Digital Libraries Laboratory

Members of the DL Lab
University of Cape Town
Department of Computer Science
Digital Libraries Laboratory @ Centre for ICT4D

May 2017

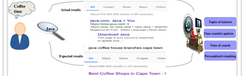
Information Retrieval Products

http://libraryweekcollection.co.uct.ac.za/ http://search.ndltd.org/

Information Retrieval

- Catherine Chavula (PhD)
- Selvas Mwanza (PhD)
- Jivashi Nagar (PhD)
- Gina Paihama (PhD)
- Mushashu Lumpu (M.CS)
- Jackson Moji (M.CS)
- King'ori Maina (M.CS, submitted!)
- Morebodi Modise (M.CS, submitted!)

Do Topics & Time Go Together?



African Language Content Generation

- Problem
  - Lack of African language resources in digital form
  - Insufficient support for African languages in software
- Proposed Solution:
  - Create a focused archive of writings
  - Encourage translation of popular content e.g., Wikipedia
  - Use gamification for long-term sustainability
    - Research focus: Will this approach work?
- MSc student: Jackson Moji

Knowledge Engineering

- Joan Byamugisha (PhD)
- Zola Mahlaza (M.CS)
- Jarvis Mutakha (M.CS)
- Cameron Kyle (M.IT)
- Zubeida Khan (PhD)
- Yamiko Msosa (PhD)
- Iman Neguib (M.IT)

Multiple choice question generation for learners

- Key Ideas
  - Using isiZulu and isiXhosa grammar engine
  - Generating MCQs from a domain using grammar engine
  - Measuring quality of MCQs
- Status
  - Proposal drafting
- MSc candidate: Jarvis Mutakha

Changes in clinical practice guidelines (Yamiko Msosa)

- CPGs change frequently to reflect emerging medical practice
- Makes it hard to maintain software systems using CPG
- Find a way to automate that
- Current main outcomes:
  - Types of changes in the (fine-grained) components of the guidelines
  - Model-driven engineering q domain-specific language (DSL) to model, annotate the computerised CPGs
  - Evaluated DSL on adequacy and usability

Educational Technology

- Lighton Phiri (PhD)
- Sinini Ncube (M.CS)
- Maryam Almukhaylid (M.CS)
- Devan Govender (M.IT)

A Comparison of Socially-Motivated Discussion Forums Models for LMSes

- MSc student: Maryam Almukhaylid

Applications of ICT4D

- Blessing Ojeme (PhD)
- Sarah Brittan (M.CS)
- Josiah Chavula (PhD, submitted!)
- Haji Ali Haji (PhD, submitted!)
- Jessica Mantel (M.IT, submitted!)

Digital Libraries

- Making information available to people
- Context-sensitive: low resource environment, different languages, different skills/culture, etc.
- 4 areas of interest:
  - Information Retrieval - HS
  - Educational Technology - HS
  - Knowledge Engineering - CMK
  - ICT4D Applications - various
- Some production systems / products

KnowledgE ENgineering team

- **Aim**: to contribute computing theory, methods, and techniques to the knowledge society
- **Focus**: knowledge engineering in its broad sense; includes ontology engineering, the Semantic Web, intelligent (rule-based, ontology-driven) conceptual modelling, and natural language generation.
- Some of the current topics:
  - OE: modularity, Test-Driven Development
  - NLG: Agglutinative languages in Sub-Saharan (data and knowledge, and mainly in health and agriculture)
  - CDM: model interoperability, temporal modelling
- More information: http://www.meteck.org/keen/

Retrieval for Related African Languages

- Context
  - Low quality search results
  - Widespread multilingualism
- Approach

- Status
  - Implementation and experiments

Information Retrieval for Unemployment

- Key Ideas
  - Specialised IR approaches for finding training opportunities
  - E.g.: summaries, focused search
- Status
  - early experiments

- PhD candidate: Gina Paihama

A Grammar Engine to Verbalize Ontologies in Runyankore

Motivation

- NLG successfully applied to generate personalized patient information

Problem

- Templates which have successfully generated personalized patient information are inapplicable to Bantu Languages

Solution: Grammar Engine

Results

NLG for financial data (Cameron Kyle)

Intelligent temporal conceptual modelling (Iman Neguib)

- Temporal data and conceptual models
  - e.g., in health, business, security
    - Dev: ,(PhDStudentAcademic) : "An Academic may have been a PhD student before, but is not a PhD Student now"
    - "a HIV positive patient may not evolve into a blood donor"
- How to convert a temporal model into a temporal database?
- Expected outcome:
  - Algorithms that ensure that the constraints of the temporal model apply to the actual database

Streamlining Technology-driven Orchestration

1) Activity Management
2) Resource Management
3) Sequencing
4) Playback Sequence

- Experimental Results
  - Effectiveness & user experience
  - Orchestration load
  - Practical usage scenarios

Investigating audio classification for automating the trimming of videos in Opencast

- Key Ideas
  - Classify audio into respective classes
  - Use classes to tag respective video frames for trimming
- Status
  - Training classifier

- MIT student: Devan Govender

Multi-stakeholder design of mobile micro-enterprise financial management system

- Key Ideas
  - Design mobile banking system for micro-entrepreneurs with users/NGO
  - Multi-stakeholder design (before co-design ideas)
- Status
  - Writing up thesis
- MSc student: Sarah Brittan

Information Retrieval and Educational Technology

- Information Retrieval (IR):
  - Search engine algorithms in African languages
  - IR for Development; Innovations in IR
  - Corpus development for African language content
  - Applied into local Digital Library applications, such as heritage archives
- Educational Technology:
  - Low-cost innovation, mobile tools
  - Focus on the teacher and not just the student
  - Modern technology and controlled studies

Social Media for Development (SM4D)

By: Selvas Mwanza

Digital Libraries in Clouds

- Key Idea
  - Digital repositories increase sharing of content
  - Can we automate creating/installing/configuring in private clouds?
- Status
  - Some experiments done, some to do
- MSc student: Mushashu Lumpu

IsiZulu & isiXhosa weather bulletins using NLG

- There is no fast and large scale producer, automated or otherwise, of weather summaries in Nguni languages
- 'Multilingual' natural language generation system
- Our focus:
  - 'Single' grammar compatible with isiXhosa and isiZulu
  - Grammatical similarities for isiZulu and isiXhosa verbs
  - Impact of phonological conditioning in weather-based NLG (**izaukuba -> izokuba**)

Ontology Engineering: Modularity (Zubeida Khan)

- Many ontologies, and large and/or expressive ones
- Problems: cognitive overload, scalability, reuse of fragments
- Solution: ontology modules
- How to create them? quality? maintainability? etc.
- Main achievements:
  - Ontology modularisation framework for ontology modularisation
  - Evaluation metrics (and dependencies between them) for ontology modules
  - Novel algorithms for creating good modules
  - Swapping OOP modules (e.g., a foundational ontology)
  - Tooling support for these tasks
- More information: http://www.theifiles.co.za/research.html

Integrated in-Content Communication for Shared Textbook Use

- Key Ideas
  - Mobile electronic textbooks
  - Discussion forums integrated with content
    - vs. standalone discussion forums
- Status
  - Currently writing up thesis

- PhD candidate: Sinini Ncube

questions, comments, ...

http://dl.cs.uct.ac.za/
http://www.meteck.org/keen/index.html

enkosi
hamba kakuhle
thank you and go well