Teaching computing for complex problems in civil engineering and geosciences using big data and machine learning: synergizing four different computing paradigms and four different management domains

Zoran Babović[1,19], Branislav Bajat[20], Dusan Barac[2], Vesna Bengin[18], Vladan Đokić[2], Filip Đorđević[20], Dražen Drašković[2], Nenad Filipović[12,19], Stephan French[9], Borko Furht[14], Marija Ilić[10], Ayhan Irfanoglu[5], Aleksandar Kartelj[2], Milan Kilibarda[20], Gerhard Klimeck[5], Nenad Korolija[2], Miloš Kotlar[2], Miloš Kovačević[20], Vladan Kuzmanović[20], Jean-Marie Lehn[17], Dejan Madić[18], Marko Marinković[20], Miodrag Mateljević[13], Avi Mendelson[8], Fedor Mesinger[13], Gradimir Milovanović[14], Veljko Milutinović[6,15,16*], Nenad Mitić[2], Aleksandar Nešković[2], Nataša Nešković[2], Boško Nikolić[2], Konstantin Novoselov[4,7], Arun Prakash[5], Jelica Protić[2], Ivan Ratković[3], Diego Rios[9], Dan Shechtman[8], Zoran Stojadinović[20], Andrey Ustyuzhanin[4,11,21] and Stan Zak[5]

*Correspondence:
vm@etf.rs

[6] University of Montenegro and Montenegrin Academy of Sciences and Arts, Podgorica, Montenegro
Full list of author information is available at the end of the article

Abstract

This article describes a teaching strategy that synergizes computing and management, aimed at the running of complex projects in industry and academia, in the areas of civil engineering, physics, geosciences, and a number of other related fields. The course derived from this strategy includes four parts: (a) Computing with a selected set of modern paradigms—the stress is on Control Flow and Data Flow computing paradigms, but paradigms conditionally referred to as Energy Flow and Diffusion Flow are also covered; (b) Project management that is holistic—the stress is on the wide plethora of issues spanning from the preparation of project proposals, all the way to incorporation activities to follow after the completion of a successful project; (c) Examples from past research and development experiences—the stress is on experiences of leading experts from academia and industry; (d) Student projects that stimulate creativity—the stress is on methods that educators could use to induce and accelerate the creativity of students in general. Finally, the article ends with selected pearls of wisdom that could be treated as suggestions for further elaboration.

Keywords: Education, Computing paradigms, Artificial intelligence, Big data, Nature-based construction

Introduction

Expanding computing-intensive problems in civil engineering, geosciences, and related fields towards higher fidelity requires hardware infrastructure that is not only a lot faster, but also a lot less power greedy, smaller in size, and more precise, than what is typically available. Most of these problems are related to simulations that support feasibility studies before a concrete engineering effort is started. As such, the computational needs are strongly fluctuating in time and require flexible and elastic access. Beyond mere hardware access, more complex problems require a higher level of algorithmic and programming expertise, proper data organizational skills, as well as management and administration skills. Students, therefore, have to learn not only how to use the modern computing infrastructure and related programming models, but also how to conduct and administrate the simulation processes, data management, and feasibility studies in a holistic manner. To address these needs, we believe, a competitive course that spans over two or more semesters, or a related program/curriculum, has to include the following components: (1) Computing with different paradigms that could cooperate and synergize on large and complex problems, (2) Management skills needed for complex software engineering management, (3) Examples of good synergy between engineering and management, and (4) Advanced student projects.

From the educational systems' point of view, this article builds on the top of trends growing in the academic and research worlds that affect modern educational systems' challenges [10].

First, the ubiquity of online access to academic resources results in the growing demand for supporting the existing online modes of education, without compromising the quality of education. Second, many scientific research fields, starting to incorporate data-driven approaches, are sharing the following traits:

> (a) The demands for expertise and tools in specific engineering and science domains, as well as in Artificial Intelligence and BigData, and
> (b) The initiatives in research that are shifting from research institutes to include small companies (e.g., there were tens of start-ups funded on AI-driven pharma research upon the success of AlphaFold).

Several of the coauthors of this article have experiences in experimenting with different teaching methodologies to address these challenges at various universities such as: Harvard, MIT, Purdue, Indiana University, University of Manchester, University of Pisa, University of Belgrade, and The National University of Singapore. We believe that a modern academic system, to be successful, should stress the following:

> (a) Incorporate online educational resources and practices that keep students motivated for further domain research upon completion of education curricula;
> (b) Give basics of data-driven research approaches that would enable students to get familiar with the current state-of-the-art techniques;
> (c) Have a short connection with a business incubator/investment program that would help students transfer/embed their knowledge/techniques into a successful business model;

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 3 of 25

(d) Generate a lot closer level of interaction between the teacher and the students, and employ even the cell phone based applications for interaction while creating jointly.

What follows is first about computing, then about management, and finally about synergy, and possible advanced student projects.

Note, however, that the same effects could be accomplished with two different courses (one on computing and one on management), followed by a research course (like UROC, the undergraduate research-oriented course, as well as its graduate equivalent, at Indiana University in Bloomington, Indiana, USA, or EE496/EE596 at Purdue University in West Lafayette, Indiana, USA, oriented both to undergraduate and graduate research).

What follows is based also on teaching experiences from periodic visits and workshops at MIT (Media Labs and CSAIL) and Harvard (Continuing Education and Undergraduate Computing), plus Imperial and Manchester universities in the UK, during four different decades.

More recent experiences described here come from universities in Serbia (Kragujevac, Novi Sad), the University of Montenegro (Podgorica), and those elsewhere in the USA and Europe: CMU, NYU, ETH, and EPFL in Switzerland, UNIWIE, and TUWIEN in Austria, Siena, and Salerno in Italy, as well as Ljubljana and Koper in Slovenia.

These experiences suggest the following important prerequisites for the success of educational missions:

(a) Fundamental knowledge in natural sciences, mathematical logic, and philosophy is absolutely needed;
(b) Educational processes should be devised so as to eliminate the fear of mathematics, which could be achieved only if textbooks are used that were written by the Planet's best textbook writers; and
(c) When it comes to nature-based engineering, fundamentals of geophysics should be taught, too;
(d) Programming skills for existing and emerging computing paradigms should be taught properly.

A Latin proverb says: "Mens sana in corpore sano". Research at the University of Novi Sad indicates that children active in sports at age 4–7, later turn into good students and creative professionals (synaptic developments are the most active at an early age and get considerably accelerated through sports). In conclusion, scientific research at universities has to be synergized with sports.

About computing

There are 4 different computing paradigms and related programming models that could bring benefits to civil engineering, geophysics, and related fields. Some of the paradigms/models are well-established, while others are newly emerging.

The 4 paradigms are: Control Flow (MultiCores as Intel products and ManyCores as NVidia products), Data Flow (Fixed ASIC-based as Google TPU and flexible FPGA-based as it initially was the case with Maxeler DFE products and recently with many

other products of other vendors), Diffusion Flow (IoT, Internet of Things, and WSNs, Wireless Sensor Networks), and Energy Flow (BioMolecular and QuantumMechanical products). For a detailed elaboration on the above-mentioned computing paradigms, see [2, 3, 7–9, 15, 21–23, 28, 32–34, 37].

These paradigms could be compared, hardware-wise, by: (a) Speed, (b) Power, (c) Size, and (d) Precision. Software-wise, these could be compared by: (a) Ease of Programming, (b) Availability of adequate system software, (c) Reusability based on the library routines, and (c) Effectiveness of development tools. These issues are best understood through a series of in-class examples and four off-class projects, with touch-and-feel effects for students.

For any given application, one of the paradigms is best suited. Some paradigms are better suited to serve as hosts, others are better suited to serve as accelerators. However, they all are best utilized if a proper type of synergy among them is induced, i.e. if they are combined in a proper manner, for the best possible final effects.

About management

The four groups of management skills of interest for this environment are: (a) Acquisition of funding and knowledge of interest for the successful completion of a project (issues covered in this course: Funding Mechanisms and Business Administration); (b) Planning of activities, both on the strategic and the tactical levels (issues covered in this course: The CMMI System and selected Agile Methods); (c) Getting Incorporated and protected (issues covered in this course: SBA and PTO methodologies for applications like: Incorporating a new business and protecting new know-how; (d) Getting the web presence with the marketing support (issues covered in this course: How to generate a commercial web site and how to organize a Mind Genomics marketing campaign for that site).

Success on these issues, in most cases, is difficult to measure, so these issues are also best mastered via a series of in-class projects. To be successful, these in-class projects have to build up on top of in-class projects from the computing/programming part of the course. For any given application, one of the above listed skills is the most relevant, but all of them are important and have to be mastered, together with the methods that could enhance creativity [4, 16].

About synergy

By the time when the synergistic effects have to be covered, the students are well educated with practical knowledge from 16 different in-class project assignments. Since the semesters are typically shorter than 16 weeks, some of the in-class projects are paired and delivered at the end of the same week.

Once the weekly in-class projects are mastered, the foundations have been built for further upgrading of the professional know-how. Of course, various in-class examples are of benefit to students, too.

At the semester's end, students have to come up with a proposed solution that synergizes the four computing paradigms and the four management skills. This solution is to be developed in the form of a skeleton for a funding proposal to one of the engineering funding agencies or to a large engineering industry that supports research.

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 5 of 25

For a given holistic project that combines the four presented paradigms, aimed at a simulation that supports a feasibility study, the students have to present the solution that they had envisioned, but also to demonstrate that they are able to accomplish the following: (a) To shed a precise light on the creative process that led to the envisioned solution (using one of the 10 creativity induction methods that they get exposed to in this course); (b) To write about their achievement (with a skeleton for a survey article and also for a research article), (c) To find the hidden knowledge in their simulation studies (using one of the 10 specified data mining algorithms in Big Data conditions); and (d) To discuss the feasibility of the future civil engineering process in the context of finances and branding (using financial calculations and branding methodologies).

For all above to be doable by students in one semester, the teacher also has to cover the following topics: (a) Ten creativity-inducing methods with examples there-off; (b) How to write survey articles and how to write research articles; (c) Different data mining algorithms applicable in geosciences, genomics, and other Big Data contexts, and (d) How to treat the finances of a future software engineering project and how to turn the final result into a brand.

The synergy of engineering and management is crucial for the success of complex missions!

This course first presents the pros and cons of each computing paradigm and then it discusses possible ways for them to synergize. After that, this course presents the pros and cons of each mentioned issue in management and then discusses possible ways for them to synergize, with examples that come from a number of universities worldwide.

The educational follow-up activities

During the semester to follow, interested students are welcome to develop their ideas from "About Synergy" Section into more detail, and possibly to come up with an article to be submitted to a proper conference or journal. Some students may select their follow-up work to serve as a thesis, on the B.Sc. or the M.Sc. level.

One of the important side effects of the university-level educational process is to teach students that, once some goal is achieved, the work towards that goal is not finished until after the results obtained, errors made, and lessons learned are elaborated in an article for public usage, through open literature.

In the efforts so far, related to the methodology presented here, some student results in the follow-up phase ended up being published in prestigious journals like ASCE (American Society of Civil Engineering) and IEEE (Institute of Electrical and Electronics Engineers).

Teaching the four computing paradigms

The first issue is to teach the essence of each relevant paradigm, and the second issue is to teach how to combine the elaborated paradigms using a proper architecture that enables their efficient interaction.

Computing paradigms

The Control Flow paradigm is based on the implementations of von Neumann in the 1940s. This paradigm is best suited for transactional computing and could be effectively

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 6 of 25

used as the host in hybrid machines combining the four paradigms treated in this course. If a Control Flow MultiCore machine is used as a host, the transactional code is best to run on the Control Flow host, while the other types of problems are best treated on accelerators based on other types of paradigms. In the case when the code works on data organized in 2D, 3D, or nD structures, a good level of acceleration could be achieved by a Control Flow ManyCore accelerator. The programming model for Control Flow machines is relatively easy to comprehend. Speed, Power, Size, and potential for high precision of Control Flow machines are well understood.

The Data Flow paradigm and the related machines were created through the inspiration that came from specific observations of Richard Feynman. This paradigm insists on the fact that computing is most effective if, during the computational process, the data are being transferred over infinitesimal distances, as in the case of computing that is based on execution graphs. The Data Flow approach, in comparison with the Control Flow approach, brings speedups, power savings, a smaller volume of computing engines, and larger potentials for higher precision. However, it utilizes a more complex programming model, which could be ported onto the higher levels of abstraction, in which case a part of the presented advantages would disappear.

The Diffusion Flow paradigm originates from the research in massive parallelism (IoT), possibly enhanced with sensors (WSNs). One important characteristic of this approach is a large area or geographical coverage, which means that, in this paradigm, it is theoretically impossible to move data over small distances, during the computing process. However, some level of processing could and should be done within IoT and WSN structures, either for data reduction purposes or for some kind of pre-processing. The final processing of Big Data can be done during the "diffusion" procedure of the collected data toward the host. If the energy for IoT or WSN is scavenged from the environment, the power efficiency becomes extremely high, while the size of the machinery is negligible, as well as its potential for the highest precisions. On the other hand, the programming model has evolved since the initial PROTO approach from MIT and has to be mastered adequately, which could be a challenge.

The Energy Flow paradigm could be used for only a very limited set of algorithms. However, for a selected set of algorithms, it could be extremely fast. No matter if the BioMolecular or the QuantumMechanical approach is used, the processing is based on the energy transformations, and the corresponding programming model has to respect the essence of the paradigm if the best possible performance is desired. For the doable algorithms, the speedup is enormous, the needed power is minimal, the size is acceptable, while the potential for precision is unthinkable of. Adequate programming models are on the rise.

The expectation is, for computing-intensive problems in civil engineering, geosciences, and related fields, that the BioMolecular paradigm could be very useful for nature-based construction engineering, while the QuantumMechanical paradigm could be very useful for highest-detail simulations.

A possible architecture for a supercomputer on a chip

The current technology could fit over 100 Billion Transistors (BTr) on a chip or even a Trillion Transistors (TTr) on a wafer. Consequently, it is possible to place (on a single

chip) both the above-mentioned Control Flow engines (MultiCore and ManyCore) and both the above-mentioned Data Flow engines (ASIC-based fixed and FPGA-based flexible).

However, other acceleration options (in the form of IoT or WSN) and specific emerging options (in the form of BioMolecular and/or QuantumMechanical) have to be placed on the chip, but easily accessible via a proper set of interfaces.

Of course, classical I/O and memory hierarchies have to be placed partially on the chip and partially off the chip and should be connectable through proper interfaces.

Therefore, regardless of the fact that 100BTr or 1TTr structures are used, the internal architecture, on the highest level of abstraction, should be as in Fig. 1.

However, the internal distribution of resources could be drastically different from one supercomputer-on-a-chip to another, due to different demands that various applications are placing before the computing hardware (transactions-oriented or processing-oriented), and due to different data requirements (memory intensive for massive Big Data of the static type, or streaming oriented for massive Big Data of dynamic type, coming and going via the Internet or other possible protocols).

The in-class projects are related to the implementation of a mathematical algorithm in all eight sub-paradigms of the four major paradigms. For that purpose, students can use local machines or the portals like Amazon AWS, Microsoft AZURE, or similar.

Teaching the advanced management

The advanced management issues covered in this section are those oriented to the acquisition of funding, planning of activities, commercializing the final results, and making sure that the dissemination process is conducted properly.

Acquisition of funding and the knowledge of interest

This section teaches how to acquire funds for a project and summarizes the major lessons from a successful MBA program. The in-class projects are: Writing a proposal for funding to come from an agency in the USA or in the European Union and writing the Analytical part of the GRE or GMAT tests (both, Analysis of an Argument and Analysis of an Issue). The first homework assignment enables students to learn all the necessary formalisms related to the writing of research proposals. The second homework assignment represents a possibility for students to practice crisp writing that stresses the most important issues without any unnecessary redundancy.
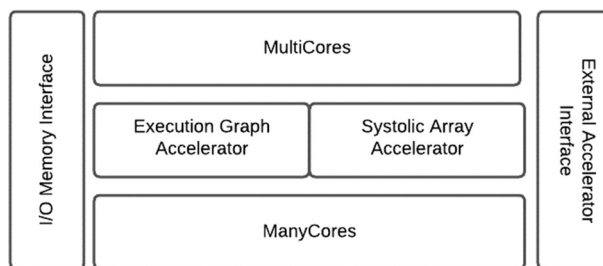

Fig. 1 Block diagram of a Supercomputer on a Chip advocated in this article

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 8 of 25

Planning of activities, both on the strategic and the tactical levels

This section teaches CMMI for strategic planning and one agile method for tactical planning. The in-class project assumes that the proposal for funding was approved. Students are asked to write a strategic CMMI plan for the entire duration of the proposal, plus a tactical agile plan for the first weeks of the work on the funded project. The CMMI is taught using a wide plethora of examples; for civil engineering students an example is used that is oriented to the construction of an airport, and for computer engineering students are given an example that is oriented to the development of a banking core. For agile planning, students are asked to span 1 month of activities at a time, with the first week given extremely precisely and the last week given with ambiguities to be resolved as the process goes on.

Getting incorporated and protected

Assuming that the project has been successfully finalized, the students have to "create" a company to commercialize the effects of the project and to protect its assets. For that purpose, the in-class projects are to write a business plan using the SBA guidelines and to protect assets (names and products) using the PTO methodology. For this part of the course, successful individuals from the industry are invited to join and share their experiences. For the SBA guidelines, an example is given related to a startup company that was successfully sold to a major mainstream company. For PTO patenting, an example is given related to one of the largest out-of-court settlements in the computing industry.

Getting a web presence with the marketing support

Assuming that the company is already incorporated, the in-class projects are to create the Web Store for dissemination of the company's product or service and to create the Mind Genomics marketing campaign for that web store. Three different approaches to the generation of Web Stores are presented (suitable for small, medium, and large companies). As far as mind genomics the stress is on the mathematics treatment behind this method, plus on the balance between statistical analysis and psychological values exchange. This last in-class project is relatively complex since it has to synergize a number of issues of importance for the marketing process.

Teaching synergy

In this part, the stress is on the inducing of creativity, on creative writing of research-oriented articles, on issues of interest for dissemination, and on the search for hidden knowledge that is buried inside the large data repositories.

Ten creativity-inducing methods

Ten methods for inducing creativity are taught and each student has to specify, for the end-of-class project related to complex simulations, which one of the 10 methods was predominantly used during the process of generating the end-of-class project [4]. These 10 methods have been classified into five different groups, with each group including two sub-methods: Mendeleevization (Inductor and Catalyst),

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 9 of 25

Transdisciplinarization (Modification and Mutation), Hybridization (Symbiosis and Synergy), Remodelling (Granularization and Reparametrization), and Unorthodoxization (ViewFromAbove and ViewFromInside). In class, students are given examples that demonstrate the essence of all ten methods.

How to write survey articles and how to write research articles

Students are taught how to write survey articles and how to write research articles. The major elements of a survey article are a rigorous classification and a systematic description. The major elements of a research article are: (a) Problem statement; (b) Existing solutions and their criticism; (c) Essence of the proposed solution; and (d) An analysis that demonstrates the superiority of the proposed solution. For the in-class project, they have to write a survey article on other similar efforts related to what they had proposed for their end-of-class project, while for the research article, they have to write only the skeleton, related to their end-of-class project.

Different data mining algorithms

Ten different data mining algorithms, applicable not only in engineering, physics, and geosciences, but also in genomics environments, are taught, and students are asked to elaborate on the utilization of one of them for the analysis of data coming from their end-of-class project. Students are taught that there is always some hidden knowledge that has to be datamined, using appropriate hypotheses created with the help of one of the ten methods to induce and accelerate creativity among students. The ten data mining methods have been selected based on the level of citations attributed to different data mining methods.

How-to's of potential importance

In this final part of the course, issues are covered that may differ from one generation to the other, having in mind the predominant profile of students. Possible topics to be covered may refer to finances, branding, etc. Students are taught how to synergize engineering, science, finances, and marketing, for the best benefit of the missions that they are a part of, in academia or in industry. In this part of the course, the initiative is given into the hands of students, their ideas are evaluated in front of all others, errors in reasoning are discussed and lessons for their future professional work are learned and elaborated using appropriate examples.

Teaching applications that need synergy

In "Classes of Problems Needing Synergy of Computing and Management" Section we first classify the applications in civil engineering and geosciences that need drastic acceleration and then in "Elaboration" Section. we elaborate on their essence. These two-course sections help the students to learn how to synergize Control Flow and Data Flow paradigms.

Classes of problems needing synergy of computing and management

Examples that follow cover simulations of Big Data problems needing Data Mining in specific or Artificial Intelligence in general, and are related to complex problems in Civil Engineering, Geo Sciences, or related fields, namely:

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 10 of 25

(1) NBCS—Nature-Based Construction Sciences.

(2) GSNC—Genomics Supporting NBCS.

(3) EQPA—EarthQuake (EQ) Information Systems for Prediction and Emergency Alarming.

(4) NCMS—Creation of new civil engineering materials and structures (CO2 and EQ sensible).

The algorithms used in these four domains could be of the following types:

(1) Statistical and stochastic processes that mimic nature (life processes of plants and animals).

(2) Genomic algorithms of the type NW (Niederman-Wuensch) or SN (Smith-Waterman) or similar.

(3) PDEs of the type FE or FD or hybrid solutions (FE = Finite Element, FD = Finite Difference).

(4) Tensor calculus and mathematical logic or hybrid algorithms (based on vectors and matrices).

For all these applications, we assume that the optimal distribution of transistors on the chip, over the involved computing resources, should be as it was given in Table. 1. This means that the flexible FPGA-based Data Flow dominates in the distribution of transistors over the paradigms implemented on the supercomputer-on-a-chip.

Of course, for applications of interest, data are provided either from the internal Memory System or from an external Internet Protocol or another stream. In a number of applications, data are provided via IoT (Internet of Things) and/or WSNs (Wireless Sensor Networks).

Elaboration

In NBCs, it is wiser to utilize biological structures that grow relatively fast, and are populated with insects that produce relatively strong nano-materials, than to build walls of concrete that are CO_2-emitting and EQ-fragile. Also, it is wiser to use fish and plankton, rather than metal structures, to protect fragile underwater developments. Before each and every investment of this type, a feasibility study has to be performed, and it is best if it is based on a simulation. However, adequate simulations could be extremely time-consuming and could last for years, even at the fasters Control Flow supercomputers of the

Table 1 A recommended distribution of resources for a chip to support Artificial Intelligence for Big Data (BTr = Billion Transistors) [24]

Chip hardware type	Estimated transistor count
4000 ManyCore Processors with related memory	12BTr
4 MultiCore Processors with related memory	4BTr
1 SystolicArray for Tensor Calculus	1BTr
1 Reprogrammable DataFlow engine	69BTr
Memory and I/O Interfaces	14BTr

world. Consequently, the solution is to switch from Control Flow to a proper combination of the other three computing paradigms. However, the students of civil engineering have to be educated so that they could utilize the benefits coming from other available paradigms.

In GSNC, the genetics of existing species have to be analyzed, to figure out what genetic characteristics bring the desired effects, so as to get to know what genetic changes could bring more effective species. Related runs of genomics code may also take years to execute, which places strong demands on the utilization of faster machines. In other words, computer simulations, based on Control Flow engines, that include enough details, could take time which is unacceptable. Again, the solution is in proper synergies of the four paradigms, and in educating the civil engineering students about advanced computing and the related management know-how.

In EQPA, models do exist of towns and cities, which are based on bricks and cement, that could be applied, in simulators, to earthquakes of specific characteristics. However, simulations of earthquakes with these models as inputs may take years, even on the fastest Control Flow machines today. Again, the simulation process could be drastically accelerated if only a proper Data Flow accelerator is used. The Data Flow machines are well suited for PDEs (Partial Difference Equations) of the type FE (Finite Element) needed for damage predictions, and for PDEs of the type FD (Finite Difference) needed for ad-hoc alarming in emergency situations.

In NCMS, new materials with desired properties and new procedures that generate higher effectiveness are best found if Artificial Intelligence algorithms are utilized and combined with standard algorithms used in research on materials and procedures. These hybrid algorithms (the standard ones enhanced with Artificial Intelligence) could be extremely computing-intensive, so again, the solution is in the synergy of several paradigms and related education. Artificial Intelligence can be used to decrease the number of iterations in computing-oriented studies, but the best effects are obtained if Artificial Intelligence is combined with fast computing paradigms [25].

Student feedback

After this course, some students end up in academia for higher degrees and sponsored research, while others join the industry, with the stress on either research or production. The experiences of students, as well as of the various educators on similar missions, are summarized next.

Academia

In its long form or short form, as already indicated, this course has been presented so far in a number of universities worldwide, and in the case of the most talented students, class projects resulted in journal articles of successful patents.

The long form, for 1, 2, 3, or 6 credits (in each case, the presented material was the same, but credits were assigned according to how heavy were the in-class projects), the course was presented at Purdue University, Indiana University, NYU, Polytechnic University of Barcelona, Technical University of Vienna (TUWIEN), University of Vienna (UNIWIE), University of Ljubljana in Slovenia, University of Montenegro, University of Belgrade, University of Kragujevac in Serbia, and Bogazici University in Istanbul.

In one of its short forms, this course was presented at MIT Media Labs, MIT CSAIL, Harvard Computer Science, Harvard Continuous Engineering Education, Columbia, CMU, University of Pittsburgh, Polytechnic University of Valencia, ETH in Zurich (both, Civil Engineering and Computer Science), EPFL (both, in Lausanne and Geneva), University of Pisa, University of Siena, University of Napoli, University of Salerno, University of Zagreb, University of Rijeka in Croatia, Koc and Sabanci universities in Istanbul, Tsinghua University and Shandong University in China.

At these courses, the presence of Civil Engineering and Geo Sciences students was exclusive or mixed with students of Computer Science or Computer Engineering, or the majority of students were of the Computer Science or Computer Engineering profiles, with the presence of some Civil Engineering and Geo Sciences students.

The students of Civil Engineering had an exclusive presence at some of the courses at: Purdue University, ETH in Zurich, the University of Ljubljana in Slovenia, and the University of Montenegro.

The students of Civil Engineering were mixed with other students at: MIT Media Labs, Harvard Continuous Engineering Education, the University of Kragujevac, and the University of Rijeka.

No difference in the level of acquisition of either computing or management skills was noticed between the students of Civil Engineering or Geo Sciences and those of other profiles. This means that the students of Civil Engineering and Geo Sciences got more out of this course since others had a "head-start" on basic computing skills.

Industry

The impact of this course on students was different for those coming from Civil Engineering and Geo Science industries than for those coming from Computer Engineering or Computer Science industries.

After the course, the students of Civil Engineering became very enthusiastic about Nature-Based Construction, and in the synergy part of the course, they offered some very innovative solutions in the domain of Nature-Based Construction. For example, a student of Civil Engineering, working on the construction of Montenegro's first motorway, specified the number of instances where Nature-Based Construction would work better than concrete in some critical places on the motorway.

A student of Geo Sciences, Biology, and Ecology, at the same university, was able to notice a number of instances when the fauna of the Shkodra Lake (shared by Montenegro and Albania) could bring a number of benefits to the ecosystem, together with solving some important engineering issues.

Actually, the students of universities placed lower on the Shanghai 1000 List were more motivated in these courses and accepted a large number of in-class projects (a notable characteristic of this course) with greater enthusiasm, on average.

After the course, the students of Computer Engineering became very enthusiastic about the design of supercomputers on a chip and several of them later joined the leading companies in that field.

Some of them were instrumental in bringing small departments of important USA companies to their small country, namely to Serbia. The Serbian departments of companies like Esperanto or Maxeler were led and highly populated by former students

of computer engineering taking this course. Companies like Texas Instruments and AMD from the USA ported some of their testings to Belgrade, to small companies there, led by former students in the course described here.

Actually, even one past vice president of Intel took this course in 80's of the previous century, while one of the current vice presidents of Qualcomm not only took this course but for some years after his B.Sc. graduation was also the Teaching Assistant of this course and got his Ph.D. from the University of Belgrade in Serbia, under the mentoring of a coauthor of this article. And finally, one of the co-owners of the USA company (Integra) that was sold for 1.075 billion dollars took this course in the past, while another past student of this course is the co-owner of the patent that holds the world's record for in-court patent settlement, in the whole history of computing (with Marvel).

Of course, the overall impact of this course would be best evaluated if both the Civil Engineering, Geo Science, Computer Science, and Computer Engineering students would do evaluations of this course after they retire.

One of the coauthors of this article, when teaching the related subjects, insisted to his students that the formula for success in engineering, and in general, reads: $M3 + C4 = I7$. For the case of nature-based construction, the M3 means: proper Materials, adequate Methods, and high Motivation. Also, education in this and other technical fields requires students to support 4C: Communication, Collaboration, Creativity, and Critical Thinking. Education is no longer a "waiting room" for our growing up or aging, but the main value of the next century and we have to acquire new knowledge every day. Information, Innovations, and Internationalization are Important Impacts for Inspiring an Intelligent World.

What helps a lot is to demonstrate how simple Control Flow solutions could be implemented using the Data Flow paradigm. This is best done by implementing several example solutions through simulation, with the overall goal to use the advantages of speed, power, size, and precision.

Although the class time of one-semester classes is relatively short, students would typically have enjoyed more to see the paradigm implemented or analyzed beyond examples of medium complexity. Nonetheless, much was still achieved in the given time frame, as the optimized approach was utilized in numerous ways. Comparison with other emerging paradigms could be of great help, too.

The following sections include examples that could be used as a student project that enhance creativity. The section to follow elaborates on the major issues of importance for the given problem. The students are taught that there should always be plenty of room for improvement based on creative thinking. Student ideas are discussed in the classroom so that each student is given an opportunity to contribute new ideas that further contribute to performance improvements.

Examples

Experiences from a number of universities worldwide are presented next, using the same template in each particular case, if and where appropriate and applicable:

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 14 of 25

a. Definition of a Big Data Problem in Civil Engineering or Geo Physics that could be solved with Machine Learning and needs the researchers of the profile built by the course described in this article.

b. Justification of the need for the presented problem, in the context of one of the current priorities in modern Civil Engineering or Geo Physics.

c. Presentation of the algorithmic math and related math logic.

d. Estimation of the possible effects (speedup, power savings, size reduction, and precision increase) that could be achieved by combining a subset of the presented paradigms, using proper management skills.

e. A positive classroom experience.

f. A negative classroom experience.

g. An example of a project needing student profiles is described here.

h. The specific problems that have to be resolved for the success of the project.

i. The topic of the presented course that is most relevant for the success of the project.

j. The examples of the engineering-management synergy that could be crucial for success.

Each point above was asked to be presented in about four sentences, with pointers to references that shed more light on the issues covered, having in mind the fact that this article defines principles and gives guidelines.

When talking about modern mathematics-oriented teaching, various types of models and related subjects could be used, by means of which certain relationships and processes could be visualized, and thus made as vivid to students as possible. According to some antique philosophers, learning is a product of the activities of those who learn, meaning that a successful educational process needs at least a part of the students to be interested in the subject taught. This holds also for current trends in modern education, based on online learning and self-education (students have to be in agreement with their understanding of the learning process used).

When talking about computing issues, the main problem in doing an interdisciplinary project is the need to understand "the language" of all disciplines involved.

For example, Civil Engineering and Geo Physics students do not master the theory of building a distributed computing system, while computer science students have a hard time understanding the mathematics of weather prediction needs.

When talking about management issues, the experience indicates that the instructor is getting better results when (s)he lets the students define their topic for the final work. Also, we need to teach students that failure is a part of the road to glory. Students must understand how to recover from a failure and how to do a trough post-mortem.

When talking about the technical side of the project, the experience indicates that most of the students have a tough time understanding the differences between cycle-by-cycle and the use of parallel, event-based simulators.

Issues in nature-based construction

The stress should be on both the algorithms and their synergistic or symbiotic interaction. Also, other issues of importance should be covered, as well, e.g., Power Engineering [11–13] in the context of NBC.

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 15 of 25

Power engineering is often viewed as being an old discipline with very little to discover. However, this cannot be further from the truth. Basic principles of the electric power grid design, operations, and planning, must be revisited as new hardware technologies, and data-enabled sensing, computing, and control, are rapidly emerging and offering flexible sustainable energy utilization, which forms an integral part of NBC's vision. All else the same, smart integration of energy resources and users is key to the NBC of our environment. The power grid can and should be viewed as a complex network system whose spatial, temporal, and functional attributes are so diverse that these can no longer be managed by the utility in a centralized manner. Instead of assuming that the utility knows full information about the grid-connected components (resources and electricity users) and can manage them in a top-down worst case approach manner, the grid of the future will have to interactively exchange minimal information with the grid users. These users, in turn, have many embedded data-enabled means for predicting, controlling, and decision-making at different degrees of spatial, temporal, and functional granularity. Such future grid users will offer major help in energy production, and delivery and will support societal objectives through increased controllability and observability, previously not harvested in extra-high voltage (EHV) power systems. We must educate our students on how to rigorously pose such evolving complex network system problems by establishing flexible protocols for NBC of our environment. There exists a major opportunity to draw on ML, AI, and other big-data tools and contribute to decarbonization by a huge number of grids demonstrating the effects of the law of large numbers. We have just begun to teach about the changing electric power grids by taking such a view, at MIT, where we have a course named Principles of Modeling, Simulation, and Control of the Changing Electric Energy Systems, in the EECS Department. These principles, once understood, will facilitate the major use of tools today's students are excited about, such as ML and AI, in the context of the overall NBC objectives. If we don't attempt to teach power systems engineering with an eye on the emerging objectives, this will represent a major roadblock to the innovative use of green resources.

It is also important to teach synergies between computing and civil engineering, on one hand, and nature-based construction and urbanism with architecture, on the other hand.

As far as Urbanism, machine learning and big data could be and are used widely both in:

(a) Understanding/monitoring/assessment of the existing urban fabric, phenomena, and behavioral patterns.

A vast amount of Earth observation data is available through satellites, including key information about the environment and its human impacts. Researchers now have access to petabytes of open and free image archives, which they can use to track environmental changes over time. However, making sense of this big data collection poses challenges for geospatial technology experts who must design tools that allow analysts to capture subtle changes in the ecosystem at the global, regional, or local scale [30].

AI has been extensively used in the analysis of satellite image time series for automated extraction of information across various fields [31], including urbanization impact assessment. Its applications include mapping urban boundaries, land use, and cover change detection, population dynamics, urban climate, air quality metrics, and more [36]. Geospatial tools utilizing remote sensing data, and interactive visual features through geospatial applications are essential in enabling analytical reasoning for urban planning through the use of available geospatial information.

(b) Design/evaluation/prediction of the new development practices.

The reasoning behind employing this type of research can be seen in the quest for securing sustainable policy-making, cost-effective allocation of resources and environmental protection, smart city governance, and social justice [20, 35]. Based on the most recent and comprehensive research on machine learning for spatial analyses in urban areas, conducted by Casali, Yonca, and Comes [6], current studies can be classified according to their primary focus on infrastructure (the highest percentage of studies), socioeconomic, land use, urban form, and environmental topics. The same study identifies the most used supervised algorithms to be neural networks, random forests, support vector machines, gradient boosting decision trees, decision trees, K-nearest neighbor, and logistic regression. The constraints for data-driven urban planning can be perceived in a-contextuality, neglecting the political and cultural dimension of the city, promoting technocratic forms of governance, reinforcing existing power relations, complying with private and market-led interest, and hence causing spatial and social inequalities and raising ethical concerns.

On the contrary, practices, and studies continue to shed light on other possible fields of application, such as social inequality, gentrification, social vulnerability, crime prediction, and exposure to risks and hazards.

As far as architecture, machine learning, and artificial intelligence have also widened the imagination and creative process within architecture, specifically through systems such as DALL-e and Mid Journey that generate images based on textual descriptions, providing grounds for combining various concepts. This practice tends to reduce the process from conception to execution, specifically by improving research, optimizing the workflow, enabling scenario testing, and maximizing efficiency in the construction process.

However, we still believe that natural intelligence is superior compared to artificial one (whether we admit it publicly or not). On the other hand, artificial intelligence is expected to be flawless, although we are aware of the fact that our natural intelligence is not like that at all (remember how many times we make a mistake in a single day—for example, drop something, miss the door, say the wrong word, etc.). Artificial intelligence (i.e., Machine Learning) produces results with errors evidently (usually with a low probability, but the event is still possible). How important these errors are to us depends on the problem we are solving, and that message has to be conveyed to the students. Therefore, in the education process, special attention must be paid to the following aspects:

- What is the maximum expected accuracy of the model, and
- How to control unallowed (large) errors that may occur.

Considering the first problem, it is very important to estimate the maximum accuracy and repeatability of the data on the basis of which the modeling is done. Given that today's complex problems involve a large number of parameters, the domain of definition of the problem itself is multidimensional and, as a rule, very large. The accuracy and repeatability of the input data set are something that must be taken care of, not to get misleading results (as in the case of simulating a game of chance, for example). Of course, as a rule, accuracy and repeatability do not have to be constant in the entire domain. Also, it should be emphasized that the accuracy of the model itself cannot be better than the accuracy contained in the data itself, on the basis of which the model is formed (e.g., one cannot model the results of measurements more precisely than the accuracy with which those measurements were carried out).

Machine learning models can approximate a certain process with a very high probability of high quality, but regardless of that, in a small percentage, very large unallowed errors could occur. For example, if we are determining the position of an autonomous vehicle on the road, it is useful to have the accuracy of the order of a millimeter, but it could be very inconvenient if an error of several tens of meters occasionally occurs (which could lead to a traffic accident). In order to eliminate these errors in practice, it is very useful to understand and use the classic optimization methods, in addition to modern models based on machine learning. That is why special attention should be paid to them in the educational process.

Also, it is important to teach students how to synergize two or more algorithms proven effective for a number of applications, like image understanding, especially in conditions when the image is corrupted by noise or damaged in a number of different ways. For example, two algorithms introduced by Stan Zak are: Generalized Brain-State-in-a-Box neural network (gBSBNN) [26] and combined Discrete Fourier Transform and neural network (DFT-NN) [14]. Their synergy envisioned for the research in NBC is as follows: For one set of conditions, gBSBNN performs better, while DFT-NN performs better for another set of conditions. The exact specification of the two sets of conditions could and should be a subject of a study.

All these examples could and should be presented before the students in each and every offering of the described course, as examples of good practice. Theoretical knowledge is best absorbed if presented through innovative examples from engineering practice.

Issues in genomics

These issues require an interdisciplinary education, covering: Genomics, acceleration of processes in computing, and data mining. The students have not only to master the algorithms of interest but also to understand the ways to accelerate them.

One example covers the experience of a student whose code, after acceleration, was slower. That happened simply because the acceleration issues were not covered in enough depth. In another case, students had great ideas on how to find correlation

between the desired properties and the genetic structure, but they did not have a proper education in data mining.

Students especially enjoyed the lectures that teach them how to use genomics to engineer the directions of mutations which leads to new species that poses specific desired characteristics of interest for nature-based construction in specific, and nature-based engineering in general.

Issues in earthquake engineering

After an earthquake, buildings might suffer significant damage without collapse (Fig. 2a). There is a need to "tag" such buildings to decide if it is necessary to evacuate their occupants based on the damage sustained by the buildings (Fig. 2b). Such tagging decisions are dependent on the damage state of the building immediately after the earthquake. Structural post-earthquake functionality is conventionally evaluated by trained engineers via visual inspection of the damage. A building is tagged "Green" (unrestricted access), "Yellow" (restricted access), or "Red" (no access) according to the severity of the observed damage. This process usually lasts several months, which leads to the situation that people live in untagged buildings that may be unsafe for them or inhabitants leave their homes (move or rent a hotel, etc.) although their property is safe.

The current process for building tagging produces big delays in the recovery process after an earthquake.

Not knowing the status of the possibility of occupying the building several months after an event increases the monetary losses and frustration of the affected people. Instead, the procedure can be changed, so that the buildings safety ("tagging") is based on the image recognition algorithm.

As a replacement for the long surveys on the field, fast production of photos from the site, could be used for the derivation of the building's safety. Machine learning algorithms would perform an assessment using big amount of data (photos) that are quickly collected on the field.

Fig. 2 a Damage and collapse of infill walls in a 2-storey RC building in Durrës during Albania earthquake (MW 6.4) of November 26, 2019 [18] and **b** building usability in the Lower town of Zagreb after the M_w5.4 Zagreb (Croatia) earthquake of March 22, 2020 [1]

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 19 of 25

For this purpose, well educated students are precious for the employment of such machine learning algorithms and their quick use immediately after the earthquake event. Also, further development of machine learning algorithms and sample preparation requires combined knowledge in the field of earthquake engineering, construction, and programming.

Carefully designed curricula that contain theoretical basis with the application through practical examples and case studies, is necessary for preparing the students for this problem in practice.

Employment of existing datasets from previous earthquakes, containing huge amounts of data, would be of great importance for successful training and development of knowledge and skills within the students.

Enthusiasm and effort of students to reach the positive results in the application of the machine learning algorithms employed for this purpose give additional motivation to professors, but also to the next generation of students. Students organizing the teams working on the developing and applying the machine learning algorithms on the case studies and the final presentations showing the results and goals reached, opens new avenues of ideas for the upcoming generations of students.

It is always negative to have some students not being interested to invest in themselves by devoting time and energy, but additionally giving de-motivating statements to their colleagues that are working hard on the problem.

Crucial skill for a successful application of this work in practice combines knowledge from the field of earthquake and construction engineering and programming. Eyes opening by problem definition, then how to get the data by measuring, then classic treatment, then modern treatment, then comparison, and implications!

One of the problems of investigating earthquakes raw data is that they occur at different scales and under a wide range of physical phenomena. Developing of novel computing techniques based on mapping the behavior of structures according to the experimental results obtained directly from the field or by testing in laboratory conditions (Fig. 3), could generate accurate models for disaster mitigation as well as structure response prediction (Fig. 4). Nowadays, there is a tendency in computer science to develop understandable software, based on numerical methods and Artificial
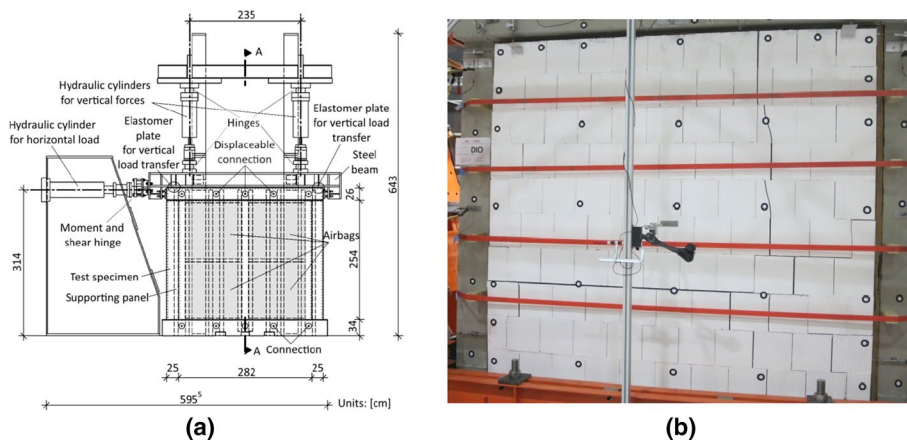

(a)   **(b)**

Fig. 3 **a** View of test set-up [5] and **b** damage of the infill wall at the applied displacement of 49.5 mm [17]
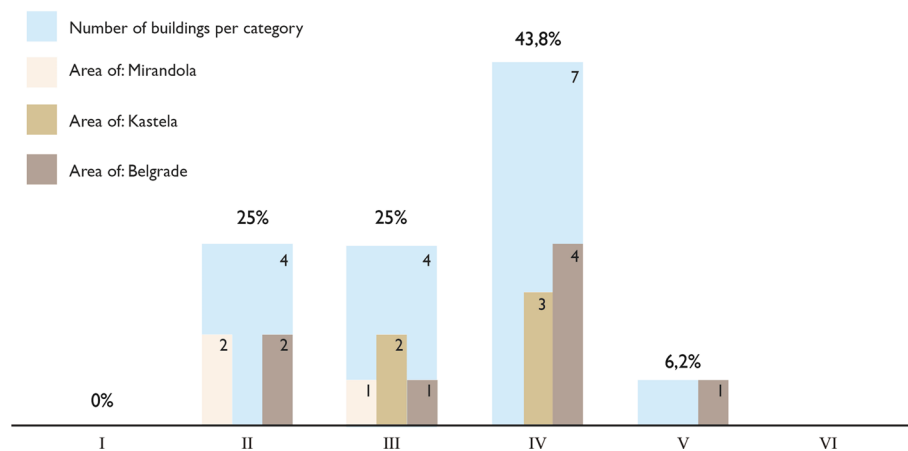
Fig. 4 Index of structural response for the three study areas [27]

Intelligence (AI) to cover currently a big gap between computing and earthquake engineering. Taking the data from a series of earthquake events, having good computational skills, and revealing novel computing methods are crucial for future education in the field of earthquake engineering.

Possession of information in computer science is equally important as the ability to use its potential. Every piece of information is also consequential in civil engineering, given that its processing often involves complex procedures and significant costs.

In addition, the cooperation between computer science and the construction industry is to be more pronounced, in parallel with an exponential increase in the interest of students over the past years.

Additional integration of these two areas into the educational frameworks would be beneficial, since the practical application of machine learning paradigms could lead to fast and accurate solutions of even complex problems in various branches of civil engineering.

Issues in materials engineering

This is what is considered to be important in the educational process:

- Learning undergraduate physics to the level of solving practical problems and describing basic principles;
- Producing a formulation of a solid epistemic scientific method based on solid mathematical formalism and supported by experimental pieces of evidence;
- Discovering and explaining new scientific phenomena in research domains such as material science, and biochemistry;
- Devising design of new materials of given properties (e.g., superconductivity, thermal conductivity, photovoltaics, …) or new experiments capable of verifying new theoretical models (e.g., electronic ion conductivity, catalytic activities of 2D materials).

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 21 of 25

A modern academic system, to be successful, should incorporate the following features:

- Online resources and practices that keep students motivated in further domain research upon completion of education curricula;
- Teaching techniques to research literature sources, mining papers/lectures that present results of relevance and importance to the given research goals;
- Basics of data-driven research approaches that would enable students to get familiar with the current state-of-the-art techniques;
- Teaching how to run computational experiments in a reproducible manner, allowing different members of the same team, or even another team, to get the same results without paying extra cost;
- Have a short connection with a business incubator/investment program that would help students transfer/embed their knowledge/techniques into a successful business model.

Further actions should depend on the feedback coming from the research processes.

Issues in simulation and prediction of weather and climate

Simulation and prediction of weather and climate (SPWC) is a unique problem in the sense that the central mechanism addressed is the one of fluid motion, for which we have the well-known PDEs, the Navier–Stokes equations—known, as this is written precisely for two centuries—but the PDEs that require initial and boundary conditions both involving a vast number of sub-problems. On top of this computing power desired is much above any possibly available, so compromises need to be made. No wonder thus that the big data approach is close to becoming competitive with that of solving the PDEs. The initial conditions in place might have happened before sufficiently close to those we observe.

Just about all methods in the use of solving the atmospheric PDEs require initial conditions at a discrete grid of points, or grid of cells, the more the better. We cannot measure desired values at specific points, so what are the values we obtain using our method of choice at a chosen grid?

Perhaps the dominant view in place is becoming that of looking for averages of the air inside the volume of the cell of our grid, referred to as Reynolds averaged values. But inside a cell, horizontally some kilometers apart from one another, processes take place for which we do not have equations. Clouds are an excellent example. Or processes with equations that need a vastly higher resolution than the one we can afford. Thus, "parameterizations" are needed, representing the impact of these processes as dependent on variables for which we are solving our PDEs.

In addition, neighboring air cells communicate via motions of a smaller space scale than the one we afford to use. Thus, parameterizations of these transports are needed. Lower boundary: Soil, vegetation, water, ice, need parameterizations each of its kind. Based on experiments and measurements, but for turbulence transports lately also on the results of Large Eddy Simulations (LES).

For success in all these subcomponents of SPWC, obviously, an excellent education in basic mathematics and physics is needed. How can we measure the skill of innovations made is a separate issue, a science of its own. While the PDEs are a core content of the SPWC models, there obviously is no analytical solution to compare against. Nature—God if one wishes—decides what is THE solution. The weather happens. Climate change as well. Regarding weather, skill in major events, like the track of a hurricane—recall hurricane Sandy of 2012 and Irma of 2017—or major weather and/or weather generating features, have understandable priority: Such as the verification of the forecast position of the upper-tropospheric jet stream [19]. Most frequently though, consistency in simulating events of an obvious economical value, e.g., the position and intensity of precipitation forecast, are given priority.

While the need for the masterly use of computer resources is for all component problems a requirement, next to choices in mathematics, advancements come from an understanding of component problems, in the community referred to as „physics". Serendipity? What should education that aims for successful innovators stress? Never accept mainstream thinking as given. Believe in yourself! Have courage. At times it takes years for an idea to come to fruition. It took 15 years for the main advancement that led to the major result of the citation just above.

The Interplay of progress in SPWC and management planning is intense, largely because of the most work in the development of SPWC models happening in government organizations. They make plans that require years of work. The "deliverable" is planned to perform better than its predecessor. Who is going to accept the responsibility if this does not really happen? The culture of this attitude needs addressing, again requiring courage.

Conclusion

This course tunes the students of civil engineering and geosciences to the potentials coming from the synergistic interaction of four different computing paradigms.

This course is discussed in the context of civil engineering and geosciences but could be easily ported to other engineering disciplines.

This course prepares students for complex educational or research missions, both in academia and in industry, for research and/or production.

This course (the approach advocated in this article) assumes the existence of a super-computer on a chip that includes some of the existing computing paradigms; namely, those which are used more frequently, and effectively interface to the other paradigms, which are used less frequently!

Researchers without ICT backgrounds sometimes claim that software engineers do not aim to comprehend the essence of the universe, and their only goal is to embed it all in a software algorithm or an Excel Spreadsheet. However, often they are not aware of the fact that, for reaching ultimate results in an area, we all already use algorithmic ways of thinking. For this reason, understanding basic methodologies and concepts of computing science could be the most valuable asset. However, this synergy should be applied in both directions; there are many examples, from genetic algorithms to NPL applications. Only together, ICT researchers and basic science researchers can understand the universe.

Finally, for a nation, the following issues are the most important: To build an age pyramid that is not upside down, to build the educational system that prepares students for their future entrepreneurial missions along the lines elaborated in [29], and to create mechanisms that motivate and praise the best educators.

In summary, these are the most important issues to be taught to opinion leaders, policy makers, educators, and students [29]:

1. Education is the job one of every country.
2. Technological Entrepreneurship (TE) could and should be taught.
3. TE = key to world prosperity and peace.
4. Failure is ok—start again.

In his addresses to young researchers, a frequent message of Nobel Laureate Jean-Marie Lehn is:

"Science Shapes the Future of Mankind! Participate!"

This invitation, coming from a Nobel Laureate, is a most effective motivation factor for young and talented, to join the avenues of research.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Innovation Center of the School of Electrical Engineering in Belgrade, Belgrade, Serbia. [2]University of Belgrade, Belgrade, Serbia. [3]Esperanto Technologies, Mountain View, CA, USA. [4]National University of Singapore, Singapore, Singapore. [5]Purdue University, West Lafayette, IN, USA. [6]University of Montenegro and Montenegrin Academy of Sciences and Arts, Podgorica, Montenegro. [7]University of Manchester, Manchester, UK. [8]Technion, Haifa, Israel. [9]Indiana University, Bloomington, IN, USA. [10]MIT, Boston, MA, USA. [11]National University of Science and Technology MISIS, Moscow, Russia. [12]Harvard University Medical School, Cambridge, USA. [13]Serbian Academy of Sciences and Arts, Belgrade, Serbia. [14]Florida Atlantic University, Boca Raton, USA. [15]University of Montenegro, Podgorica, Montenegro. [16]Montenegrin National Academy of Sciences and Arts, Podgorica, Montenegro. [17]ISIS - Universite de Strasbourg, Strasbourg, France. [18]University of Novi Sad, Novi Sad, Serbia. [19]University of Kragujevac, Kragujevac, Serbia. [20]Faculty of Civil Engineering, University of Belgrade, Belgrade, Serbia. [21]Schaffhausen Institute of Technology, Rheinfall, Switzerland.

Babović *et al. Journal of Big Data*   (2023) 10:89

Page 24 of 25

References

1. Atalić J, Uroš M, Šavor Novak M, Demšić M, Nastev M. The Mw5.4 Zagreb (Croatia) earthquake of March 22, 2020: impacts and response. Bull Earthq Eng. 2021;19(9):3461–89. https://doi.org/10.1007/s10518-021-01117-w.
2. Babovic ZB, Protic J, Milutinovic V. Web performance evaluation for internet of things applications. IEEE Access. 2016;7(4):6974–92.
3. Benenson Y. Biomolecular computing systems: principles, progress and potential. Nat Rev Genet. 2012;13(7):455–68.
4. Blagojević V, Bojić D, Bojović M, Cvetanović M, Đorđević J, Đurđević Đ, Furlan B, Gajin S, Jovanović Z, Milićev D, Milutinović V. A systematic approach to generation of new ideas for PhD research in computing. Amsterdam: Elsevier; 2017.
5. Butenweg C, Marinković M, Salatić R. Experimental results of reinforced concrete frames with masonry infills under combined quasi-static in-plane and out-of-plane seismic loading. Bull Earthq Eng. 2019;17(6):3397–422. https://doi.org/10.1007/s10518-019-00602-7.
6. Casali Y, Yonca NA, Comes T, Casali Y. Machine learning for spatial analyses in urban areas: a scoping review. Sustain Cities Soc. 2022;12:104050.
7. Chegini H, Naha RK, Mahanti A, Thulasiraman P. Process automation in an IoT–fog–cloud ecosystem: a survey and taxonomy. IoT. 2021;2(1):92–118.
8. Feynman RP. Quantum mechanical computers. Optics News. 1985;11(2):11–20.
9. Flynn MJ, Mencer O, Milutinovic V, Rakocevic G, Stenstrom P, Trobec R, Valero M. Moving from petaflops to petadata. Commun ACM. 2013;56(5):39–42.
10. Furht B. NSF-funded traineeship in data science technologies and applications. Boca Raton: Florida Atlantic University; 2022.
11. Ilić MD. Dynamic monitoring and decision systems for enabling sustainable energy services. Proc IEEE. 2010;99(1):58–79.
12. Ilic MD. Toward a unified modeling and control for sustainable and resilient electric energy systems. Found Trends Electr Energy Syst. 2016;1(1–2):1–41.
13. Ilic MD, Lessard DR. A distributed coordinated architecture of electrical energy systems for sustainability. https://www.dropbox.com/s/2s4bgcr4bympq3b/Ilic_Lessard_EESGatMITWP_dec302020%20%20Copy.pdf?dl=0 2022.
14. Hui S, Zak SH. Discrete Fourier transform based pattern classifiers. Bull Polish Acad Sci Tech Sci. 2014;62
15. Kotlar M, Milutinovic V. Comparing controlflow and dataflow for tensor calculus: speed, power, complexity, and MTBF. In High Performance Computing: ISC High Performance 2018 International Workshops, Frankfurt/Main, Germany, June 28, 2018, Revised Selected Papers 33 2018 (pp. 329–346). Springer International Publishing.
16. Laudon KC, Laudon JP. Essentials of management information systems. London: Pearson; 2015.
17. Marinković M, Butenweg C. Innovative decoupling system for the seismic protection of masonry infill walls in reinforced concrete frames. Eng Struct. 2019;197:109435.
18. Marinković M, Baballëku M, Isufi B, Blagojević N, Milićević I, Brzev S. Performance of RC cast-in-place buildings during the november 26, 2019 Albania earthquake. Bull Earthq Eng. 2022;20(10):5427–80. https://doi.org/10.1007/s10518-022-01414-y.
19. Mesinger F, Veljovic K. Topography in weather and climate models: lessons from cut-cell eta vs. European centre for medium-range weather forecasts experiments. J Meteorol Soc Japan. 2020;98(5):881–900.
20. Mulligan K. Computationally networked urbanism and advanced sustainability analytics in internet of things-enabled smart city governance. Geopolit History Int Relat. 2021;13(2):121–34.
21. Milutinovic D, Milutinovic V, Soucek B. The honeycomb architecture. IEEE. Computer. 1987;20(4):81–3.
22. Milutinovic VE. Mapping of neural networks on the honeycomb architecture. Proc IEEE. 1989;77(12):1875–8.
23. Milutinovic V, Kotlar M, Stojanovic M, Dundic I, Trifunovic N, Babovic Z. DataFlow supercomputing essentials. Cham: Springer; 2017.
24. Milutinovic V, Kotlar M. Handbook of Research on Methodologies and Applications of Supercomputing. IGI Global; 2021. 393 p
25. Ngom A, Stojmenovic I, Milutinovic V. STRIP - a strip-based neural-network growth algorithm for learning multiple-valued functions. IEEE Trans Neural Networks. 2001;12(2):212–27.
26. Oh C, Zak SH. Large-Scale Pattern Storage and Retrieval Using Generalized Brain-State-in-a-Box Neural Networks. IEEE Trans Neural Networks. 2010;21(4):633–43
27. Predari G, Stefanini L, Marinković M, Stepinac M, Brzev S. Adriseismic methodology for expeditious seismic assessment of unreinforced masonry buildings. Buildings. 2023. https://doi.org/10.3390/buildings13020344.
28. Sengupta J, Kubendran R, Neftci E, Andreou A. High-Speed, Real-Time, Spike-Based Object Tracking and Path Prediction on Google Edge TPU. In 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS). 2020. p. 134–5.
29. Shechtman D. Education for Entrepreneurship. The Keynote Speech, at the 30th IEEE TELFOR Conference, November 15 and 16, Belgrade, Serbia, 2022; 1–2.
30. Simoes R, Camara G, Queiroz G, Souza F, Andrade PR, Santos L, et al. Satellite image time series analysis for big earth observation data. Remote Sens. 2021. https://doi.org/10.3390/rs13132428.
31. Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S, Brisco B. Google earth engine for geo-big data applications: a meta-analysis and systematic review. ISPRS J Photogramm Remote Sens. 2020;164:152–70.
32. Trifunovic N, Milutinovic V, Salom J, Kos A. Paradigm shift in big data supercomputing: dataflow vs. controlflow. J Big Data. 2015;2:1–9.
33. Trifunovic N, Milutinovic V, Korolija N, Gaydadjiev G. An appgallery for dataflow computing. J Big Data. 2016;3:1–30.
34. Vázquez F, Fernández JJ, Garzón EM. A new approach for sparse matrix vector product on NVIDIA GPUs. Concurr Comput Pract Exp. 2011;23(8):815–26. https://doi.org/10.1002/cpe.1658.
35. Wang Z, Zhu Z, Xu M, Qureshi S. Fine-grained assessment of greenspace satisfaction at regional scale using content analysis of social media and machine learning. Sci Total Environ. 2021;776:145908.

36. Wentz EA, Anderson S, Fragkias M, Netzband M, Mesev V, Myint SW, et al. Supporting global environmental change research: a review of trends and knowledge gaps in urban remote sensing. Remote Sens. 2014;6(5):3879–905.

37. Wu SD, Kempf KG, Atan MO, Aytac B, Shirodkar SA, Mishra A. Improving new-product forecasting at intel corporation. Interfaces. 2010;40(5):385–96. https://doi.org/10.1287/inte.1100.0504.

Publisher's Note