htm>

B.Buchanan 2/21/12 18

### Full-Text Medical Records at Pitt: MARS

- History & Physical + Discharge Summaries
- Lab Results
- Pharmacy Orders
- Operative Reports
- Outpatient Progress Notes
- Radiology & Cardiology Reports
- Financial Charge Details

B.Buchanan 2/21/12 19



### Natural Language Understanding

- De-ID System: Making Full Text Records Anonymous
  - [G.Cooper, B.Buchanan, P.Hanbury, M.Saul]
  - Heuristics find any of 18 HIPAA-specified identifiers
  - Dictionaries of place names
  - Medical terms with proper names preserved
  - Same tags used in multiple replacements of same name
- [Evaluation of a Deidentification (De-Id) Software Engine to Share Pathology Reports and Clinical Documents for Research Dilip Gupta, MD,<sup>1</sup> Melissa Saul,<sup>2</sup> and John Gilbertson, MD<sup>1</sup>. Am J Clin Pathol 2004;121:176-186]
- Neg-Ex: Determining Negations in Medical Text
  - [Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. *Journal of Biomedical Informatics* 34, (2002) 301-310.]
- IBM's Watson: Statistical NLP
- HealthMap (UW): Reading and Summarizing Text

B.Buchanan 2/21/12 21

### WATSON: Coping with the Volume of Information

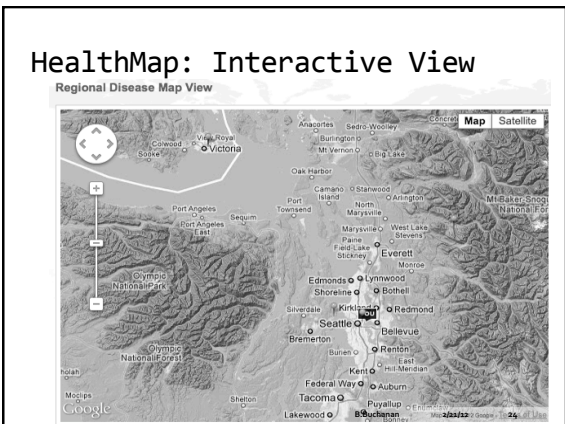
- Start with credible sources (e.g., textbooks, encyclopedias, peer-reviewed jnls)
- Associations of terms within 200 million pages of current documents will suggest correct answers
- Augment associations with some knowledge of grammar, language use (e.g., metaphors), real world

B.Buchanan 2/21/12 22

### HealthMap: Reading Text from GoogleNews + Other Sources

- Disease alerts within 100 miles of Seattle, Washington, United States in the past month.
- **Gastrointestinal**
- **Jan 16**  
Victoria, British Columbia, Canada  
Dozens remain isolated in Victoria hotel rooms after Norovirus outbreak - Victoria Times Colonist
- ...
- **Respiratory**
- **Jan 13**  
Cowlitz County, Washington, United States  
More cases of whooping cough documented in Cowlitz County - Longview Daily News
- **Jan 13**  
Grant County, Washington, United States  
Grant whooping cough outbreak downgraded - TheNewsTribune.com
- ...

B.Buchanan 2/21/12 23



## II. How Can Small Apps Make a Difference?

### II-A. Follow the Money

- In 2000, health care spending rose to \$1.3 trillion dollars, or an average of \$4,637 per person.

Some Documented Costs

B.Buchanan 2/21/12 25

### Costs: Physicians' Time

- 15-17% of their time in direct patient care.
- 64-67% of their time in indirect patient care (divided between reviewing results, performing documentation, and engaging in communication)
- Reviewing results and carrying out documentation are clearly information-focused in nature, which means that physicians spend about 35-40% of their time engaged with information

[Bill Hersh  
<http://informaticsprofessor.blogspot.com/2012/02/is-medicine-information-science.html>]

B.Buchanan 2/21/12 26

### COSTS: Waste in the System

Here are some of the contributors to the \$1.2 trillion being leaked out of the system.

OVERTESTING	\$210B
PROCESSING CLAIMS	UP TO \$210B
IGNORING DOCTOR'S ORDERS	\$100B
INEFFECTIVE USE OF TECHNOLOGY	UP TO \$88B
HOSPITAL READMISSIONS	\$25B
MEDICAL ERRORS	\$17B
UNNECESSARY ER VISITS	\$14B
HOSPITAL ACQUIRED INFECTIONS	\$3B

["Health Care's Wasted Dollars"  
<http://money.cnn.com>  
 Aug. 10, 2009 ]

B.Buchanan 2/21/12 27

### COSTS: Hospital ER Visit (2008, MEPS)

- Average expenses for a visit to the Emergency Room were \$1265 in 2008

- Ages 45 to 64	\$1681
- Ages 65 and up	\$1554
- Children under age 18	\$351 - \$412

B.Buchanan 2/21/12 28

### COSTS: Claims Processing

- Inefficient claims processing costs as much as \$210 billion annually
- "We spend a lot of time and money trying to get paid by insurers," said Dr. Terry McGenney, a Kansas City, Mo.-based family physician.
- "Every insurance company has its own forms," McGenney said. "Some practices spend 40% of their revenue filling out paperwork that has nothing to do with patient care. So much of this could be automated."

B.Buchanan 2/21/12 29

### Costs: Non-Compliance

- 10-25% of hospital & nursing home admissions, resulting in 340 deaths per day
- 20% of unintentional pregnancies in the US at a cost of \$2.6 billion
- 3 times as many doctor visits & \$2000 per year in additional costs compared to patients who follow their treatment plan
- 33-69% of all medication-related hospital admissions in the US at a cost of \$100 billion

B.Buchanan 2/21/12 30

### COSTS: Medical Errors

- 98,000 Deaths, \$29 billion in costs
- Solution = Computerized Medical Records
- Systems Investment of \$115 billion over 15 years can produce yearly savings of \$81 billion from efficiency and error avoidance

B.Buchanan 2/21/12 31

### COSTS: ICU STAY


- Mean length of stay in ICU (2002) = 8.5 - 14.4 days (without or with mechanical ventilation)
- Cost on 3<sup>rd</sup> and subsequent days (2002) = \$3184 - \$3968 (without or with mechanical ventilation),

B.Buchanan 2/21/12 32

### HOW CAN SMALL APPS MAKE A DIFFERENCE?

- Find a small, well-defined problem
- Create a niche app that saves time or money
- Multiply its uses across the U.S.
- THEN Multiply the number of apps

II-B. Some Hypothetical Examples



B.Buchanan 2/21/12 33

### THE MULTIPLIERS

- # Hospitals (U.S.) ~ 12,000
- # Clinics, Ctrs, etc. ~ 36,000
- #Physicians & Surgeons > 972,000 (2009)
- # Staffed Beds ~ 1 m
- # Hosp.Admissions > 37 m
- # E.D. Visits (2008) > 123 m
- Hospital Charges > \$725 b

B.Buchanan 2/21/12 34

### EXAMPLE: Clerical Time

- Small app that saves 5 min. of clerical time on every admission – used in half the hospitals
- At \$24 per hour, 5 min. = \$2
- \$2 x 0.5 x 37m admissions = \$37m annually

B.Buchanan 2/21/12 35

### EXAMPLE: PATIENT EDUCATION -- When to Visit an E.D.

- Small app that lets patients avoid a trip to an E.D. 1% of the time, reducing #visits from 123m visits for 300m people to 122m annually.
- Assume \$569 per visit (median cost in 2008)
- \$569 x 1m = \$700m savings

B.Buchanan 2/21/12 36

**EXAMPLE: PATIENT EDUCATION – Compliance on Meds**

- Record schedule & doses – with reminders
- Explain package insert information – common, harmless side effects vs serious ones

Save a fraction of \$100 billion (from the 33-69% of all medication-related hospital admissions in the US)

B.Buchanan 2/21/12 37

**EXAMPLE: Claims Processing (\$210b / yr)**

- Small set of apps that fill out forms (one per form)
- Save 5 min. per form x 6 forms per day
- = 30 min./day x 900K physicians & surgeons

B.Buchanan 2/21/12 38

**Example: Transcription Time**

- **“Allscripts Broadens Speech Recognition Choices In EHRs: EHR vendor now will offer both M\*Modal and Dragon’s software to its customers.”**  
[By Nicole Lewis InformationWeek, February 01, 2012]
- 1 minute / note X \$1 / min. X millions of notes

B.Buchanan 2/21/12 39

**EXAMPLE: Adverse Drug Events**

- Over 770,000 people are injured or die each year in hospitals from adverse drug events (ADEs)
- Costs estimated at \$1.56 to \$5.6 billion annually.<sup>47</sup>
- Patients who experienced ADEs were hospitalized an average of 8 to 12 days longer than patients who did not suffer ADEs, and their hospitalization cost \$16,000 to \$24,000 more.
- “Anywhere from 28 percent to 95 percent of ADEs can be prevented by reducing medication errors through computerized monitoring systems.”

B.Buchanan 2/21/12 40

**ADE’s (contd.)**

- In 1992, 567 ADEs cost LDS Hospital \$1.1 million in direct expenses (not including liability costs or the cost of injuries to patients). If 50 percent of these ADEs had been prevented, LDS Hospital would have saved \$500,000 a year.<sup>2</sup>
- Brigham and Women’s Hospital would have saved \$480,000 annually if the 17 percent decrease in ADEs had been applied hospital-wide; this figure does not include the costs to patients of injuries, malpractice costs, or the expense of additional work required to correct medication errors and treat patients who suffer from ADEs.<sup>23</sup>
- Doctors at Wishard Memorial Hospital who used computer order-writing workstations had 13 percent lower inpatient charges (about \$900 per hospitalization) than those who used paper forms.<sup>25</sup>

B.Buchanan 2/21/12 41

**Example: Avoiding Errors with Reminding Systems**

- WellPoint Inc., which operates Blue Cross Blue Shield plans in 14 states, will use Watson’s vast health care database and quick calculating power to guide treatment decisions for its 34.2 million members.
- In its new job, Watson will swiftly compare patients’ electronic records to a mammoth library of textbooks and medical journals, and WellPoint’s history of treatments to generate a list of options and the rationale behind them.”  
[ABC News, Sept. 12, 2011]

B.Buchanan 2/21/12 42

### III. Challenge for AI

- Recommend the Apps
  - Reviews & Recommendations (text + ratings) ✓
  - Specific Medical Context
  - Specific Features
  
- Learn to Distinguish the Best ✓
  - For Each Medical Context

B.Buchanan 2/21/12 43

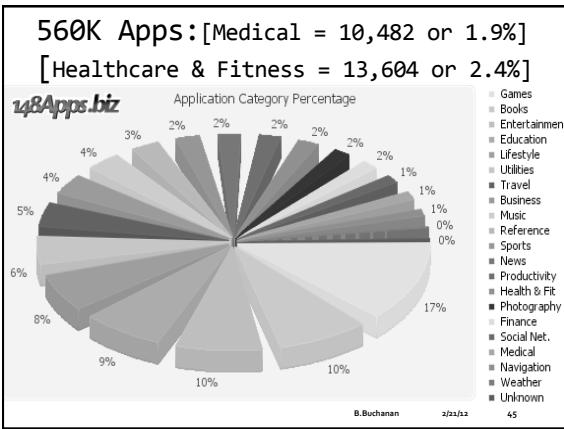
### CHALLENGE: TOO MANY APPS

- e.g., "Finding Good Apps for Children With Autism" 400-700, depending on the list
- "...it's daunting to know where to begin when sorting through the thousands of programs available. "There are so many apps, and not all of them are great," ... To that end, several good Web sites have been created to review special-needs apps."



- [NYTimes, Nov. 29, 2011]

B.Buchanan 2/21/12 44



### CHALLENGE: RATE OF NEW APPS

- Total Apps in the App Store Currently Available for Download = **560,990**
- Number of Active Publishers in the US App Store = **134,074**
- Count of Application Submissions January, 2012
  - This Month (Games): 3,598 ( 116 / day )
  - This Month (Non-Games): 19,430 ( 627 / day )
  - This Month (Total): 23,028 ( 743 / day )

B.Buchanan 2/21/12 46

### NewsFinder: Sorting the Wheat from the Chaff

- Goal = alert subscribers to "interesting" articles about AI from credible sources in the last week
- Problems with keyword searching
  - (a) Too many hits (e.g. >8000 mentioning 'artificial intelligence' or 'robot' or 'data mining' from Google News)
    - Many duplicates (e.g., stories about Watson)
    - Many local stories (e.g., high school robot teams)
  - (b) No sense of why a story is interesting
  - (c) No sense of the breadth of AI (e.g., ethics, history,...)

B.Buchanan 2/21/12 47

### NewsFinder

- Train 19 SVMs (one for each topic; e.g., agents, ethics, history, machine learning) – Keep Learning!
- Remove 583 Stopwords; Represent as vector (~180 words)
- Assign topic names to ~170 articles found online at 37 sources (+GoogleNews) each week
  - Remove untagged stories
  - Must not contain any of 458 "Bad Words"
  - Must contain at least 2 (of 107) "Whitelist" term
  - Remove very similar stories (>cosine similarity thresh.)
- Select 12 or fewer of the most relevant stories


B.Buchanan 2/21/12 48



### NewsFinder Extensions

- QUESTION: Can NewsFinder be modified to find interesting stories in medical informatics?
  - New Whitelist;
  - New set of topic categories
  - New set of preferred sources

FUTURE: Use same procedure to read reviews and select best apps



B.Buchanan 2/21/12 49

### Conclusion: Finding relevant information is half the battle

The next half includes

- Selecting the most relevant information for a specific context
- Selecting the best apps for the specific context
- Integrating the results from running several apps

B.Buchanan 2/21/12 50

# FINI

Thank you.

B.Buchanan 2/21/12 51