

Computational Technologies For Accreditation In Higher Education

Aboubekeur Hamdi-Cherif

Abstract— Academic accreditation and assessment in Higher Education (A3-HE) is, above all, a social status meant to acknowledge that an institution or program is following recognized and requested quality criteria issued from common good practice. In a previous work, we described the main processes involved in A3-HE. Two main issues were reported. First, heavy and tedious paperwork characterizes actual academic processes. Second, subjective judgments might interfere with the processes. Indeed, both the internal self-examination undergone by institutions / programs and the external reviewing processes made by recognized accrediting bodies are prone to errors and subjective biases as they are largely based on rules of thumb human judgments – despite the presence of standards. In this paper, we describe a set of computational technologies to address these issues. Emphasis is made on technologies spanning (crude) data, information, refined information including decision support, ultimately leading to the most refined and expensive piece of information, *i.e.*, knowledge and its discovery in large and diversified databases over the Web, based on cloud computing solutions. A human-machine interactive knowledge-based learning control system for A3-HE is our far-reaching goal. However, the A3-HE processes are too complex to be addressed by computerized systems alone. As a result, scaling up to real-life applications still require much time to reach tangible implementations.

Keywords— Academic accreditation, Database management system (DBMS), Information system (IS), Decision support system (DSS), Knowledge base control system (KBSC), Data mining, Web technologies.

I. INTRODUCTION

Academic accreditation and assessment in Higher Education (A3-HE) formally acknowledges that an institution or program is following recognized quality standards. It is a complex, lengthy and tedious set of intertwined processes involving people at different levels of responsibility within the community, costly machines, various artifacts, and more often than not, heavy clerk work and expensive resources in terms of time, money and other unquantifiable elements. In most countries of the world, academic accreditation processes are

being conducted by government organizations such as a Ministry of Education / Higher Education or non-governmental non-profit agencies [4], [24]. In order to address the issue of heavy paperwork and subjectivity during the periodic self-examination, or during the reviewing process or at the final test for accreditation, as reported in [11], we define some computational technologies that are shown to help. There are indeed many ways in which computational science and technologies can be applied in academic accreditation processes enhancements. We describe here the most important components and show how computers can be useful in this domain, although it must be also stressed that the A3-HE problem is too complex to be addressed by computerized machines alone. These latter can play a significant role towards alleviating the encountered multi-faceted issues, leading to an acceptable solution. Not pretending to delve into all the details of these highly complex and interdisciplinary issues, we suggest using a few entry points available to computational and control sciences. In a previous work, our aim was to contribute to a unified view of the processes involved in A3-HE for accrediting bodies around the world [11]. Now, we focus on computational technologies *per se*. These latter involve different levels of actions spanning all possible outcomes [27] ranging from gathering and exploitation of (crude) data, to the generation of information, to the obtainment of refined information including decision making, and ultimately to the most refined piece of information, *i.e.*, knowledge, its discovery and its subsequent usage for controlling the academic processes. The actual technology-oriented efforts only span minor processes within the A3-HE, *e.g.* [8], [13], [15] [17]. Hence this paper. In the next section, we give an overview of the importance of data leading to rational decisions in A3-HE information systems. Section 3 discusses the transition from knowledge to its discovery, including Web technologies. Section 4 is devoted to advanced networking including cloud computing and its interplay with Open Networking Foundation (ONF). Section 5 discusses the knowledge-based control learning system (KBLC) as an ultimate architecture with the possible impacts the proposed technologies are thought to induce on future A3-HE. The paper ends with a conclusion summing up the main results and pointing towards some potential future developments.

Manuscript received March 26, 2011.

A. Hamdi-Cherif is with Computer College, Qassim University, PO Box 6688, 51452 Buraydah, Saudi Arabia. Tel. +96663800050, ext. 3301. His permanent address is with Faculty of Engineering, Computer Science Department, Université Ferhat Abbas, Setif (UFAS) 19000, Algeria. email: shrief@qu.edu.sa, elhamdi62@gmail.com

II. BASIC TECHNOLOGIES

A. Related Works

There exist automated tools for similar settings such as university rankings delivered by the Academic Ranking of World Universities ARWU [<http://www.arwu.org>], or Times Higher Education University Rankings (THE [<http://www.timeshighereducation.co.uk/>], or Webometrics Ranking of World Universities (WRWU, [<http://www.webometrics.info/>]), to cite but a few. There are also automated tools certification and assessment for some professions [5]. In Industry, surveys confirm that certification and accreditation (C&A) approaches in private and government sectors differ widely. However, they now share a general trend towards automated processes replacing costly manual approaches. In the US Department of Defense (DoD), the Defense Information Technology Systems Certification and Accreditation Process (DITSCAP) and the newly-established Department of Defense Information Assurance Certification and Accreditation Process (DIACAP) guides C&A and specialized information systems. A study has been undertaken that provides an analysis of inflexible and flexible design approach costs of both automated and the currently employed C&A manual processes prescribed by policy. The study deals with the Global Positioning System Joint Program Office (GPS JPO) [7]. These efforts, although too specific can provide useful guides for the development of corresponding future A3-HE computational frameworks. The main issue to date is that there is no such counterpart for A3-HE.

B. From Data Upwards

We stress the stratification of A3-HE processes into levels. This stratification motivates for the introduction of incrementally-sophisticated technologies. The proposed solution to the issues addressed in A3-HE settings is to be considered under six hierarchical and complementary levels, namely data base management system (DBMS), information system (IS), decision support system (DSS), knowledge-based system (KBS), data mining (DM) and knowledge base control system (KBCS). The main aim is to integrate, within A3-HE, learning methods, encompassed by the knowledge base learning control system (KBLCs) framework.

C. Overview of Ad Hoc Technologies

The knowledge-based learning control system (KBLCs) layers are depicted in Figure 1. The proposed hierarchical study spans gathering and management of data, generating information, producing knowledge, discovering knowledge, then controlling A3-HE processes using human-machine interaction, proactively. In order to address the A3-HE processes from a computational perspective, we adopt a bottom-up approach, *i.e.* starting from the simplest to the most complex technologies. Because we are interested in architectures and not in specific implementation, some of the proposed technologies, considered here, are still in their early infancy and did not reach sufficient maturation for application in simple contexts, let alone in A3-HE settings. The driving technology in the present approach is knowledge base control system (KBCS), considered as the core of knowledge base

learning control system (KBLCs). There exist several feedback loops. The main ones are the outer feedback based on the interaction with Web technologies, including pervasiveness, based on mobile-like technologies and cloud computing. Machine learning processes (ML) and common sense acquisition constitute the middle and inner feedback loops, respectively.

The results given by these technologies integration are meant to be used by all parties, *i.e.* by institutions or programs seeking A3-HE status, by decision makers at different levels of accountability including government policy and legislature makers.

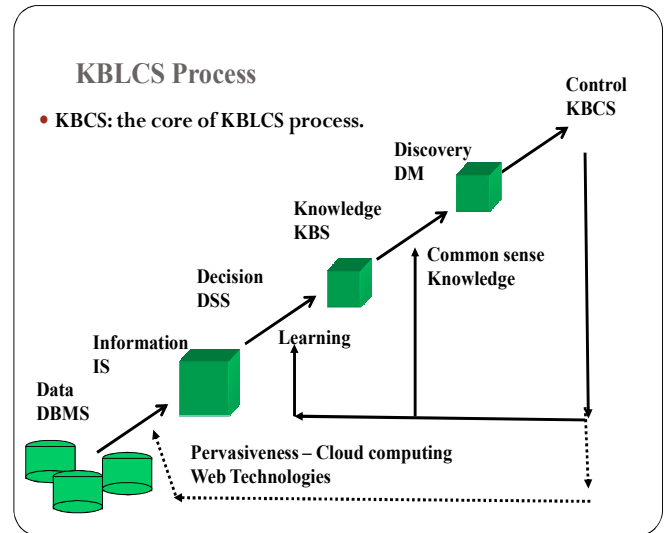


Figure 1 Knowledge-based learning control system process

D. Database Management System (DBMS)

Any A3-HE project or implementation has to start with the most basic element *i.e.* crude data. In any A3-HE context, data is very heterogeneous. For instance, at the most basic level, crude data might consist of number of students per year in a given program who have completed/dropped courses, with their grades, courses with major failure rates, faculty completing course schedules as planned, the number of highly qualified faculty wishing to join/leave the course/institution, the impact of the program on employers, to cite but a few scattered pieces of data. In most cases, data has to address at least curricula contents, students, faculty, staff, and employers. As the volume of data grows, it becomes necessary to use a database management system (DBMS), *i.e.*, a coherent computational environment that controls the creation, use, and the maintenance of integrated collection of different data records and files known as databases. A DBMS allows end-users and other software to store and retrieve data in a friendly and structured way. Indeed, as data becomes more and more diversified, the techniques used to design and implement DBMSs have become more sophisticated to cater for large-scale heterogeneous data and noise [6]. Instead of having to write computer programs to extract information, most systems offer the possibility of using special query language or fourth-generation programming language (4GLs) along with other

application development features. Traditional DBMSs whether they are relational, object-based, network, hierarchical, or multidimensional, are considered as intelligence-free programs. By intelligence-free program, we mean a program that does not infer new conclusions not explicitly stored in the computational environment.

E. Information System and Decision Support System

A series of technical advances in recent years has increased the amount of data that decision makers can record about different characteristics of any academic enterprise and its by-products such as students, research contributions, and community involvement or impact on it. It has to be noticed that any academic institution is literally flooded in a deluge of data. This data is, of course, vital for advancing our appraisal of the enterprise. Despite the fact that DBMSs are unavoidable in any A3-HE setting, they are far from being sufficient. Indeed, on top of DBMS technology, we need a more structured information as a form of processed data. This naturally leads to the use of information system (IS) [18]. On the other hand, there is no proper decision without information. Because information is becoming more and more sophisticated, proper decisions are becoming more and more difficult to take. This issue is handled through decision support system (DSS) For instance, if we consider an institution or a program as an enterprise evolving in a competitive marketplace, then it becomes obvious that it has to consider constantly-changing external parameters to stay alive and competitive. For that, information system (IS) and decision support system (DSS) technologies are vital.

1) Information System (IS)

Let us note at the outset that information system (IS) is not to be confused with its own discipline - information systems. In a very broad sense, the term information system (IS) refers to the interaction between people, algorithmic processes, data and technology. The term refers not only to the information and communication technology (ICT) within an organization, but also to the way in which people interact with this technology in support of the organization's business processes [18]. As a result, IS involves both the infostructure and the infrastructure that supports it [9].

2) Decision support system (DSS)

When exactly to take a given decision within the institution or program? What precise items are to be changed within the institution/program? What curriculum to adopt to face fierce competition and market trends? These issues are addressed within a decision support system (DSS) framework. A typical DSS consists of three principal components: a data warehouse server, online analytical processing (OLAP) and data mining tools, and back-end tools for populating the data warehouse. The data warehouses contain data consolidated from several operational databases and tend to be orders of magnitude much larger than operational databases. Typically, the data warehouse is maintained separately from the organization's operational databases because analytical applications' functional and performance requirements are quite different from those of operational databases. Data warehouses exist

principally for decision support applications and provide the historical, summarized, and consolidated data more appropriate. Due to the fact that data warehouse construction is a complex process that usually takes many years, perhaps decades, we can instead build data marts, which contain information for specific departmental subsets [3]. For example, an alumni follow-up data mart may include only type of degree, time of employment after graduation, and the employer. Several departmental data marts can coexist with the main data warehouse and provide a partial view of the warehouse contents.

3) Example of DSS/A3-HE interaction

As an example of DSS used for accreditation, we can cite the AEFIS™ (Academic Evaluation, Feedback, and Intervention System <http://www.goaefis.com/about-aeffis/faqs/>). AEFIS is a web-based assessment management solution and Instructional Decision Support System (IDSS) that supports personalized learning by delivering real-time information to administrators, faculty, students, and alumni and industry to enhance curriculum development and help in the accreditation processes. The core of AEFIS™ 3.0 is the so-called Instructional Decision Support System™ (IDSS™). It is a web-based system that gathers information to help instructors make educational decisions based on student data to improve instructional practices. However, AEFIS™ has no machine learning layer.

III. FROM KNOWLEDGE USE TO ITS DISCOVERY

This section gives the basic outlines for addressing two fundamental issues. First, it investigates the contribution of data mining technologies to A3-HE for the discovery of hidden patterns in data. For that, we describe the computational contributions to A3-HE and show how it can help in developing increasingly-complex tools for addressing knowledge discovery and issues. Second, we consider control issues within the framework of knowledge base system (KBS) and machine learning frameworks, ending with knowledge base learning control system (KBCLS) approach.

A. Knowledge-Based System (KBS)

Knowledge can be considered as the one of the most important asset in our modern society and particularly in academia. It is omnipresent in all endeavors of computing practice: from the rise of knowledge management to the Semantic Web and from the new blog culture to the knowledge economy. This penetration has made proper knowledge base systems (KBSs) and their engineering processes one of the most requested features. A KBS is a specialized software that uses artificial intelligence (AI) methods such as expert system in problem solving processes. It is composed of a rule base, a fact base and an inference engine. The most precious part of any KBS is its of expert knowledge embodying human knowledge, while a typical inference engine can be obtained freely from the Internet, e.g., the "C Language Integrated Production System" (CLIPS <http://clipsrules.sourceforge.net/>). A KBS has also couplings and is designed to facilitate

knowledge retrieval in response to specific queries, or to transfer expertise from one domain of knowledge to another.

B. Data mining

1) Defining Data Mining Field

Data without the pertinent knowledge that conveys it might be useless. Knowledge can be seen as the patterns or characteristics of the data. Raw data is sometimes meaningless because what we want is the knowledge hidden in the data and not the data *as such*. That is why a new technology has emerged in the mid 1990's to deal with the discovery of knowledge from data. It is called knowledge discovery in databases (KDD) or simply data mining (DM). Uncovering hidden information is the fundamental goal of data mining. A distinctive aspect of actual data mining systems is their extensive use of the Web and the manipulation of diversified data. In A3-HE, data might be constituted of student or faculty retention rate over a given period of time within a program or institution. The reasons behind these retentions rates constitute the knowledge extracted from DBMSs using specialized methods such as classification, clustering or association [12].

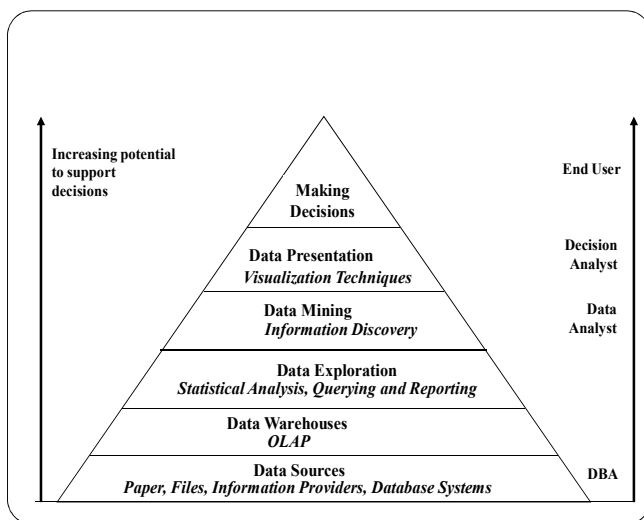


Figure 2 Data mining field overview

2) Data Mining Process And Tasks

(i) Process

Data mining process is based on the following steps: data collection, data preprocessing, data mining proper, information interpretation, and visualization.

(ii) Tasks

The data mining tasks are categorized as follows.

- *Classification* decides the class/group for each data sample. For example, students withdrawing from a given program can be classified based on their secondary schools.
- *Clustering* is to regroup similar data into a finite set of separate clusters/categories. This task is also referred to as *segmentation*. In machine learning, it requires *unsupervised learning* i.e. that the clusters are not known in advance.

- *Association, link analysis* or *affinity analysis* is to tell whether a set of data depends on the rest. An association rule can be written as $A \rightarrow B$ where both A and B are a data set.

- *Summarization* or *characterization* is a simple description of a large data set. It is desirable that representative information of a data set can be obtained, so that we can easily have a general view of the data set.

- *Text mining* is used when the data to be mined are text instead of numerical data. It originated from information retrieval (IR) of the library science. Keywords are used to find related documents from a document database. For more advanced applications of text mining, classification, clustering, and association techniques are utilized. Text mining can facilitate the re-examination of the literature to link the facts that were already known [16]. In A3-HE, text mining can be used in the analysis of reports written in the self-examination by institutions or by accrediting bodies.

3) Examples of Commercial Data Mining Systems

The data mining technology is now well established and several major companies are offering various products and solutions.

(i) IBM SPSS Modeler

IBM® SPSS® Modeler is a data and text analytics workbench that helps in the building of predictive models without programming. It has been used in modeling to improve business outcomes in areas like customer relationship management (CRM), marketing, resource planning, fraud and risk mitigation, medical research, law enforcement and national security. Its main characteristic are:

- Discover patterns and trends in structured or unstructured data by using visual interface supported by advanced analytics. It allows to model outcomes and understand what factors influence them.
- Visualize the data mining process using graphical interface. From this interface, the user can access both structured (numbers and dates) and unstructured (text) from a variety of sources, such as operational databases, survey data, files.

[<http://www.spss.com/software/modeler>]

- (ii) *Microsoft SQL Server 2008 R2*, is a DBMS that incorporates manageability, development productivity, and business intelligence. The Server incorporates multiple data mining functions, including association mining, classification (using decision tree learning, naive Bayes, and neural networks algorithms), regression trees, sequence clustering, and times series analysis. Moreover, the Server also supports the integration of algorithms developed by third-party vendors and application users.

[<http://www.microsoft.com/sqlserver/en/us/>]

(iv) Oracle Data Mining (ODM)

Oracle Data Mining (ODM) is an option to Oracle Database 11g Enterprise Edition. It enables to build and deploy applications that deliver predictive analytics and

new insights. Applications can be built on ODM's SQL APIs that automatically mine Oracle data and deploy results in real-time-throughout the enterprise. Because the data, models and results remain in the Oracle Database, data movement is eliminated, security is maximized and information latency is minimized. Oracle Data Mining models can be included in SQL queries and embedded in applications to offer improved business intelligence. Data analysts can access Oracle data using the optional ODM's GUI and explore data to find patterns, relationships, and hidden insights. ODM provides a collection of in-database data mining algorithms that solve a wide range of business problems. Anyone who can access data stored in an Oracle Database can access ODM results-predictions, recommendations, and discoveries using SQL-based query and reporting tools.

<http://www.oracle.com/technetwork/database/options/odm/index.html>

C. A3-HE Issues And Data Mining

In A3-HE, knowledge discovery is of utmost importance. Computer Science program coordinators and curriculum committees are confronted to a dilemma. On the one hand, the organization and content of a scientific discipline should not depend on the whims of employers alone. On the other hand, the new technological environment is under the pressure of globalization and contains forces beyond the control of governments, let alone curriculum committees, and marketplace effects are forcing academic environments in terms of enrollment, faculty positions and external research funding [14], [19]. Data mining is a framework that helps in alleviating this dilemma by making a rapprochement between academic and professional aspects of the issue.

D. Web technologies

1) From Web 1.0 to Web 2.0 Technologies

All the previous issues dictate the urge for further integration of multiple technologies such as Web 1.0 / Web 2.0 technologies. There was a natural transition to Web 2.0 after the so-called dot-com crash in Fall 2001. The main Web 2.0 technologies include wikis, blogs, RSS filters, folksonomies, mashups, podcasts, crowdsourcing, social networks, and virtual worlds [2]. A summarized comparison between Web 1.0 and Web 2.0 technologies relevant to A3-HE is given in the Table 2 [<http://oreilly.com/web2/archive/what-is-web-20.html>], [20].

For example, use Web 2.0 technology in A3-HE to identify degree and accreditation mills through social networks and blogs, to cite but a few. For the dangers of degree accreditation mills, refer to [11].

Web 1.0	Web 2.0
Britannica Online	Wikipedia
personal websites	blogging
domain name speculation	search engine optimization
page views	cost per click
screen scraping	Web services
publishing	participation
content management systems	wikis
directories (taxonomy)	tagging ("folksonomy")
stickiness	syndication

Table 1 - Web 1.0 vs. Web 2.0

2) Example of DBMS/Web 1.0 interaction

At a higher level of the A3-HE processes, data / information might consist of the codification of good practices as proposed by institutions for discussion / amendment / adoption and might address issues related to market trends, constraints imposed by legislatures, globalization and others.

As an interaction between DBMS and Web 1.0 technologies, we can cite the so-called "good practice database" that was developed and is now managed by the Australian Universities Quality Agency (AUQA). The Database is intended to gather information about good practices in the Australasian Higher Education sector publicly available. It is intended that such information will serve as a general guide to the sector to assist it in its efforts to improve. The Database is not intended to present standards for the sector nor will AUQA audit an institution/agency against entries in the Database.

A new batch of Good Practices was launched in September 2010, including entries from seven Higher Education institutions. Submissions covered a range of topics and now with over 200 entries, the database provides an important tool to benchmark institutional practices. Invitations for batch 14 have been distributed and accepted with 19 good practices launched on 10 February 2011 and include a number of contributions which will assist the sector to compare practices among its institutions. (AUQA <http://www.auqa.edu.au/gp/terms.php>).

3) Example Of DSS/Web 2.0 Interaction

Consider the following decision-based issue. Is a given program/course worthwhile teaching? This issue is reduced to the following: What Web 2.0 technologies say? An important part of the answer can be supplied by wikis, blogs, crowdsourcing, social networks, involving relevant community members such as faculty, employers, students, professionals, parents, and the general public. Web technologies can help in

shaping the final decision by encouraging, accepting, amending, or discouraging such a program/course.

IV. ADVENT OF CLOUD COMPUTING

A. Cloud Computing Characteristics

Another major technology that is susceptible to address stringent A3-HE issues is cloud computing. This novel technology provides a way to develop applications in a virtual environment where computing capacity, bandwidth, storage, security, and reliability aren't issues because users don't need to install the software on their own system. In a virtual computing environment, users can develop, deploy, and manage applications, paying only for the time and capacity used while scaling up or down to accommodate changing needs or external requirements. Cloud computing has the following characteristics:

(i) Distributed applications

The storing of, and accessing to applications and computer data often through a Web browser rather than through specific software installed on every personal computer or server.

(ii) Pervasiveness

Internet-based computing whereby information, IT resources, and software applications are provided to computers and mobile devices on-demand.

(iii) Service-oriented application (SOA)

Using the Internet to access Web-based applications, Web services, and IT infrastructure as a service.

The applications (e.g. the email service Gmail [<http://www.gmail.com>]) live on a remote "cloud server" that anyone can access through the browser like any other website. Rather than every user having their own copy to install maintain and upgrade, a single cloud application at a single location is shared by millions of users, or more, on a daily basis (<http://www.cloudcomputingdefined.com/learn/>). Although cloud computing is a relatively new technology, some solutions have appeared recently. For instance, IBM™ has partnered with Amazon™ Web Services to give users access to IBM software products in the Amazon™ Elastic Compute Cloud (EC2) virtual environment. This is product-level code with all features and options enabled. <http://www.ibm.com/developerworks/cloud/ec2.html>

B. From Cloud to Open Networking Foundation (ONF)

1) Open Networking Foundation (ONF)

On 21st of March 2011, six companies that own and operate some of the largest networks in the world announced, the formation of the Open Networking Foundation (ONF), a nonprofit organization dedicated to promoting a new cloud-focused initiative approach to networking called Software-Defined Networking (SDN). The founding members of ONF are Deutsche Telekom™, Facebook™, Google™, Microsoft™, Verizon™, and Yahoo!™. Joining these six

founding companies in creating ONF are 17 member companies, including major equipment vendors, networking and virtualization software suppliers, and chip technology providers. ONF, through SDN, allows owners and operators of networks to control and manage their networks to best serve their needs. ONF's first priority is to develop and use the so-called OpenFlow protocol [<http://www.openflow.org/>], [21]. The main functionalities of OpenFlow is to seek to increase network functionality while lowering the cost associated with operating networks through simplified hardware and network management [<http://www.opennetworkingfoundation.org/>].

2) Example of Cloud Computing DBMS Environments

Microsoft SQL Azure Database™ represents an interaction between traditional DBMS and cloud technology. It allows to build and extend cloud-based relational database service built on SQL Server. It allows to gain access to a available, scalable, multitenant database service hosted by Microsoft™. The SQL Azure™ features are:

- No software to install or hardware to set up.
- Availability and fault tolerance.
- Provisioning and deployment.
- Application development.
- Scalable databases based on business needs.
- Multitenant support.
- Integration with Microsoft Visual Studio™.
- Transact-SQL development support.
- Familiar, relational database model.
- ADO.NET, ODBC, JDBC, and PHP support.

<http://www.microsoft.com/sqlserver/en/us/editions/azure.aspx>

C. Cloud Computing / ONF And A3-HE

A3-HE large databases surely contain interesting patterns waiting to be discovered. Data mining, as expanded above, is a suitable technology helping in this discovery. On the other hand, cloud computing, or its newer expression ONF, through interaction with human experts can lead to novel theories about academic processes and their modeling/control entailing the possibly of designing *ad hoc* tools for addressing some of the A3-HE pending issues.

V. A CONTROL APPROACH TO A3-HE

Both human-made and pure human systems exhibit collective behaviors amongst individuals in which some controllers seek a certain form of equilibrium. Based on this idea, we propose a framework whereby the A3-HE is considered as a control system.

A. Control Theory Viewpoint

1) A3-HE As A Process To Be Controlled

In control theory, a controller is a device which monitors and affects the operational conditions of a given dynamical system. The operational conditions are the output variables of the system which can be affected by automatically adjusting certain control variables. In our situation, the A3-HE processes

are equipped with a controller for measuring the output variables, *i.e.* A3-HE outcomes which are feedback for corrective action. In this example, the assessment is the controller that delivers the appropriate values of the control variables to the program/institution system and therefore directs the activities of the A3-HE processes. The actions taken for improving the outcomes are the desired outcomes, or setpoints. In agent-based terminology, the A3-HE is viewed as a multiagent entity producing actions (actual outcomes) under the influence of inputs (objectives, specifications) within a certain environment (community) and in which a set of human expertise is coded so that the difference between the actual outcomes and desired ones are kept to a minimum. Benchmarking helps in the minimization process.

2) Adaptive Control and Software Engineering Interaction

An important part of A3-HE is the development of software. Recently an integration of control methods in software has been set out. In this context, efforts are deployed in order to apply some control theory methods such as adaptive control to software engineering at large. Diverse research topics have been integrated into a coherent subject matter. As software systems grow in complexity, some of the central challenges include their ability to self-manage and adapt at run time, responding to changing user needs and environments, faults, and vulnerabilities. Control theory approaches provide some of the answers to these challenges. These topics address issues such as requirements engineering, software development processes, pervasive and autonomic computing, service-oriented architectures (SOA), on-line adaptation of software behavior, testing and Quality of Service (QoS) control. These efforts better explain how the adaptive control approach can be applied to improve the quality of software systems as well as future theoretical and experimental challenges for researchers in this area [26].

B. KBCLS

As far as knowledge is concerned, the hierarchical control processes consider two layers. The first is knowledge base control system (KBCS) and the second is knowledge base learning control system (KBLCS). Some of the prominent applications is pattern discovery. Patterns often concern the categories to which situations belong. For example, is a loan for a given student creditworthy or not? Will a certain segment of students population ignore an elective course or take it? Will a course taught by a given instructor yielded high, medium, or low failure rate on a given batch of students? All these issues are handled by machine learning procedures and incorporated in a KBCLS.

1) KBCS For A3-HE

Intelligence-based or AI-based programs are the basic characteristics of KBCS. Indeed, for complexity reasons, 'brute force' use of procedural heuristics-free algorithms is useless. Clearly, because of the nature of any A3-HE context, a more 'intelligent', *i.e.* heuristics-based approach is required to solve problems since human expertise codification is

necessary. KBCS principles and tools enhance the A3-HE existing solutions because it incorporates human expertise in the form of code. As to learning control, it is traditionally linked with machine learning technologies such as fuzzy systems, neural networks, genetic algorithms, among others [28].

2) Machine Learning

Machine learning is an adaptive process that makes computers improve from experience, by example, by analogy, or otherwise. Learning capabilities are therefore essential for automatically improving the performance of a computational system over time on the basis of previous history. A basic learning model typically consists of the following four components:

- learning element, responsible for improving its performance,
- performance element, or decision support system (DSS) responsible for the choice of actions to be taken, whether these actions are decisions or controls,
- critical element, which tells the learning element whether the criteria are met within some critical boundaries,
- problem generator, responsible for suggesting actions that could lead to new or informative experiences [23].

The importance of learning will grow as the external environment continues to generate and integrate large quantities of diversified data.

3) Supervised vs. unsupervised learning

As far as relevant machine learning is concerned, there are basically two categories of learning schemes, namely supervised learning and unsupervised learning. Supervised learning learns the data with a known answer at hand. From the control standpoint, this is a typical feedback control system. The parameters are modified according to the difference of the actual output and the desired known output, or expected answer. Classification falls into this category. On the other hand, unsupervised learning learns without any knowledge of the outcome. Clustering belongs to this category. It finds data with similar attributes and aggregates them in the same cluster [1].

The main familiar machine learning methods such as decision trees learning (DTL), and support vector machines (SVM) have proved very useful in addressing both classification and clustering problems. But machine learning techniques usually handles relatively small data sets because the learning procedure is normally very time-consuming. To apply the techniques to data mining tasks, the problem with handling large data sets must be overcome [16].

4) Some Products for Machine Learning

Many machine learning commercial and free products and solutions are now available.

a) *Commercial Products*

(i) *CART*[®]: is data mining software developed by Salford Systems. *CART*[®], is a decision tree that automatically sifts large, complex databases, searching for and isolating significant patterns and relationships. This discovered knowledge is then used to generate predictive models for applications such as finding best prospects and customers, targeted marketing, detecting credit card fraud, and managing credit risk. The most recent 2008 release, *CART* 6.0, includes modeling automation technology supposed to accelerate the process of generating accurate and robust models for deployment in core business functions. [<http://salford-systems.com/cart.php>].

(ii) *See5*, for Windows platform, developed by RuleQuest and its Unix counterpart *C5.0* are decision trees or sets of if-then rules data mining tools for discovering patterns that delineate categories, assembling them into classifiers, and using them to make predictions. *See5/C5.0* has been designed to analyze databases containing hundreds of thousands of records and hundreds of numeric, time, date, or nominal fields. *See5/C5.0* also takes advantage of computers with up to eight cores in one or more CPUs (including Intel Hyper-Threading) to speed up the analysis. *See5/C5.0* classifiers are expressed as decision trees, forms that are generally easier to understand than other machine learning methods such as neural networks, for instance. *See5/C5.0* developers provide C source code so that classifiers constructed by *See5/C5.0* can be embedded in the organization's own systems. [<http://www.rulequest.com/see5-info.html>]

b) *Open-Source Free Products Over The Internet*

(i) *Weka*, [10] is a collection of data mining tasks and machine learning algorithms. The algorithms can either be applied directly to a dataset or called from any Java code. Furthermore, *Weka* contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also open for developing new machine learning schemes. *Weka* is open source software issued under the GNU General Public License, [<http://www.cs.waikato.ac.nz/ml/weka/>]. A book has been recently published teaming machine learning algorithms with applications using *Weka* [25].

(ii) *Tanagra*

Tanagra [22] is a free data mining software for academic and research purposes. It proposes several data mining methods for exploratory data analysis, machine learning, statistical learning, and database. As a successor of *SIPINA*, *Tanagra* implements various supervised learning algorithms, especially an interactive and visual construction of decision trees along with some other paradigms such as clustering, factorial analysis, parametric and nonparametric statistics, association rule, feature selection and construction algorithms. *Tanagra* is

an open source project as researchers can access to the source code, and add their own algorithms, as far as they agree and conform to the software distribution license. Because *Tanagra* is an experimental platform, it does not include, presently, what makes all the strength of the commercial software in this domain : a wide set of data sources, direct access to data warehouses and databases, data cleansing, and interactive utilization. [<http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html>]

C. *KBLCS Impacts On A3-HE*

We believe that the integration of previously-described technologies will advance our knowledge of A3-HE processes based on the most powerful technological tools available to computer and control systems scientists, entailing a better modeling and control of the A3-HE processes. The impacts on many fields of research are expected to be important, not only on academic processes as such but also on the community as a whole. We expect impacts on the following fields of research and technology.

- i. *DBMS/Data Mining*: More structured organization of data for efficient response to queries. For instance, to develop ways to index or otherwise preprocess the data to make searches more efficient.
- ii. *IS/DSS*: To provide intelligent interfaces that will assist the user in framing and executing queries.
- iii. *KBS/KBCS/KBLCS*: To further formalize A3-HE within knowledge base systems (KBSs) and engineering frameworks with machine learning capabilities.
- iv. *Web Technologies*: To further integrate A3-HE within actual Web/cloud computing technologies and solutions.

VI. CONCLUSION

Academic accreditation is above all a recognition by the community that a given academic institution meets some expected quality standards and generates recognized quality products such as students, research production, and services. We have shown how to proactively increase the collaboration between human decisions and computerized A3-HE processes. On top of the multitude of methods and tools, it remains highly expected that the technologies reported here will uncover even more useful hidden structures in A3-HE processes. In addition to actual query methods now available, however intelligent these might be, future public A3-HE databases have to include an array of “what-if” simulation scenarios capable of producing highly anthropomorphic behavior through the knowledge base learning control system framework. Further integration of these technologies into A3-HE will remain indeed a challenging task for years to come.

VII. ACKNOWLEDGMENT

The author would like to acknowledge moral support from Qassim University and particularly from Computer College during the three-year long implementation of the academic accreditation processes. Help and support in the form of seminars have been provided by the National Commission for Academic Accreditation and Assessment (NCAAA) (<http://www.ncaaa.org.sa>), Riyadh, Saudi Arabia.

References

- [1] Alpaydin, E., "Introduction to Machine Learning", MIT Press, 2004.
- [2] Andriole, J.S., "Business Impact of Web 2.0 Technologies", *CACM*, 58(2):67-79, December 2010.
- [3] Chaudhuri, S., U. Dayal, V. Ganti "Database technology for decision support systems", *IEEE Computer*, pp. 48-55, December 2001.
- [4] Council for Higher Education Accreditation, "Recognition of accreditation organizations: a comparison of policy and practice of voluntary accreditation and the United States department of education", CHEA, January 1998.
http://www.chea.org/pdf/RecognitionWellman_Jan1998.pdf. Retrieved 2009-11-06.
- [5] The Chartered Institution of Water and Environmental Management (CIWEM), Version 7 – November 2009, <http://www.ciwem.org>.
- [6] Elmasri, R. and S.B. Navathe, "Fundamentals of Database Systems", Pearson Education, 6th Edition, April 2010.
- [7] Falkenthal, D. "Flexible Design Analysis of Automated GPS Certification and Accreditation Tools", ardent.mit.edu/real_options/.../Falkenthal-GPS_AP-complete.pdf
- [8] Fuller, U., J.C. Little, B. Keim, C. Riedesel, D. Fitch, S. White "Perspectives on developing and assessing professional values in computing". *SIGCSE Bulletin*, 41(4), January 2010.
- [9] Gowan, A. B. MacDonald, H. Reichgelt, "A configurable assessment information system", *SIGITE '06 – Proc. of the 7th Conf. on IT Education*, ACM, October 2006.
- [10] Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten; "The WEKA Data Mining Software: An Update", *SIGKDD Explorations*, 11(1):10-18, 2009.
- [11] Hamdi-Cherif, A., "A computational framework for academic accreditation and assessment in higher education (A3-HE) – Part 1 academic processes", Invited Conference / Book Chapter In: *Recent Advances in Artificial Intelligence, Knowledge Engineering and Database (AIKED'10)*, Cambridge, UK, February 22-25, 2011, pp. 403-410, <http://www.worldses.org/online/2011.htm>
- [12] Han, J., M. Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann Publishers, 2006.
- [13] Harkiolakis, N., "An automated approach to assessment management of higher education institutions", In *ICCOMP'07 – Proc. of the 11th WSEAS Int. Conf. on Computers*, July 2007.
- [14] Hoganson, K., "Computer science curricula in a global competitive environment", *J. of Computing Sciences in Colleges*, 20(1), October 2004.
- [15] Hou Yung-Chi, A., 'Student learning outcomes and quality assurance of higher education in Taiwan', *Proc. of Conf. on QA and Student Learning Outcomes of HE in Asia-Pacific Region*, Taiwan, 2009.
- [16] Romero, C., S. Ventura, M. Pechenizkiy, R. Baker, (Eds.), "Handbook of Educational Data Mining", Series: Chapman & Hall/CRC Data Mining and Knowledge Discovery Series, October 2010.
- [17] Iyer, V. G., N. E. Mastorakis, "Some innovative practices on quantity and quality of engineering education and research", *EDUCATION'09 - Proc. of the 6th WSEAS Int. Conf. on Eng. Education*, July 2009.
- [18] Kroenke, D M., "Experiencing MIS" Prentice-Hall, Upper Saddle River, NJ, 2008.
- [19] Lenn, M.P. "Global trends in quality assurance in higher education", *World Education News & Reviews*, 5(2), Spring 1992, pages 1 and 21.
- [20] O'Reilly, T., "What is Web 2.0? - Design patterns and business models for the next generation of software", <http://oreilly.com/web2/archive/what-is-web-20.html>, 09/30/2005, retrieved January 2011.
- [21] Pfaff, B. *et al.*, "OpenFlow Switch Specification Version 1.1.0 Implemented (Wire Protocol 0x02)", February 28, 2011.
- [22] Rokotomalala, R. "TANAGRA: a free software for research and academic purposes", *Proceedings of EGC'2005, RNTI-E-3*, vol. 2, pp.697-702, 2005 (in French)
- [23] Russel, S., P. Norvig, "Artificial Intelligence – A Modern Approach", Prentice Hall in Artificial Intelligence, 2nd Edition, pp. 51-54, 2003.
- [24] Westerheijden, D. F., B. Stensaker, M. João Rosa (Editors) "Quality assurance in higher education: trends in regulation, translation and transformation", Springer, 2007.
- [25] Witten, I.H., E. Frank, M.A. Hall "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Third Edition, ISBN 978-0-12-374856-0, January 2011.
- [26] Wong, W.E., B. Cukic (Eds.), "Adaptive Control for Software Quality Improvement", *Series on Software Engineering and Knowledge Engineering*, World Scientific, To appear in Summer 2011.
- [27] Yaverbaum, G., D. Lidtke, H. Reichgelt, S. Zweben "Outcomes-based computing accreditation: program assessment", *SIGCSE'07 – Proc. of the 38th SIGCSE Tech. Sym. on Comp. Sci. Education*, March 2007
- [28] Zilouchian A., M. Jamshidi (Eds.). "Intelligent Control Systems Using Soft Computing Methodologies", CRC Press LLC, 2001.

Websites accessed as of March 2011

Examples of KBS

CLIPS: <http://clipsrules.sourceforge.net/>)

Example of DSS: <http://www.goaefis.com/about-aeffis/faqs/DBMS/Data Mining Tools/ Machine Learning Software>

<http://www.microsoft.com/sqlserver/en/us/>

<http://www.oracle.com/technetwork/database/options/odm/index.html>

CART: <http://salford-systems.com/cart.php>

See5/C5.0: <http://www.rulequest.com/see5-info.html>

Weka: <http://www.cs.waikato.ac.nz/ml/weka/>

Tanagra:

<http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html>

Cloud computing

(<http://www.cloudcomputingdefined.com/learn/>)

Examples of Cloud Computing Solutions

<http://www.ibm.com/developerworks/cloud/ec2.html>

http://www.chea.org/pdf/RecognitionWellman_Jan1998.pdf

<http://www.ciwem.org>

<http://www.opennetworkingfoundation.org/>

<http://www.openflow.org/>

<http://www.microsoft.com/sqlserver/en/us/editions/azure.aspx>

Examples of University Ranking Systems

ARWU - <http://www.arwu.org>

THE – <http://www.timeshighereducation.co.uk>

WRWU, <http://www.webometrics.info/>

Web 2.0

<http://oreilly.com/web2/archive/what-is-web-20.html>



Aboubekeur Hamdi-Cherif was born in Setif, Algeria. He received BSc (Honors) Electrical Engineering and MSc in Electronic Control Engineering both from Salford University, Manchester, England, in 1976 and 1978, respectively, and PhD degree in Computer Science from Université Pierre et Marie Curie Paris 6, (UPMC), Paris, France, in 1995. He is member of IEEE and ACM.

He worked as Control Engineer with Algerian Petroleum Company SONATRACH. He taught at Ecole Supérieure des Transmissions, Algiers, Algeria, at Université de Bretagne Occidentale, Brittany, France, at Ecole Supérieure Libre

des Sciences Commerciales Appliquées (ESLSCA), Paris, France, and at Ferhat Abbas University, Setif, (UFAS), Algeria. While at UFAS, he was Director of Postgraduate Studies and Scientific Research, reporting to the Vice-Rector. He is now with Computer College, Qassim University, Saudi Arabia, where he is Associate Professor and Deputy Head of Computer Science Department. He is in charge with the implementation of the academic accreditation processes with NCAAA. He supervised about 90 BSc student projects, and 10 master and doctoral students. His research interests mainly concern control, machine learning, and academic processes automation.

In 1998, he was awarded the best technological thesis prize by UFAS. Dr. Aboubekeur received financial support for his research projects from Université Pierre et Marie Curie Paris 6, Paris, France, from UFAS, from Qassim University, and from King Abdulaziz City for Science and Technology (KACST), Riyadh.