Evalu@: An Agnostic Web-Based Tool for Consistent and Constant Evaluation Used as a Data Gatherer for Artificial Intelligence Implementations

Fernando Yepes-Calderon[1,2(✉)], Juan F. Yepes Zuluaga[1,2], and Gonzalo E. Yepes Calderon[1]

[1] GYM Group SA, Carrera 78A No. 6-58, Cali, Colombia
fernando@gym-group.org
[2] Strategic Business Platforms,
417 Heritage Park Tr Nw, Kennesaw, GA 30144, USA

Abstract. Evalu@ is a software development created under the model-view-controller pattern and is meant to be executed in a client-server architecture. It is benefited from the worldwide coverage of the Internet and acts as an evaluating gadget and a data centralizer. Evalu@ is initially conceived as a solution to the lack of assistant tools while running the quality programs in industrial environments. Later, due to its high degree of generalization in the setup of evaluations schemes, the software was successfully flavored to suit the willingness of entrepreneurs in other fields. Recently, some Machine Learning features have been added and are being tested to close the monitoring cycle by not only keeping track of the evaluation items chronologically; but also being capable of classifying and predicting outcomes based on previously gathered data.

Keywords: Data analysis · Artificial intelligence · Research-industry gap diminishing

1 Introduction

In our highly competitive world, commercial and personal work-related activities are pushed to excellence. This tendency seems to be reasonable in a capitalist environment where quality in the services and currency are equally appreciated [11]. However, how to define quality and moreover, how to accomplish it?

During the last century, several regulating organizations were created around the world to provide elements that can lead companies to proceed with high quality. Initially, the principal goal was to provide normalization so all the entities in a particular field not only could speak in the same terms but also, they could have a subtle agreement regarding the form of the delivered, whether it is a service or a product [12]. As everyone wants to have the best for a price, rapidly, the quality initiative became global. Consequently, the International Federation

H. Florez et al. (Eds.): ICAI 2019, CCIS 1051, pp. 73–84, 2019.
https://doi.org/10.1007/978-3-030-32475-9_6

of the National Standardizing Association (ISA) was created in 1926, dissolved during the second world war and re-established as the International Organization for Standardization (ISO) [8]. The ISO defines a group of statements that nowadays go beyond the primary purpose providing the bases for a clear understanding of quality with global traceability. Within the ISO methods, the 'ways of doing' that guarantees clients' satisfaction are defined, while assuring people's integrity and healthy conditions in the creation of a product or a service. Despite the worldwide acceptability of ISO, it lacks a consistent-technological assisting tool that facilitates the application of standards [18]. Designing a tool with enough flexibility to follow a highly dynamic environment that supposes the application of standards in diverse production fields is cumbersome. To assure the desired flexibility, the presented solution breaks the paradigm of *customized solutions* [15] used by software developers, and introduces the agnostic-black-box (ABlaBo) concept. In the AblaBo, the solutions go from the simple to the complex and get perfectly fitted to particularities while the users interact more time with the tool.

This manuscript presents an Internet-based service available in www.evalualos.com that uses ABlaBo to assist users in the tasks of evaluation, monitoring through time, generation of customized quality criteria, easy visualization, generation of predictions, unlimited creation of indexes and implementation of programmed feedback among many other operative details. The tool is designed to cover all kind of evaluations and to keep historical records of the results chronologically organized. We also present some success cases and introduce the applicability of machine learning (ML) techniques to provide prediction capabilities. With this particular feature, in addition to empowering the tool with an outstanding added value, we shorten the gap between research and real-world applications, while giving to final users the possibility to predict the behavior of their evaluating items; an aspect that traduces to saving resources and capital.

2 Materials and Methods

2.1 Software Designing Pillars and Implementation

2.1.1 EV-Boxes Creation (Containers of e-Items)

In order to provide the desired flexibility so that any group of entities (people or objects) could be evaluated, the design exploits the benefits of the persistence engine by mapping the forms entries into the database. The form-entry mapping is a standard solution used in successful applications like Moodle [3,4] and also in popular content management systems like Joomla or Drupal [10]. The user also defines this form entries through configuration files written in Excel (CSV) files, which gives several advantages including; structural simplicity, reduced size, cross-platform usability, and off-line planning [17]. Besides, in Evalu@, and willing to avoid database transaction overheat, the evaluating items' *(e-items)* descriptors and evaluations *(e-surveys)* are saved as a string containing array-like syntax as generalized in Eq. 1. Furthermore, this text pattern is exploded

to be presented to the final users in an organized manner. Also, the users are not limited in any aspect regarding the configuration of e-items and e-surveys. Shortly, the application can host several e-item boxes, where unlimited e-items can be held, and these e-items can be evaluated with unlimited e-survey profiles. No matter how complex the setup scheme is, Evalu@ will continue being fully operative; from here, its agnostic nature. See in Fig. 1 two configuration files to create e-item boxes.

Item to evaluate	Descriptor	ages: photo. For date	Option 1	Option 2	Option 3	Option 4	Option 5
Extinguisher							
A	Serial	IRP-E002					
	Brand	Badger	Ecosan	ABC	Amerex	Zemer	
	Date of purchase	DD/MM/YYYY					
	Capacity	1 kg	2.5 kg	5 Kg	10 Kg		

Item to evaluate	Descriptor	ages: photo. For date	Option 1	Option 2	Option 3	Option 4	Option 5	Option 6
Senior Soccer								
B	Name	Daniel						
	Lastname	Reina						
	DOB	DD/MM/YYYY						
	Position	Goalkeeper	Defense	Midfielder	Attacker			
	Laterality	Right	Left	Ambidextrous				
	Schoolarity	None	Primary	Secondary	BS	MS	MD/PhD	

Fig. 1. Excel files to create two different e-item boxes. In panel A, the e-item box to grouping extinguishers. In panel B the e-item box collects soccer players. Even when the mechanisms of this flexibility are bulky, the system does not create database-transactional overheat

$$F(j) = \begin{cases} [key_j : value_j] & \text{if value exist} \forall key_j \in N \\ Continue & otherwise \end{cases} \quad (1)$$

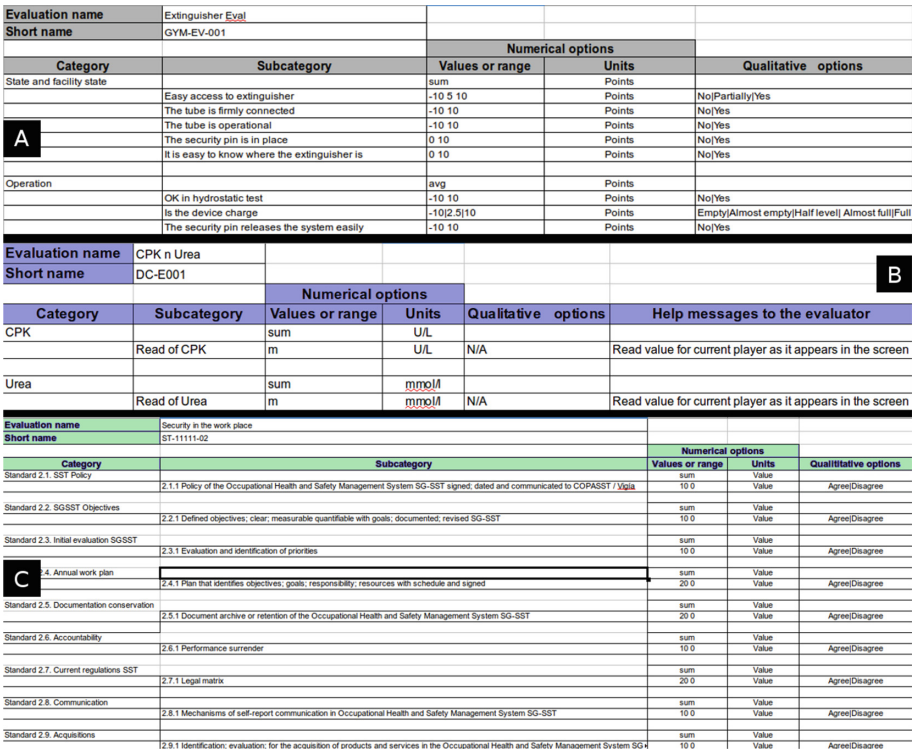
Where Key_j replaces a descriptor or an evaluation criteria, while $value_j$ holds the corresponding value.

The average time to measure over 100 repetitions querying an entry of $j = 8$ fields organized in the array-like fashion, was of 0.1448 ms. The same exercise with an equivalent data persistence using separated fields (rows) per entry, resulted in 0.1510 ms. Although the time is not considerably different, the used strategy will keep the record in one line, regardless of how many descriptors or evaluation elements are required for a given e-item box. These timing experiments are performed directly at the command line in a Mysql server (Dell, i7 processor) running Ubuntu Linux 16.04 so web latencies timing variations associated with accessing remote database servers are avoided.

2.1.2 Evaluations (e-Surveys) Creation

Regarding the e-surveys, generalization in the construction of the tests is accomplished based on the same mechanism used in the e-items box creation. Several e-survey criteria can be nested in an unlimited category-subcategory structure; these criteria will have a value in the e-survey that is quantitative by nature

but can also be presented to final users in qualitative form. Also during the configuration of the e-surveys, the user might define a range of possible values for each criterion by explicitly defining an array of values or implicitly by defining a range in the form *[initial value; step; final value]*. In both cases, the values will be presented in a select box – an HTML-form control – that will be placed in front of every evaluating criteria. The user has one more option regarding the configuration values. When placing an "m" character at the values field in the numeric options (see Fig. 2 **- Panel B**), the system will interpret the entry as used define option, and instead of posting a select box, the HTML-form control will be an input text.



Evaluation name	Extinguisher Eval			
Short name	GYM-EV-001			
		Numerical options		
Category	Subcategory	Values or range	Units	Qualitative options
State and facility state		sum	Points	
	Easy access to extinguisher	-10 5 10	Points	No|Partially|Yes
	The tube is firmly connected	-10 10	Points	No|Yes
	The tube is operational	-10 10	Points	No|Yes
	The security pin is in place	0 10	Points	No|Yes
	It is easy to know where the extinguisher is	0 10	Points	No|Yes
Operation		avg	Points	
	OK in hydrostatic test	-10 10	Points	No|Yes
	Is the device charge	-10|2.5|10	Points	Empty|Almost empty|Half level| Almost full|Full
	The security pin releases the system easily	-10 10	Points	No|Yes

A

Evaluation name	CPK n Urea				
Short name	DC-E001				
		Numerical options			
Category	Subcategory	Values or range	Units	Qualitative options	Help messages to the evaluator
CPK		sum	U/L		
	Read of CPK	m	U/L	N/A	Read value for current player as it appears in the screen
Urea		sum	mmol/l		
	Read of Urea	m	mmol/l	N/A	Read value for current player as it appears in the screen

B

Evaluation name	Security in the work place			
Short name	ST-11111-02			
		Numerical options		
Category	Subcategory	Values or range	Units	Qualititative options
Standard 2.1. SST Policy		sum	Value	
	2.1.1 Policy of the Occupational Health and Safety Management System SG-SST signed; dated and communicated to COPASST / Vigia	10 0	Value	Agree|Disagree
Standard 2.2. SGSST Objectives		sum	Value	
	2.2.1 Defined objectives; clear; measurable quantifiable with goals; documented; revised SG-SST	10 0	Value	Agree|Disagree
Standard 2.3. Initial evaluation SGSST		sum	Value	
	2.3.1 Evaluation and identification of priorities	10 0	Value	Agree|Disagree
Standard 2.4. Annual work plan		sum	Value	
	2.4.1 Plan that identifies objectives; goals; responsibility; resources with schedule and signed	20 0	Value	Agree|Disagree
Standard 2.5. Documentation conservation		sum	Value	
	2.5.1 Document archive or retention of the Occupational Health and Safety Management System SG-SST	20 0	Value	Agree|Disagree
Standard 2.6. Accountability		sum	Value	
	2.6.1 Performance surrender	10 0	Value	Agree|Disagree
Standard 2.7. Current regulations SST		sum	Value	
	2.7.1 Legal matrix	20 0	Value	Agree|Disagree
Standard 2.8. Communication		sum	Value	
	2.8.1 Mechanisms of self-report communication in Occupational Health and Safety Management System SG-SST	10 0	Value	Agree|Disagree
Standard 2.9. Acquisitions		sum	Value	
	2.9.1 Identification; evaluation; for the acquisition of products and services in the Occupational Health and Safety Management System SG-S	10 0	Value	Agree|Disagree

C

Fig. 2. Excel files to create two different e-surveys. In panel A, An evaluation designed to grade extinguishers. In panel B, an evaluation to follow metabolic variables that may be used in the senior soccer players shown in Fig. 1 **- panel B**. In panel C, an evaluation that follows the literal 2 of the 1111 standard (SG-SST) acronym of "Sistema de gestion, seguridad y salud en el trabajo"

2.2 Association of e-Item Boxes and e-Surveys

Since the e-item boxes and e-surveys are independently created, one can associate an e-survey to several e-item boxes accomplishing re-usability. Another consequence of this assignment freedom, an e-item box can have several e-surveys associated, flexibility appreciated in environments where the e-items are evaluated in more than one aspect, and those aspects need to be analyzed independently (Fig. 3).

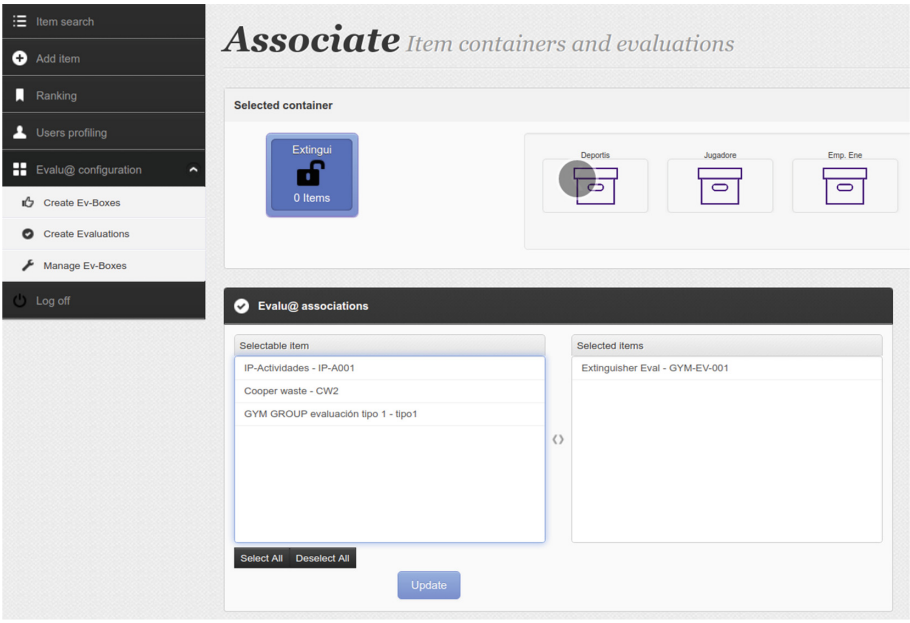


Fig. 3. Association of e-item boxes and evaluations. Here the e-item box of extinguishers is about to be linked to the extinguishers' e-survey that was previously uploaded to the system using the excel template depicted in Fig. 2 **- Panel A**

2.3 Interoperability, and Data Sharing

Evalu@ has a simple but useful module that complies with Remote Procedure Call (RPC) [2]. The protocol to share the information was chosen to be JSON due to both, its transparent translation from native programming objects and its simplicity of treatment in the client console [7]. Through the JSON-RPC specification, other applications may be benefited from the tracking capabilities of Evalu@ and produce their reporting schemes. By the moment of writing this document, Evalu@ can share JSON chains holding individual e-item profiles, including their descriptors and historical records of evaluation. Bulky JSON sharing capabilities are also supported. Besides, Evalu@ can be operated in smartphones and tablets with a responsive cascade style sheet (CSS) design. Also, e-surveys can be posted publicly using permalinks in the evalu@ site and iframes in external URLs [6].

2.4 Classification and Prediction Capabilities

Evalu@ has been given classification and prediction capabilities. These specifications are based on support vector machines algorithm (SVM) [16]. While trying to obtain substantial information from unknown data, one might feed the SVM algorithm with features that do not provide any differentiation power. To avoid the proliferation of useless features, a step of feature reduction is accomplished through an F-test. After the F-test, the K best features are selected and feed to the SVM algorithm. The K value depends on the number of e-items available (not in the system, but the grouping box). This K to number of e-items (N_e) relation is defined by $K = ceil(N_e * exp(-4))$. In this way, the machine learning strategy is always well-formulated since there will always be more samples than features. If the system does not have enough e-items to satisfy the formulation of K, the SVM feature is not offered in the user interface.

2.5 Fields of Applicability

At the moment of writing this manuscript, two fields of applicability that have nothing in common are in use of the system, something that remarks the excellent level of generalization reached by Evalu@ and make us confident of its acceptability and usability. In a further stage, other application fields will be targeted.

2.5.1 Health and Security in the Workplace

Security in the workplace is an essential aspect for companies worldwide and is regulated by the ISO. Despite the directives are self-explanatory, and there is no doubt about the concepts, the applicability is still a pending asset. With the flexibility of Evalu@, an e-item box can be created to group all the e-items by function and then, create different evaluation profiles for each e-item box. Having this centralized information benefits the companies in many aspects, including online management, live reporting, remote monitoring, and the confidence of having the system under control. Also, the prediction capabilities provided through the ML implementation in mode "supervised classification" can suggest when an e-item should be replaced, taken into maintenance or will have high/low performance in case the evaluation is performed on people.

2.5.2 Information Athletes

The process of formation in sports follows some qualification aspects that are strongly associated with the ideology of trainers. Some theory says that players must be evaluated according to the purposes of the microcycles and mesocycles [5] other more pragmatic approaches pretend to form the players in their physical, technical, and overall, mental capabilities [13]. Even though these schemes are associated with soccer, other disciplines are adopting them to educate their prospects. No matter what the approach is, the evaluating stage is of outstanding importance because it says not only the state of the players but their evolution.

By now, coaches and trainers are aware of the state of the players, but few of them can say something about their evolution without incurring in qualitative statements biased by personal perceptions [9]. Even worse, players usually ignore the plan, how they are being evaluated, and their performances according to the evaluation criteria. All these problems are solved with the data centralization and chronological reporting capabilities provided by Evalu@. In this field, the administrators are benefited from the non-supervised classification function, where groups are automatically created among the whole data accordingly to the factors of evaluation [1]. This automatic grouping can be later correlated with overall performance and thus create training plans targeting particular groups necessities, something practically more feasible than individualizing the training.

2.5.3 High-Performance Athletes and Scouting

High-performance athletes of almost any discipline are subject to a market that is continuously looking for excellence. The selection process uses scouts that travel around the world with excel sheets [14]. Then, the information is centralized, and decisions are made with the stored information. This process is expensive, sluggish, and susceptible to human mistakes. With Evalu@, the information is updated instantly, and it is available for administrators and collaborators in the act. Grading responsibilities are evident even in grading sessions that happened time ago. Collaborators can be in any part of the globe as well as the e-items. The administrators are provided with tools compare players on a timing axis to include concepts of regularity in their decisions. Ranking capabilities are also included, and the final user can build filters to profile the ideal player and to create the ranking according to those preferences. The prediction capabilities in this matter are used to define when a prospect is ready to jump to first divisions leagues and when, according to his/her historical performance, it would be better to start the marketing tasks.

3 Results

In this section, some success cases that are currently running are presented. As the system continues gathering the data, the prediction capabilities and other AI-based functions will be enabled.

3.1 Industrial Security

3.1.1 Security Network

A beta-testing stage is currently being executed at a company in the agronomy field that produces sugar as a primary activity and has run diversification tasks to produce paper and fuel. The grouping items in this particular environment consist of fire extinguishers, hydrants, emergency beds, and ambulances. The creation of these boxes was accomplished by the security head officer using the configuration built-in wizard. The company is currently running evaluations

Fig. 4. Evolution in time of an e-item in the extinguishers box. The results correspond to evaluations performed periodically using the e-survey depicted in Fig. 2 **- Panel A**

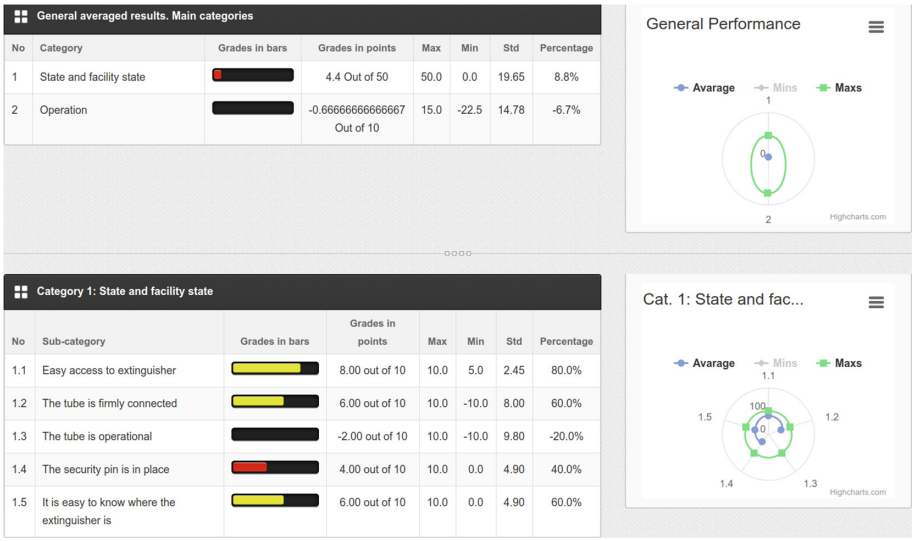


Fig. 5. Part of the average analysis. The data is generated automatically with the information individually shown in Fig. 4

using Evalu@ to certify his security network. See Fig. 4 for extinguishers profiling and evolution in time.

For companies maintaining a security network, it is crucial to report the average analysis as it is shown in Fig. 5. This analysis is provided as a default function by the software.

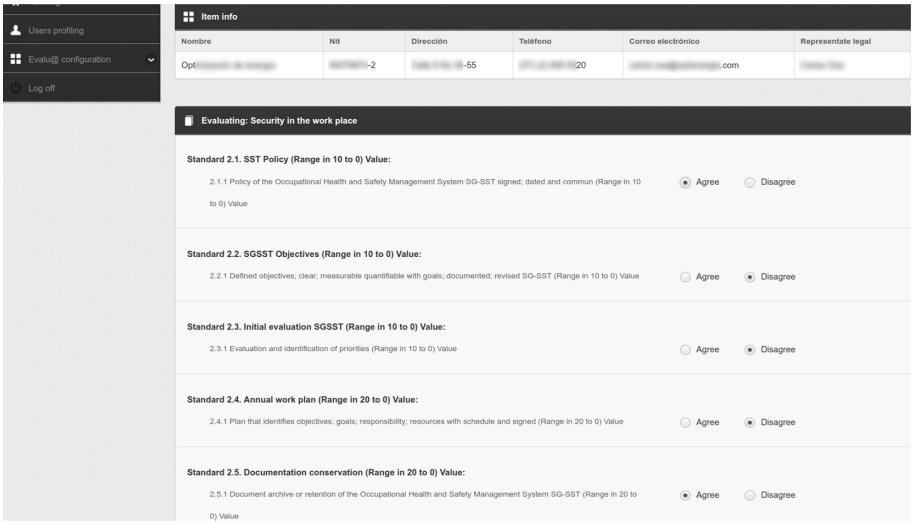


Fig. 6. Evalu@'s inputting form for the e-survey depicted in Fig. 2 **- Panel C**

3.1.2 Complying with the SG-SST Regulation

Due to the flexibility delivered by Evalu@, it is used in the same company to comply with the regulation SG-SST. The Fig. 6 is a pragmatic evidence of one the uses listed in Sect. 2.5.1.

Within Evalu@, analyzing the strong and weak points of the company, and how the correction measurements impact the performance is easily accomplished by just glancing at the evolution in time report depicted in Fig. 7. Note how the agnostic nature of the software is evidenced.

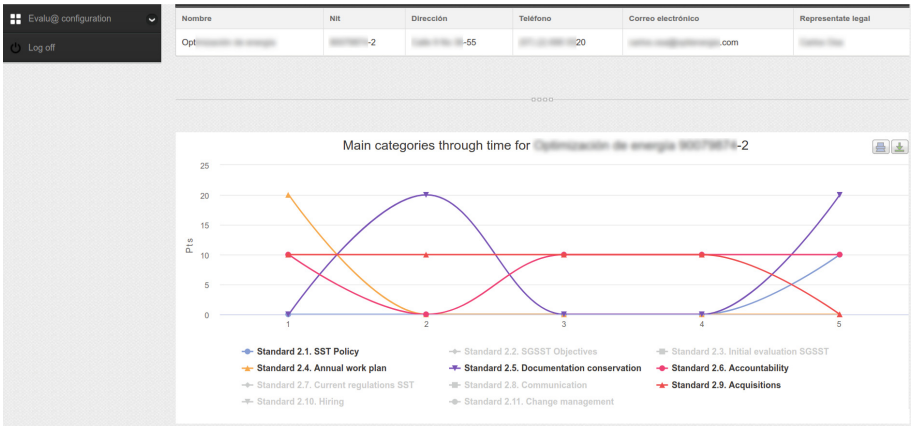


Fig. 7. SG-SST evolution in time for the literal 2 of the standard "Integral operation".

3.2 Sports Scouting

Evalu@ has already passed the first round of tests in an international scouting company where the goal is to have a record of high-performance soccer players. Specific modules such as the versus visualization and the multifeatured ranking help the scouting company to find the right player for a demanding team. Ideally, this activity requires extensive coverage, considering that good players can be anywhere in the world. In this matter, the operation was far away from the desired due to operational problems linked to data transport and data reliability. Now, this aim is feasible due to the inherent global coverage of Evalu@. See the use of the versus mode in Fig. 8.

Fig. 8. The versus module has been specifically designed to allow direct comparison of e-items among the e-survey criteria. Some information has been intentionally blurred to protect the evaluation criteria as requested by the company administrator.

4 Discussion

As a response to a recurrent industrial requirement, we have designed and created Evalu@. This system is a powerful tool that reaches a high degree of generalization. The keystone of Evalu@ is strongly related to is flexibility; that makes this solution usable in different fields of application. This tool is still being developed, but its modularity provided by a model-view-controller assures that further development will not affect the correct operation of previously launched specifications. Additionally, Evalu@ provides standard interoperability by implementing the RPC specification. Through the RPC, a simple coded request to evalu@ will return a complete JSON structure that can be used for creating new flavored angles of the same application. An ML implementation is envisaged, and we are strategically collecting more data. The ML approach represents a good step towards narrowing the gap between research and implementation. Evalu@ is being tested in two companies at the moment. They both see potential in this tool for data centralization that will undoubtedly improve their indexes of operation. Regarding success case 1, this is something that saves significant amounts of money currently invested in quality assurance tasks but moreover, the personal costs when the security network fails. Regarding success case 2, the use of this tool increases scouting accuracy that is traduced in credibility, thus profitability. In the following, our team will continue developing associated modules and target the field of athletes in minor divisions, a field that has been already contacted and from which good acceptability has been received.

5 Conclusions

Because of the high coverage provided by the Internet and its associated programming platforms, several applications have migrated from desktop to the browser environments, where a vast range of services are available. Other modern gadgets such as tablets and smartphones facilitate the data generation in almost any daily situation. Data has become massive, difficult to interpret in bulky fashion, but still meaningful. Our job consists of turning this information into valuable leads for companies, that is understandable to final users and provides new features that can assist humans in the decision-making scenario. Evalu@ has been created with all these principles in mind, assisting, to best of our knowledge, an unmet field. Further extensions include the development of hardware gadgets to increase off-line usability. Once data is centralized, AI gadgets can be developed to contribute to the creation of generalized knowledge.

References

1. Barshan, B., Yüksek, M.C.: Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units. Comput. J. **57**, 1649–1667 (2014)

2. Birman, K.P.: Remote procedure calls and the client/server model. In: Birman, K.P. (ed.) Guide to Reliable Distributed Systems. Texts in Computer Science, pp. 185–247. Springer, London (2012). https://doi.org/10.1007/978-1-4471-2416-0_6
3. Chourishi, D., Buttan, C.K., Chaurasia, A., Soni, A.: Effective e-learning through moodle (2011)
4. Huang, C.-C., Wang, Y.-M., Wu, T.W., Wang, P.A.: An empirical analysis of the antecedents and performance consequences of using the moodle platform. Int. J. Inf. Educ. Technol. **3**(2) (2013). https://doi.org/10.7763/IJIET.2013.V3.267
5. Comfort, P., Matthews, M.: An Introduction to Periodisation. Wiley, Hoboken (2010)
6. Florencias-Oliveros, O., et al.: Real-life power quality transients (2017). https://doi.org/10.21227/H2Q30W
7. Ghosh, D., Sheeh, J., Thorup, K.K., Vinoski, S.: Programming language impact on the development of distributed systems. J. Internet Serv. Appl. **3**(1), 22–30 (2012)
8. Heires, M.: The international organization for standardization (ISO). New Polit. Econ. **13**(3), 357–367 (2008). https://doi.org/10.1080/13563460802302693
9. Hvistendahl, J.: The effect of placement of biasing information. J. Q. **43**, 647–654 (1966)
10. Patel, S.K., Rathod, V.R., Prajapati, J.B.: Performance analysis of content management systems- Joomla, Drupal and WordPress. Int. J. Comput. Appl. **21**(4), 39–43 (2011)
11. Lambin, J.J.: Capitalism and sustainable development. SYMPHONYA Emerg. Issues Manag. **2**, 3–9 (2009)
12. Renard, M.-C.: Quality certification, regulation and power in fair trade. J. Rural. Stud. **21**, 419–431 (2005)
13. Mathieu, J.E., Heffner, T.S., Heffner, T.S., Salas, E., Cannon-Bowers, J.A.: The influence of shared mental models on team process and performance. J. Appl. Psychol. **85**, 273 (2000)
14. Moore, P.: Scouting an antrhopology of sports. Anthropologica **46**, 37–46 (2004)
15. Müller, J., Krüger, J., Enderlein, S., Helmich, M., Zeier, A.: Customizing enterprise software as a service applications: back-end extension in a multi-tenancy environment. In: Filipe, J., Cordeiro, J. (eds.) ICEIS 2009. LNBIP, vol. 24, pp. 66–77. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01347-8_6
16. Pedregosa, F., et al.: Scikit-learn: machine learning in python. J. Mach. Learn. Res. **12**, 2825–2830 (2011). http://dl.acm.org/citation.cfm?id=1953048.2078195
17. Shafranovich, Y.: Common format and MIME type for comma-separated values (CSV) files. RFC **4180**, 1–8 (2005)
18. Zeng, S.X.: Overcoming barriers to sustainable implementation of the ISO. Manag. Audit. J. **22**, 244–254 (2007)