# Cognitive-analytics-driven-personalized-learning

# 1 Introduction

## Personalized Learning Characterized By Several Attributes:

1. No predefined order for learning
2. Pre-test and Post-Test need to be done, before move on and receive the credit
3. Test question generated and assessed automatically
4. Student got feedback to promote reflective learning
5. The content that has been reusing would be the central concepts

Personalization, eLearning, online learning, digital learning are same term

Lack of interaction among learners and personalization of learning to suit students’ goals and styles of learning are two challenge.

Big data store data that is generate by machine and contributed by human through 2.0 applications such as social media and crowdsourcing.

Big data has been using in multiple work field for different purpose, in educational aspect, it has also known as academic analytics, learning analytics and educational data mining.

Online learning frameworks offer guidance in designing, developing, operating, and evaluating eLearning systems. Proposed by Khan.

# 2. Darwin Information Typing Architecture

Darwin Information Typing Architecture is an Xml data model for authoring and publishing

## Design and development personalized learning requires:

1. Identify and developing material that have fine granularity( the scale or level or detail in a set of data)
2. Make sure the material match with student, different learning style and gender responsive
3. Bloom’s taxonomy?
4. Question generate automatically, using generation algorithms

## Two qualities for reusing teaching and learning materials:

1. Fine granularity: a higher level of detail, or a finer (<https://www.quora.com/What-does-granularity-mean-in-computer-science-computer-networks>).
2. Context-freeness: (context-freeness grammar) means all of its production rules have a single non-terminal on their left hand side (<https://softwareengineering.stackexchange.com/questions/253454/what-does-context-free-mean-in-the-term-context-free-grammar>).

Fine granularity and context-freeness are not orthogonal to each other.

Darwin Information Typing Architecture (DITA) uses eXtensible Markup Language (XML) base for structuring, developing, managing, and publishing content.

DITA requires topic based authoring and will create a topic map.

Content filters help to search and select topics for inclusion in topic maps

# 3. Automated Question Generation

Question bank has serious shortcomings. Question bank aren’t good enough and the question become simple problem after few semester used

Generate question become necessary and important for personalized learning

## 3.1 Existing Approaches to Question Generation

Two major approaches to automatic question generation research

1. Generating questions from natural language texts
2. Generate data and questions in a more structured and constrained contexts

Predictive questions: a set of problem that user might asked in the domain

Query templates: provide answer to the question base on the answer that extracting from the knowledge base.

Ontologies: an ontology is a formal naming and definition of the types, properties, and interrelationships of the entities that really or fundamentally exist for a particular domain of discourse (<https://en.wikipedia.org/wiki/Ontology_(information_science)>)

IndusMarker, an automated generation system that generate short answer

OntoQue system user knowledge inherent generate semantically correct answers

## 3.2 Proposed Approach to Question Generation

Template variable: required input parameter for question generation

# 4 Cognitive Analytics and Personalized Learning

There is some problem with precise algorithmic solution

Big data enables a new paradigm for solving these problems

## 4.1 Data-driven Solutions

Algorithms for POS tagging fall:

Rule-based: Rule-based algorithms make use of rules, which are either crafted manually or learned through an iterative process (really common in machine learning)

Stochastic: Stochastic POS algorithms are data-driven and are based on supervised learning models such as Hidden Markov model, Log-Linear model and Conditional Random Field

Stochastic perform better than rule based algorithms

Data-driven approaches is better than traditional and rules-based methods, it doesn’t require changes to the application

Domain adaptation is also another problem, in data-driven ,the problem is solve by retraining the parser with text from the relevant domain

Cognitive systems learn from environment and interact with human naturally, focus on finding reasonable solution rather than best solution

Cognitive analytics

## 4.2 Proposed Framework

5 layers of proposed framework

1. Foundational technologies
2. Storage of large data, enforces data integrity and privacy requirement, also provide basic data services
3. Libraries for the data analysis, visualization and graph algorithms, also provide machine learning and artificial intelligence libraries

First three layer could also be used for other applications

1. Automated question generation and assessment, personalization, taxonomies and ontologies
2. Interface, provide authentication service to control access to the system. Authorization, API

# 5. Related Work

Learning analytics (LA) concepts and methods are drawn from following fields: academic analytics, action analytics, educational data mining, recommender system, and personalized adaptive learning.

Reference base on four mention

1. What: identifies data collection, and management
2. Who: the target stakeholder
3. Why: motivation for data analysis
4. How: type of processing performed on the collected data

Learning Object (LO): Different difficulty (easy, medium, difficult), learning style (visual, auditory, kinesthetic), labels (theory, example, exercise, assessment)

Learning analytics toolkit, using graphical indicators to record down user usage and all different kind of learning information and use them to improve.

Burden and Kearney describe approach to mobile learning. Research of authentic learning

# 6. An Interactive System for Personalized Learning

Layer 1 computing infrastructure: using compute cluster or hosting platform, provides parallel libraries and MapReduce frameworks

Layer 2 data layer: store structured and unstructured data, using rational and NoSQL databases.

LMS Data: data collected by the learning management system

Assessment data: keep tracks of learner assessment score, times, attempts

DITA base teaching and learning materials are store in layer 2

Layer 3 extracts information and knowledge from the diverse data sources that reside in layer 2

Machine learning algorithms

Natural language understanding libraries

Domain taxonomies and ontologies

Layer 4 3 types of analytics

* Educational data mining base data analytics(EDM)
* Learning analytics(LA)
* Cognitive analytics(CA)

Layer 5

* Personalization
* User modeling : capture user information
* Automated question generation
* Contextualized feedback to learners

ISPel use these data with taxonomies and ontologies to give better options that are available to learners.

Layer 6 Interfacing with the outside world

Authentication and authorization function provide authenticating users and controlling access

ISPeL three type of user: student, teacher and administrators