Abstract — Aspects of curriculum formation, supportive framework development and a scheme of delivery are introduced, using a web-semantic driven approach accompanied by a model of human perception, information delivery and assessment supportive schemes. Shown how semantic web research can help with monitoring of knowledge base, defining structural knowledge dependencies, detecting the change of knowledge state and estimating functional or structural deviations. The impact of Information and Communication Technology (ICT) and web-media on human perception and knowledge formation is analyzed to avoid use that might impair human development through brain/sensory/perception models. The visual, textual and audio channels of information delivery are analyzed together. Shown how web-semantic approach can be supported by modification of assessment process with better tuning to knowledge formation and student ability making assessment frameworks based on automated support for reliable and quality of knowledge delivery. Structure of further research is suggested.

Keywords - curriculum; web-semantic search; human perception; signal processing; assessment methods and tools

I. INTRODUCTION: CONCEPTS & OBJECTIVES

The high education curriculum has not utilized the speed and intensity of the technological and scientific revolution boosted by ICT. At the same time semantic Web research might help in an upgrade of the curriculum, making possible fully automated monitoring of the knowledge base. The next generation of curriculum development requires:

- Revision of the process of curriculum design including as information processing scheme; use of semantic web searches based on keywords combined with the current state of knowledge define a core; modification of teaching schemes; incorporation of system support for maintenance of core of knowledge using both: historical and current information, checking the structural knowledge dependencies and state transitions in a single model.

- Study of the impact of application of ICT on the human creativity, differencing two dependent models:
  a) Web/ICT based systems of knowledge delivery and consumption
  b) information delivery to brain in a practical and effective way.

Bullet b), in turn, requires an understanding of the structure of the brain in terms of information delivery and perception. There are three main channels of knowledge delivery: visual (imagery), textual and auditory; all must be considered together.

Load of new curriculum development and delivery should be estimated using performance/workload monitoring model and delivery control as form and cost of workload change for lecturers and students. Load should be monitored - we need to know how much innovation we can carry / cope without killing ourselves.

Clear, new web-based curriculum development should be supported by an automated process for certification of the results of education, the quality and quantity of knowledge and skills, assisted using web-based semi-automatic assessment scheme. This process needs to be 'self-tuning' on the level of student knowledge as well as to provide guaranteed coverage.
of the required scope of knowledge and significantly reducing cost. All these points are inter-related and require co-design and co-development, i.e. considered as one approach.

II. INTEGRATED APPROACH

This integration of theoretical formation of curriculum, a framework for development, a research of a process of delivery, models of the workloads and automatic supportive assessment scheme can be integrated if all these processes will be considered as an information processing system or model, including the human as a part of this process. A knowledge delivery through curriculum should consider properties of human brain: this helps to exercise an adjustment of speed of information delivery to student abilities. Figure 1 below presents a bird-view on our approach followed by further explanation.

![Figure 1. WEDUCA approach bird view](image)

III. WEB SEMANTIC FOR THEORY OF CURRICULUM

Making curriculum enhancement is possible using a web-semantic approach considering a knowledge delivery system by analogy with [1] and [2]. That System might be analyzed through closely connected functions: Definitive, Characteristic and Predictive. A Definitive Function (DF), nominates terms, DF answers the question “What is it”? The second function describes the interrelations between Definitions and thus characterizes them. It is called the Characteristic Function (CF). CF answers the question “how are these definitions connected”? The Predictive Function (PF), in turn, answers the question “What if”? DF and CF and their application are essential elements of knowledge. Application of definitions and characterization of their interrelations enables predict behavior of a learned element.

The evolution of a subject area can be described through a joint Glossary \{DF\} of terms, as well as an essential Glossary \{DFe\}; through their change we can observe features such as merging, moving and separation of a curriculum core. A moving core occurs when curriculum shows growth when using some descriptive elements (DFi) and decrease in the use of others. This process of core motion can lead to the separation of a discipline into several new disciplines. This separation should be detected early before it becomes significant, as the discipline can loose its predictive function (PF) and thus become obsolete.

The other process of changing the core is merging of several cores of various disciplines. This process can be caused by significant discoveries and inventions, practical demand or technological revolution or an accelerating fusion of several disciplines where the sum is greater than the parts.

Fusion becomes visible by the growth of the Predictive Function - resulting in the applicability of a new discipline to society.

In order to make continuing process and establish theoretically useful concepts one has to develop a sort of engine to make the motion of knowledge visible – a knowledge seismometer that warns that the curriculum needs to be adapted to the motion of the knowledge landscape that is happening.

A set of initial terms/keywords is given defined manually, taken from the curriculum.

Collect text from the Internet with predefined keywords (formation of body of texts for various time slots).

Design context vectors that determine meaningful closeness and dependency between keywords at various moments of time and time slots (point and evolution dependency).

Describe keywords on a 2-dimensional chart.

Searching clusters of closely related keywords.

Searching of the hierarchical structure of cluster dependency (hierarchical clustering).

Develop a closed curve that includes clusters of keywords (this is a landscape version and change of form and size of this curve is landscape motion).

Those new terms that appeared inside this closed curve and thus became closer to the central keywords are candidate terms for inclusion in the curriculum.

Those keywords that are more distant from central keywords are candidate terms for exclusion from the curriculum. After deletion these candidate terms the curve changes its form and thus motion of the curriculum is defined.

When structures of clusters become virtually disconnected clusters are too remote then it indicates a process of core separation - formation of a core emerge for new discipline. Monitoring of evolution of core is possible algorithmically. Testing and initialization of proposed algorithm can be approached using as a starting point three ACM curriculums for computer science of 2001, 2005 and 2012 respectively:

- Define keywords from curriculum 2001, search new candidate terms using the proposed algorithm.
- Define keywords from curriculum 2005, compare with previous results, search new candidate terms.
- Define keywords from curriculum 2012, compare with previous results and define new terms and topics to be included into curriculum.

The more new candidate terms are included into new curriculum the more visible core motion and landscape change we have – indicating changes required into next generation of curriculum. As a trivial example of checking of conceptual and semantic links between terms and core migration in computer science area using the web using Multidimensional scaling (performed by a tool R script) one can see initial mapping of...
cluster of knowledge, Figure 3. Checking of moving for term “friend” might be helpful for formation of landscape change as a whole.

A. Visualization of dynamics

The Web-enhanced Design of University Curricula (WEDUCA) approach represents the curriculum in the form of several clusters of terms that dynamically change over time in context, quantity and even meaning. In order to define, analyze and “identify” clusters there is a need to understand the measures of clustering between entries and terms thus measuring semantic connectivity and distance, usually called the “semantic similarity of terms”. The most modern approach to calculate this is based on the use of distributional semantic models [20], successfully used for word similarity, word clustering, automated thesaurus generation, word sense disambiguation, and query expansion. For WEDUCA purposes the first application of distributional semantic models comes from analysis and dynamic monitoring of curriculum evolution.

Distributional semantic models, variously known as vector spaces, semantic spaces, word spaces, corpus-based semantic models, all rely on some version of the distributional hypothesis, stating that semantic similarity between two words/terms can be modeled as a function of the degree of overlap among their linguistic contexts. Having web astronomical and constantly growing amount of textual material we can provide the basis for identifying linguistic contexts and semantic dependency as well as differences (of terms) evolving over time. This will enable the observation of curriculum migration with growing precision.

A subject map (or landscape) is a kind of visual representation of the important subjects and terms in the target domain and their mutual conceptual relationships and connectivity, on a 2-dimensional chart. The reduction of dimension and scale of the original semantic space allows the representation of the distances among terms on a subject map.

Semantic Vector Space enables visualization of super-scale volume of data, simply because distances between terms on the subject map depend on distribution of context vectors, which components depend on distribution of large amount of invisible on the map (landscape) word combinations. This way a relatively small number of visible terms on the map might depend on distribution within knowledge domain of millions of invisible word combinations.

This makes possible to create a compact map of a subject, aggregating information of ‘super-scale’ body of texts. Introduction of new information into the body of texts might substantially change distances and dependencies as a whole between different terms on the map, making visible the impact of new information. This also makes possible to detect and process an early identification of missed terms/keywords in the curriculum.

Thus maps of subject area, based on Semantic Vector Space model becomes essential and unavoidable instruments for monitoring of distances and differences of super-scale and large-scale bodies of texts and data.

Checking of moving for term “Friends” on Fig.3 might be helpful for observation of landscape change as a whole. Space position of terms on the map defines their semantic connection. Terms from Computer science area are inside enveloping curve that evolves permanently, together with modifications of dependencies between terms. External terms in relation to enveloping curve also depend on semantic connectivity. Thus the external term “Friends” is connected with the new cluster ”Social networks”. The term “Social networks” is relatively recent and therefore resides at the border of computer science area - enveloping curve. This kind of dependencies are visually detectable and can be analyzed using 2-dimensional chart. We can consider it as simplest test of work for our model.

IV. KNOWLEDGE DELIVERY ASPECTS

The development of two areas of knowledge formation and delivery: human perception and a theoretical analysis of ways to deliver knowledge efficiently should be included into design of curriculum. A problem of searching the core of knowledge has a long history. Surprisingly, we still follow a lot of ancient scientists, their results & methods, hardly developing our own and new. To quote Jonathan Swift [3]:

“ancient virtuosos were carrying a spirit in casks and barrels which was out of great assistance upon long life voyages.”

In the past people had time to think, to reflect, time to react. They had time to use their associative memory and brain. But this gentle approach happens less and less. We are now swamped by information the Internet, like children in a sweetshop. To cope with this information avalanche we have to introduce semi-automated ways to identify, track and adapt knowledge drifting. In other words use the web-semantics to extract knowledge from web and manage it is crucial. ICT driven gargantuan information noise thrown at us at an enormous rate and volume using for this all possible delivery channels and is overloading our brains - “aliqno modo essentiae adhaerere” (in some way affect the substance). An ever increasing range of technologies are delivering the information deluge and are unwittingly involved in degrading our ability to filter and follow the knowledge we desire from it. To counter balance workload on the brain efficiently we should consider both: brain organization and theories of control and signals.
A. Human perception aspect

Each student brain is capable of absorbing substantial amount of knowledge. The knowledge delivery process can use any combination of channels of delivery: via text, via audio or visual. 'Text' should be separated from 'visual' and considered as special channel. All three processes of information absorption consume brain energy; all have a different level of information density and all need to be effectively used in a concert to achieve the best results. As an example, let’s look at the visual channel in a bit more detail and understand where this input is processed and how much energy the visual path of information delivery requires. The visual cortex deals with visual image sensing, while recognition and interpretation of “what we see” uses other segments of brain. So, to an extent our abilities to think while we concentrate on the screen are actually reduced or even blocked. This is only part of the problem, the perception of textual, audio and possibly other sensory information also may be needed to contribute to the recognition of knowledge but all must be correlated together.

![Diagram of human brain](image)

Figure 4. Visual recognition within of human brain

1 – Areas of vision; 2 – Eye perception (retina); 3 – optic nerve (nervus opticus); 4 -optic chiasma – (chiasma opticum); 5 - Left visual tract (tractus opticus sinistrum); 6 -Right visual tract (tractus opticus dextrum); 7 – lateral geniculate body of thalamus (corpus genicula tumlaterale thalami optic); 8 – Four-hill area (corpora quadri gemina) – internal regulation of visual sensors; 9 - optic radiations (radiation optica); 10 - commissura fornix (corpus callosum, commissura anterior) - connecting links; 11 - visual cortex of occipital lobe

Main function of the visual system is the transformation of energy produced by direct or reflected light. Photosensitive cells absorb light and cause a nerve signals to transfer information to the brain. These photoreceptors (cones and rods) transform light by photochemical reaction into nerve impulses. The variations in intensity in time and space provide the information that the brain interprets to create its changing images of the environment being viewed. Figure 4 illustrates transfer of primary visual information in more detail.

The visual cortex includes optic elements (retina), the optic nerve and optic tracts that transfer signals within the brain. The visual image is captured by retina as an inverted image (we see ground and sky upside-down). When signals go through internal areas and visual cortex, the signals are transformed into an image that we interpret as being right-side-up. Left and right eye optic nerves (n.opticus) go through the whole brain, form a cross and through visual area of the thalamus reaches visual cortex. Eyes optic nerves, optic tract, optic chiasma are primarily signal transfers connecting the visual area of the thalamus.

When visual signals reach visual cortex the image information is analyzed, recognized and interpreted in terms of knowledge. Zones 7 and 8 are inside our brain; they serve like “routers” of higher order of signals, leading to visual cortex. Visual recognition is not instant, but a process that includes above-mentioned actions of centers of a brain (Figure 4). This demonstrates that “digesting” of visual information occupies several areas of our brain and consumes a lot of brain energy.

Note that brain blood supply and neuron system were developed through hundreds of thousand years of evolution making human sapiens. Intensive use of brain segments for real time processing of visual information requires a large quantity of energy, reducing our power to think or create - pushing us back down along evolution line toward a homini.

Thus from one point of view we have to analyze how the learner’s brain works with the information delivered and then how this information is converted into new knowledge (new in terms of “customer”, based on the previously mentioned definitive (DF), characteristic (CF) and predictive (PF) functions of knowledge use, which are presented in [10]). We should address also the transformation of introduced above and earlier [15],[16] functions DF into CF and PF, as society needs more “active animals”, not just website followers; besides, virtual reality does not make real world better. Extending further A. Schopenhauer and Pope [22]:

“...readers will not be read…”

We might, regretfully, state that watchers in foreseen future will not be able even to read. The example of teenagers glued to video games for hours on end completely lost and locked into their own private world is all to common!

Thus handling of incoming information and investigating how knowledge is created inside our brain is a part of scheme of knowledge delivery and formation from the physiological point of view. The next steps in this direction are:

- Create a model and monitoring scheme of information flow through visual tracts, tracing activation of segments of brain;
- Experiment of formation of delivery using various channels;
- Set up experiments how visual information acts on brain information “digesting”;
- Balance delivery of information through audio and text;
- Create a model of workload of a brain when mentioned information channels are used;

The result of this segment of research will insight the inputs for design of a control model of information delivery.

---

1 homini - animal that looks like human
From theory of information point of view the question is: “What is the best combination of text, visual and audio information to maximize the efficiency of the delivery of meanings and knowledge?” Text, visual and audio are processed completely by different parts of brain, thus the effect of separating these channels (Figure 5) needs to be evaluated.

This separation includes information volume and density, which channel might bring and strength of impact. Questions such as “how much text we should have in the process of knowledge delivery?” and “how many letters, words and which structure and fonts of text sequence would be most effective” are much more important than we think.

So far nobody has provided a rigorous answer to even this simple question: “Which layout of text information should be efficiently delivered: one, two or three columns?” Book design experience [4] and web-design, see, for example [5], where the role and importance of proportions, visual and colors used described are good starting points. Note here that modeling of information delivery even in terms of text has not been considered quantitatively and systematically.

The formation of performance flow for attention control and the role of empty space was developed by Stanislavski's [6,7] and Gaston Bashlyar's [8,9] systems. This formation of new knowledge and information delivery should be considered together with the aim of maximum efficiency in use of perception and learning abilities. The outcome of this research might have immediate commercial application – for web-design, education and art.

Three theories here might be used together: the theory of filters, the theory of signal processing and the theory of information. All mentioned channels (visual, audio and text) have different properties in terms of information density and in terms of absorption of delivered information within brain. The fundamental question here is: How to combine text and visual information to maximize delivery of knowledge?

C. Knowledge delivery using text

There is no doubt that factors such as the amount of text, its appearance and format (including fonts, color, size, number of columns) have an impact on us, they depend on each other and our ability to absorb/learn, for example: Size of sentences, Length of words all impact, Intensity of visual support (might have negative or positive impact on brain).

Text recognition is one of the fastest processes in the brain, it involves both primarily sensor groups - cones or rods (the latter for black and white sensing) that are much higher in volume - easier to perceive. Text translation in meanings within the brain activates associative memory and other areas, called sometimes “thinking schemes” [20] in contrast to visual recognition ones.

Text recognition for different languages varies in amounts of patterns (letters or symbols): Chinese, Japanese languages involves recognizing thousands of hieroglyphs while for Greek, Latin and their derivatives number of letters to recognize is only in the range 27 to 33. It has also been discovered and already documented [10] that the amount of symbols required for remembering and recognizing text and meaning has strong impacts creativity (in our terms, the predictive function gets weaker).

Clearly, such an amount of patterns, as well as texts (bible, other religious texts), requires memorizing and requires more space and brain energy, as well as increasing the length of recognition algorithms and inefficient use of energy used to decode symbols.

If we apply an algorithmic scheme of meaning decoding it becomes clear that cost of decoding has a double-hit overhead: cost for decoding by itself - cost of energy taken from the brain for decoding instead of thinking. To understand how efficiently we can process information into knowledge a set of experiments is needed to answer the questions:

- What amount of symbols and in which font and format “mean” brain can reliably absorb in given time slot?
- How many brain zones (areas) are activated to “digest” a text?
-What amount of energy from central nervous and blood supply systems both needed “to absorb and interpret” text by brain?

Then the different approaches of information delivery can be analyzed using required power, volume and time as independent variables. Taking into account that repetition and practicing might be both involved delivery of knowledge “at once” is not panacea.

Some knowledge only lodges in the brain and stays there for a long time when you practice it. For example in the area of design patterns: if you only hear and see them and you never code them you can never understand them in depth. The proposed model has to take into account constructivist approach as well [18,19,20].

Every person constructs his or her own personal reality by communicating with and interpreting the environment, it defines speed and cost of individual knowledge absorption.

In principle, the basis of analysis of information delivery via audio is similar to visual and text: analysis of segments of brain involved and estimation of audio information; amount of required energy to consume/process information. Audio flow of information might be measured and analyzed in terms of:
The open questions related to audio form of knowledge delivery are:

- How long can speech be delivered with understanding? Is the rule of golden 20 min justified?
- What if we use text and audio? Does it make an impact?
- How much it extend attention time?
- Audio impact on alpha rhythm “massaging” of the brain
- Impact on information delivery when text delivered via audio and video means (for example subtitles ...);
- Content vs. the length of speech, density of the meaning? (math, programming, philosophy)?
- Impact of repetition - how it works here with text? Audio & text? Only audio? Only visual?
- Web-delivery of knowledge? How it might be maximized in efficiency?

This opens the question of whether services like YouTube can ever be really effective delivery platforms for education. When this phase is relatively clear, semantics of phase one is in need of further “fine tuning” i.e. how do we know what the goals are needed? Two key dimensions are here to consider: science and industry where the most of our students will work.

In addition, neither [11] nor [12] nor [13] even mention resource dependencies of the discipline and the means for its delivery, as well as formation of the curriculum and its required assessment schemes. The delivery process mentioned above heavily depends on student’s innate ability to learn. For weaker students one has to prepare sort of spoon-feeding forms of teaching and frequent assessments. For the more academically inclined the model of curriculum delivery and assessment needs to different with higher level of abstraction and self-study involved, supported infrequent assessments. To some it up with an old adage: “horses for courses”.

There is a need to prepare a quantifiable model of workload of the student and the lecturer taking into account the range of abilities and prior knowledge of the intended students. Clearly that longer journey costs more and it requires greater efforts, thus quantitative analysis and models of curriculum delivery is justifiable subject of research.

V. WEB-SUPPORTED ASSESSMENT

Up to now, except for pure math disciplines, the rigorous formation of a set of essential questions relevant to the current curriculum has been very rare. A logically cohesive sequence of the questions should be arranged for the assessment, mapped to the process of delivery of the discipline; even a semi-automatic procedure might be quite useful, as it reduces teacher’s workload in formation of assessment and, at the same time, guarantees quality.

The ability of student knowledge should be formally measured, analyzed and terminated when PFs > PPr. Indices stand for: s - student and tr - threshold). Threshold indicates the required level of understanding, proven by assessment through analysis of abilities developed in the student’s mind.

The two equations below define coarse content in terms of essentiality and completeness by making joint and essential glossaries:

\[ DFj = DF_1DF_2 \ldots DF_{n-1}DF_n \]

\[ DFe = DF_1DF_2 \ldots \cap DF_{n-1}\cap DF_n \]

The building of a set of questions around an essential glossary DFe forms an essential set of questions CFe and this constitutes a complete formal preparation of the required assessment. Also the combination of DFe and CFe form a core of the discipline assessment.

A. Modes of training/assessment

Assessment using web has two major advantages. Firstly, it can be introduced into the learning process in a training mode, thus becoming an integral part of learning and increasing self-assurance and confidence of learner. During the learning process the student is able to make his or her own judgment using a “training mode” of assessment and there is also the advantage that this can be done anywhere in the world and at anytime.

Secondly, in a secure environment during a defined and scheduled assessment time the student has much more confidence regarding the subject and process of assessment as a benefit of their ‘assessed’ previous training. Training mode might be organized with at various depths and coverage of assessment making available, in principle, unlimited amount of generated questions on the subject.

An interesting research direction is to consider the application of AI for the formation of narrower, special purpose created assessments for any student during training mode, for checking and improving own knowledge on different segments of the subject taught. To achieve this a new web-based tool and formation of semantic channeling and tunneling as a function of student knowledge and progress is crucial. We strongly believe that our computer age needs revision and radical improvements regarding how teaching is achieved as well as a revision of the theoretical principles of curriculum design and automated support for the formation of relevant assessments.

B. Individual Assessment Tool

A supportive net for further development of theory of curriculum design for computer science, invented theoretically [1], estimated in terms of reliable delivery [2] initially tested [14] and developed up to implementation 15],[16] during last few years needs an assessment re-development. A special assessment tool automatically tuned for individual automated assessment is required. Temperature of the problem grows along with grow of number of students (up to 250 - 300 per
module so far), while web-classes will attract thousands [17], or even millions of students, e.g. teaching English in Europe and Asia existing methods and schemes will not be helpful.

The automatic assessment tool we propose can be used in a form of multiple choice answer approach (MCA) [16]. This approach assumes several logically connected (say m) correct answers from n possible options.

The development of special individual assessment tool (IAT) to create and run MCA would be useful across a wide spectrum high education disciplines. An automated individual reenter-able scheme of self-assessment with IAT helps students to assess themselves during semester using their gadgets for checking own performance, seek relevant advice or even pass module as soon as they are ready.

First platform independent version of IAT was released for Apple IPAD, see screenshot on Figure 6 with implementation of:

- Multiple choice answer approach, MCA
- Penalty functions, their forms and quantification
- Iterations of learning and assessment: how it should be
- Individualization of assessment

Figure 6. First version of IAT application of WEDUCA

VI. CONCLUSION

- Proposed approach provides a theoretical basis for formation of the curriculum for high education using web-semantic methods and schemes for clustering web-based information
- Monitoring and observing changes in knowledge domain, including segmentation and landscape evolution are possible to exploit using web-semantic searching methods making automatic curricular adjustment possible
- A framework for curriculum development and delivery including human perception features and three ways of information delivery (audio, visual and text) is described toward improvement in quality and efficiency of personal knowledge formation; brief description of essential models are given

- A formation of automatic supportive assessment mechanisms with substantial and justifiable coverage as well as prototype individual assessment tool are described and illustrated.

ACKNOWLEDGMENTS

The framework described here was instantiated in a prototype form as a version 2.0 of the WEDUCA platform, which is available for download at http://teachflowapp.com

REFERENCES