Information Systems in Management
Primary version of the journal is an electronic version.

Editor
Department of Informatics, Warsaw University of Life Sciences – SGGW

Editorial Committee
Dr hab. inż. Arkadiusz Orłowski – Editor-in-Chief
Dr Piotr Łukasiewicz – Executive Editor
Dr inż. Tomasz Ząbkowski – Scientific Editor
Prof. nadzw. Kris Gaj – Linguistic Editor
Dr hab. Wiesław Szczesny – Statistical Editor

Editorial Council
Dr Oguz Akpolat – Mugla Sitki Kocman University, Turkey
Prof. dr hab. inż. Ryszard Budziński – Uniwersytet Szczeciński
Prof. dr hab. Witold Chmielarz – Uniwersytet Warszawski
Dr hab. inż. Leszek Chmielewski – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Prof. Jeretta Horn Nord – Oklahoma State University, USA
Prof. Frederick G. Kohun – Robert Morris University, USA
Prof. Yuiry Kondratenko – Black Sea State University, Ukraina
Prof. Alex Koohang – Middle Georgia State College, USA
Prof. Vassilis Kostoglou – Alexander Technological Educational Institute of Thessaloniki, Greece
Prof. dr hab. Marian Niedźwiecki – Uniwersytet Łódzki
Dr hab. inż. Arkadiusz Orłowski – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie – Chairman
Dr hab. inż. Joanna Paliszkiewicz – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Prof. Kongkiti Phusavat – Kasetsart University Bangkok, Thailand
Prof. Josu Takala – University of Vaasa, Finland
Dr hab. inż. Ludwik Wicki – Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Prof. dr hab. inż. Antoni Wiliński – Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

Address of the Editor
Faculty of Applied Informatics and Mathematics, WULS – SGGW
ul. Nowoursynowska 166, 02-787 Warszawa, Poland
e-mail: isim@sggw.pl, www.ism.wzim.sggw.pl

ISSN: 2084-5537
e-ISSN: 2544-1728

Wydawnictwo SGGW
ul. Nowoursynowska 166, 02-787 Warszawa, Poland
e-mail: wydawnictwo@sggw.pl, www.wydawnictwosggw.pl

<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waldemar Karwowski, Piotr Wrzeciono</td>
<td>227</td>
</tr>
<tr>
<td>TOWARDS THE DATA STRUCTURE FOR EFFECTIVE WORD SEARCH</td>
<td></td>
</tr>
<tr>
<td>Piotr Korneta</td>
<td>237</td>
</tr>
<tr>
<td>STOCK MANAGEMENT AS A CRITICAL SUCCESS FACTOR FOR POLISH CROP PRODCERS</td>
<td></td>
</tr>
<tr>
<td>Roman Pietroń</td>
<td>246</td>
</tr>
<tr>
<td>MICRO- VS. MACRO SCOPE SIMULATION IN BPM MODELLING</td>
<td></td>
</tr>
<tr>
<td>Paulina Rutecka, Karina Cicha</td>
<td>258</td>
</tr>
<tr>
<td>CUSTOMER’S TRUST MANAGEMENT IN POLISH E-COMMERCE – CASE STUDY</td>
<td></td>
</tr>
<tr>
<td>Waldemar Sobiecki, Anna Kurzydlowska</td>
<td>269</td>
</tr>
<tr>
<td>ENHANCEMENT OF SCRUM-BASED PROJECT MANAGEMENT LEARNING EXPERIENCE BY USING WEB APPLICATION</td>
<td></td>
</tr>
<tr>
<td>Tomasz Woźniakowski, Magdalena Nowakowska, Rafik Nafkha</td>
<td>286</td>
</tr>
<tr>
<td>CONTEXT-SOCIAL MODEL OF PROSUMPTION IN E-COMMERCE – ANALYSIS OF A PROTOTYPE SOLUTION FOR CLOTHING INDUSTRY</td>
<td></td>
</tr>
</tbody>
</table>
TOWARDS THE DATA STRUCTURE FOR EFFECTIVE WORD SEARCH

WALDEMAR KARWOWSKI, PIOTR WRZECIONO

Department of Informatics, Warsaw University of Life Sciences (SGGW)

In the paper problem of searching basic forms for words in the Polish language is discussed. Polish language has a very extensive inflection and effective method for finding base form is important in many NLP tasks for example text indexing. The method for searching, based on open-source dictionary of Polish language, is presented. In this method it is important to design a structure for storing all words from dictionary, in such a way that it allows to quickly find basic words forms. Two dictionary structures: ternary search tree and associative table are presented and discussed. Tests are performed on the six actual and three crafted artificial texts and results are compared with other possible dictionary structures. At the end conclusions about structures effectiveness are formulated.

Keywords: Text Indexing, Associative Array, Ternary Search Tree

1. Introduction

Today, English is the dominant language not only in the Internet but also in many areas, primarily connected with science and technology. However, at the same time, the national languages are commonly used and generally the number of published documents is increasing. A huge number of documents prevents person from viewing them directly. Automatic analysis of the texts can help in reviewing big sets of documents. Especially text indexing is an important issue today, it helps to organize documents and it is widely applicable in knowledge management. During indexing, the very important task is to reduce inflectional and derivationally
related forms of a word to a common base form. The process of reducing words to their base or root form is possible with stemming algorithms. The Polish language has a very extensive inflection and stemming is not an easy task in this situation. Effective stemming algorithms are designed for English, but they are not useful in Polish. Review of stemmers for Polish language is made in [10]. There are available two stemmer tools: Stempel [9] - algorithmic stemmer for Polish language and Morfologik [7] - morphosyntactic dictionary for the Polish language, which has the stemmer tool. The problem of automatic text indexing in agriculture domain, for documents in Polish language, was discussed by authors in [3, 4, 11]. Because dictionaries in Morfologik or Stempel were not big and integration with our system was not easy, we designed special custom method. We have to note that generally stemming extracts rather root that base form, but for our goals we need very simple stemming just finding base form for word. Our method involves use an open-source dictionary of Polish language [8]. This dictionary is an open project licensed under the GPL and CC SA licenses. It is continuously updated and contains more than 200000 records. Every record consists of base form and all inflected forms derived from it. We prepared collection of all words with pointers to base form; as a result, we can easily find the base form for any word. The basic problem is effective search for words in the collection, because we have together almost 4 million words. Simple method involves storing sorted words in a huge array together with pointers to base form and using the binary search. This method was applied in [3, 4, 11], it is memory consuming but works quite well. To improve searching we need more specialized structures than sorted array. Dictionary structure based on a trie was presented and tested in [5]. The primary disadvantage of trie structure is huge amount of needed RAM memory. In this paper we want to present two other structures: ternary search tree (TST) and custom dictionary based on associative array, which requires much less memory than trie and at the same time ensures relatively high efficiency. Main goal of the research was to implement, test and compare search efficiency for all mentioned structures. The results allow us to design and implement the best dictionary for the indexing task.

The rest of this paper is organized as follows: in Sect. 2 the dictionary structure based on a ternary search tree is presented and discussed. In Sect. 3 the associative array is shortly described. In Sect. 4 subsequently testing results for sorted list, trie, TST and associative array based on hash table are presented and discussed. Summary and final remarks are formulated in Sect. 5.

2. Dictionary structure based on a ternary search tree

In [5] we presented that a subsequent search of words in a sorted array, which contains almost four million items, takes some time, but the structure based on the trie makes the same task many times faster. The main disadvantage of structure
based on trie is its huge size in memory. The ternary search tree is a kind of search tree where nodes are arranged in a manner similar to a binary search tree, but with three children: left, middle and right. Similarly to trie, each node in the TST stores only one character (a part of key) [1]. The left child stores character value which is less than the character in the current node. The right child stores character which is greater than the character in the current node. Middle child stores next character in the word. Additionally each node stores flag which denotes possible end of word. In other words each node in a ternary search tree represents a prefix of the stored strings. All strings in the middle subtree of a node start with that prefix.

![Ternary Search Tree Diagram](image)

**Figure 1. Words with the common prefix in TST**

In the Fig. 1 ternary search tree for words: z, ziąb, ziem, ziemí, ziemia, ziemią, ziemniak, ziemniaka, ziemniaki, zim, and zima, is presented. The thick line is used for nodes which are the end of the word. For example node e represents prefix zie, node n represents prefix zienn. Because we compare Unicode values of characters,
value of the letter q (0105 in hex) is bigger than value of the letter e (0065 in hex) what we can see in the Figure 1. We can compare unicodes because in searching, unlike in sorting, order of letters in the Polish alphabet is not important. Moreover we have to note that TST unlike trie depends on order of adding words. The dictionary based on TST allows to store big and small letters, which make possible to distinguish own names. In structure based on trie, such situation almost doubled size of every node which significantly increases its size in memory.

3. Custom associative array structure based on hash table

An associative array is a known data structure based on key/value pairs. The associative array has a set of keys and each key has a single associated value. For presented key, the associative array will return the associated value. An associative array is also called a map or a dictionary. Associative arrays are often implemented based on hash tables [2, 6]. There are many ready implementations of associative arrays in many programming languages, especially in Java or .NET libraries.

In our situation we have pairs (word, index) where index means index in table with indexes to base forms table. Of course word is a key and index is a value. We have to note that we cannot immediately points to base forms table, we need additional table with indexes to base forms table, because in Polish language we have many homographs. For example word mają has possible base forms: mać (verb), mieć (verb) or maja (noun - Lithodes maja). This additional table was implemented also in dictionary implementation based on trie described in [5].

To store hashed keys we prepared huge table with size over 4 million cells (i.e. bigger than number of words), we chosen 7199369 because it is prime number. For every word we count hash code in the interval [0, 7199368]. In other words at the last step of counting hash code for a given word, we take hash(word) as hash(word) mod 7199369. Such code hash(word) is taken as an index in the table, but cell value is additional index to array which stores pairs (word, index). Because it may happen that a few words have the same hash code (i.e. we have conflict), from this reason array which stores pairs (word, index) has additional attribute next i.e. it stores triples (word, index, next). Attribute next points to next cell where word has the same hash code.
In the Fig. 2 indexes in the table S are hash codes, for example $h_1 = \text{hash}(w_1)$, $h_2 = \text{hash}(w_2)$, $h_2 = \text{hash}(w_3)$ and $h_2 = \text{hash}(w_4)$. Table B stores triples (word, index, next). Hash code $h_1$ is unique, it points to triple at index $i$ in the table $P$ it is $(w_1, i_1, -1)$. Index $i_1$ points to the table with indexes to base forms table, but in this situation, attribute next in this triple is equal -1 (it denotes that we have not next words with the same hash code). Hash code $h_2$ is not unique, in such situation $j$ points to the first word $w_1$ (i.e. triple with word $w_1$) with hash code $h_2$ in the table $P$ it is $(w_2, i_2, k)$. Index $i_2$ points to the table with indexes to base forms table, but attribute next has value $k$. It means it points to the next word $w_3$ with the hash code $h_2$. Analogously next attribute has value $l$ and points to third word with hash code $h_2$ i.e. $w_4$. Because we have not more words with hash code $h_2$, triple $(w_4, i_4, -1)$ has attribute next equals -1, it means that it was last word with hash code $h_2$.

In other words we can say that to every cell in table $S$ we attach list (chain) of pairs (word, index). We have to note that table $S$ is huge to reduce to minimum number of conflicts. Average performance for associative array is constant, but maximum number of conflicts sets a worst-case performance.

4. Comparing effectiveness of dictionary structures

As in [5] our application was implemented in C#. There are 204050 base forms and total 3953809 words in dictionary [8], but without homographs we have 3801058 different words. We prepared six versions of our application based on: sorted array, trie structure, TST, associative array, .NET Dictionary class and .NET HashSet class. The first two versions were tested in [5], we tested them again because they were slightly improved in comparison with [5]. The following two versions are based on structures described in the previous sections, the last two versions based on .NET Framework 4.7.2. We tested applications based on Dictionary
and HashSet classes because they are very similar to our implementation of associative array.

First we measured amount of RAM memory needed to store whole dictionary structures. Dictionary with sorted array of strings takes about 150MB of RAM memory, dictionary with trie structure takes more than 1.1GB of RAM memory, dictionary structure based on TST takes about 250MB of RAM memory, our associative array takes about 200MB of RAM memory, dictionary based on .NET Dictionary class takes about 250MB of RAM memory, similarly solution based on .NET HashSet class.

Next we prepared to measure average performance (time complexity) of searching in our applications. Theoretically binary search is equivalent to searching in balanced binary search tree and average performance is $O(\log n)$ where $n$ is the number of words ($\log_2 3953809 = 21.9148117485834 \approx 22$). For binary search standard Array.BinarySearch method from .NET library was used. In trie structure searching performance does not depend on number of words, it depends on longest word (i.e. number of characters). In our dictionary [1] this number is 39 (niedzięwięćdziesięciopięciopółścieniowego), of course average performance is $O(1)$. For TST structure average performance is $O(\log n)$ and worse $O(n)$, but many depends on words and their order. In our case deepest branch of TST is 106. Of course it is bigger than 22 for binary search but in TST every time we compare only one pair of characters but for binary search pair of whole words. Associative array has constant searching performance $O(1)$. Many depends on longest chain of words with the same hash code, in our situation this number is 6, and we have 2953044 different hash codes which means that most often we have only one element chain. In associative array we used standard .NET method GetHashCode() defined for strings, obtained value is taken without sign and modulo 7199369. For .NET Dictionary and HashSet searching performance is $O(1)$. We have to add that for HashSet we had to defined EqualityComparer class and defined method GetHashCode() (for pair (word, index) we take word.GetHashCode()).

To practically compare the search results of two dictionary structures, we selected several publications from Agricultural Engineering Journal (Inżynieria Rolnicza) exactly the same that were used in the paper [5]. “Text A” is “Information system for acquiring data on geometry of agricultural products exemplified by a corn kernel”; “Text B” is “Assessment of the operation quality of the corn cobs and seeds processing line”; “Text C” is “Methodological aspects of measuring hardness of maize caryopsis”; “Text D” is “Evaluation of results of irrigation applied to grain maze”; “Text E” is “Extra corn grain shredding and particle breaking up as a method used to improve quality of cut green forage”; and “Text F” is “Comparative assessment of sugar corn grain acquisition for food purposes using cut off and threshing methods”. Additionally, we have prepared three “artificial” texts. “Text X” contains two thousand times word contains; this word is not present
in the Polish language dictionary. “Text Y” contains two thousand times word *niewybielaty* and “Text Z” contains two thousand times word *niewybielatych*.

The results of the test are presented in Table 1 – Table 9, similarly like in [5]. In the header we put the number of words in the particular text. The measure is the number of processor ticks. Every test was taken two times: for one thousand loops, and ten thousand loops. The reason is that .NET Just In Time compiler prepares methods before the first run, if we run method next time, compiled method code is in memory. The result is influenced by a certain overhead during the first loop, it is something like \( \text{overhead\_tics} + n \times \text{tics\_for\_one\_loop} \). For bigger \( n \) we can better estimate average number of tics for one loop taking ratio tics/number of loops. Of course from many reasons, connected with .NET environment, number of tics can differ between two runs, but differences are not significant. We did the tests many times for every case and the results differed insignificantly.

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>33591897</td>
<td>342741</td>
<td>975962</td>
<td>160745</td>
<td>292288</td>
<td>243338</td>
</tr>
<tr>
<td>10000</td>
<td>316131531</td>
<td>3485992</td>
<td>9736250</td>
<td>1563227</td>
<td>2887232</td>
<td>2432310</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>53229999</td>
<td>500002</td>
<td>1355375</td>
<td>250009</td>
<td>469488</td>
<td>394600</td>
</tr>
<tr>
<td>10000</td>
<td>467090971</td>
<td>5009019</td>
<td>13488726</td>
<td>2452428</td>
<td>4665948</td>
<td>3940382</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>45597531</td>
<td>429401</td>
<td>1190847</td>
<td>213183</td>
<td>397152</td>
<td>331165</td>
</tr>
<tr>
<td>10000</td>
<td>406983513</td>
<td>4248994</td>
<td>11779211</td>
<td>2072389</td>
<td>3870395</td>
<td>3284391</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>28353364</td>
<td>250092</td>
<td>725084</td>
<td>130667</td>
<td>232243</td>
<td>211289</td>
</tr>
<tr>
<td>10000</td>
<td>256743505</td>
<td>2509729</td>
<td>7216749</td>
<td>1279460</td>
<td>2294568</td>
<td>2044434</td>
</tr>
</tbody>
</table>
We summarized all results in the Table 10, taking approximate number of ticks, leaving only two important digits, to show general tendency. We can observe that searching results with associative array are the best. Anyway they are comparable with versions based on .NET structures: Dictionary and HashSet. Version based on trie is about 2 times slower, but still comparable with associative array. Searching with TST is significantly slower, about 15 times. This confirmed our previous theoretical considerations. Binary search is very slow regarding to all other methods, for example associative array is more than 200 times faster than binary search. We can observe that for example for texts D and E the TST searching was relatively slow, it means the tree structure affects search time. For associative array and .NET

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>32357241</td>
<td>327902</td>
<td>894725</td>
<td>158679</td>
<td>292624</td>
<td>240832</td>
</tr>
<tr>
<td>10000</td>
<td>30765139</td>
<td>3255707</td>
<td>8889389</td>
<td>1543999</td>
<td>2928692</td>
<td>2422897</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>39847235</td>
<td>356195</td>
<td>1097629</td>
<td>178936</td>
<td>323735</td>
<td>281710</td>
</tr>
<tr>
<td>10000</td>
<td>353018299</td>
<td>3530443</td>
<td>10278096</td>
<td>1739347</td>
<td>3182355</td>
<td>2787582</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>48424413</td>
<td>331754</td>
<td>1600371</td>
<td>101902</td>
<td>114812</td>
<td>205132</td>
</tr>
<tr>
<td>10000</td>
<td>459046035</td>
<td>3411198</td>
<td>16180765</td>
<td>1038634</td>
<td>1145758</td>
<td>2000187</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>48844099</td>
<td>664130</td>
<td>1606728</td>
<td>224155</td>
<td>434106</td>
<td>331085</td>
</tr>
<tr>
<td>10000</td>
<td>48845906</td>
<td>6727924</td>
<td>16121447</td>
<td>2087378</td>
<td>4192262</td>
<td>3161444</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of loops</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>55643747</td>
<td>750973</td>
<td>1708222</td>
<td>262994</td>
<td>437585</td>
<td>332897</td>
</tr>
<tr>
<td>10000</td>
<td>503264504</td>
<td>7589691</td>
<td>17459160</td>
<td>2601833</td>
<td>4290230</td>
<td>3261498</td>
</tr>
</tbody>
</table>
structures results are rather proportional to number of words in tested text. In this situation results for text X, word contains hash code not presented in dictionary and negative result is obtained immediately.

<table>
<thead>
<tr>
<th>Text</th>
<th>Binary search</th>
<th>Trie search</th>
<th>TST search</th>
<th>Associative array</th>
<th>NET Dictionary</th>
<th>NET HashSet</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>31000</td>
<td>340</td>
<td>970</td>
<td>160</td>
<td>290</td>
<td>240</td>
</tr>
<tr>
<td>B</td>
<td>46000</td>
<td>500</td>
<td>1300</td>
<td>250</td>
<td>460</td>
<td>390</td>
</tr>
<tr>
<td>C</td>
<td>40000</td>
<td>420</td>
<td>1100</td>
<td>200</td>
<td>380</td>
<td>320</td>
</tr>
<tr>
<td>D</td>
<td>25000</td>
<td>250</td>
<td>7200</td>
<td>130</td>
<td>230</td>
<td>200</td>
</tr>
<tr>
<td>E</td>
<td>30000</td>
<td>320</td>
<td>8800</td>
<td>150</td>
<td>290</td>
<td>240</td>
</tr>
<tr>
<td>F</td>
<td>35000</td>
<td>350</td>
<td>1000</td>
<td>170</td>
<td>320</td>
<td>280</td>
</tr>
<tr>
<td>X</td>
<td>45000</td>
<td>340</td>
<td>1600</td>
<td>100</td>
<td>110</td>
<td>200</td>
</tr>
<tr>
<td>Y</td>
<td>48000</td>
<td>670</td>
<td>1600</td>
<td>200</td>
<td>420</td>
<td>310</td>
</tr>
<tr>
<td>Z</td>
<td>50000</td>
<td>750</td>
<td>1700</td>
<td>260</td>
<td>430</td>
<td>330</td>
</tr>
</tbody>
</table>

5. Conclusions and future work

We examined six dictionary structures for text analysis in particular for indexing text. Tests have shown that the structure based on the associative array makes searching faster than other structures. Five structures but trie utilize similar amount of RAM memory. For our purpose dedicated association table is the best choice. If somebody does not want to implement dedicated structure classes from .NET library are relatively good. However, the trie structure and TST are still useful in tasks such word completion or error correction. Trie is very good choice if we have limited dictionary, for dictionaries like [8] the better choice for mentioned tasks is TST which compromises the advantage of fast completion and reasonable RAM amount. The main conclusion is that our application should be developed based on associative array but parallel version with trie structure can be useful in special task like morphology study.

REFERENCES


STOCK MANAGEMENT AS A CRITICAL SUCCESS FACTOR FOR POLISH CROP PRODUCERS

PIOTR KORNETA

Faculty of Management, Warsaw University of Technology

The key objective of critical success factors is to filter out excessive information reaching organizations so that management can focus on several most critical areas. Both scholars and practitioners employ most frequently expert interviews as to identify critical success factors. The aim of this study is to show how quantitative methods can contribute to a more efficient critical success factors identification. This study uses a sample of observations relating to 300 Polish crop producers in a 5-years period between 2013 to 2017. The findings of this study show clearly that the lower the inventory levels the higher the profitability and the growth of sales revenues of Polish crop producers.

Keywords: critical success factors, performance measurement, performance management systems, inventory

1. Introduction

Last three decades have witnessed considerable development of performance management systems, among which balanced scorecard become the most well-recognized one. Currently performance management systems are defined as dynamic and balanced systems, which facilitate support of decision-making processes by gathering and evaluating relevant information. Although there is a huge number of published performance management systems, the vast majority of these frameworks rely highly on measurement and critical success factors [4, 13, 21]. The concept of critical success factors for companies was discussed by Ronald Daniels in
1961, who claimed that information systems must focus on a limited number of factors. These factors, if addressed properly, should both, ensure an organization’s success and prevent the organization from receiving too much information [6, 17].

According to the literature, critical success factors can stem from: industry, environment, competitors, partnerships with clients or suppliers, information systems, strategic planning, products, process management, working capital management, knowledge management, environmental or country specific factors, financing, and other sources. The large number of critical success factors described in literature results from the fact that it is widely acknowledged that critical success factors should be tailor-made for each organization. Hence, one size fits all rule is not applicable for critical success factors [1, 19, 26].

With respect to above, it can be concluded, that critical success factors, constitute a fundamental role of performance management systems, and if are managed well, should directly improve the company’s performance. The company’s performance, in turn, can be measured with, inter alia, profitability or growth measures.

The aim of this paper is to study if the level of inventories at crop producers can be considered as critical success factor. Inventories in crop producers’ industry are subject to several significant risks, including volatility of prices and various sources of impairment. As to achieve the aforementioned objective this study employs relevant statistical test.

The rest of this paper is organized as follows. In the next, second section a brief literature review is being provided, based on which relevant hypotheses are developed. The following sections provides methodology for this study. In the fourth section the results are being presented and discussed. In the last section of the paper conclusions are being provided, practical implications, limitations of the study and the directions for further research.

2. Theoretical background, hypotheses development

The literature on inventories is vast and continues to develop. Significant portion of papers relates to various aspects of stock management. The number of papers relating to stock levels optimization is also considerable. In general, the papers on stock levels optimization can be divided into those aiming to optimize stock levels within one organization and among a chain of cooperating companies, so called supply chains [3, 20, 24]. The number of papers measuring relationships between stock levels and profitability or growth of companies in various industries are considerable less numerous, while for some industries or countries insufficient, which justifies this study. According to Blinder and Maccini higher stock levels should improve the profitability of companies. This should be achieved through reduction of production interruptions in manufacturing cycle, which ought to protect abnormal costs of products, provide protection against price fluctuations and
prevent loss of business resulting from non-availability of raw materials. That sequence of cause and effects should finally lead to profitability improvements [5].

The majority of scholars however, postulate lower inventory levels are associated with higher profitability. The findings of these studies use either statistical test or regression analysis and are undertaken in various countries and various industries [9, 11, 15].

The selection of inventory levels for crop producers in this study, is primarily interesting because of the two more reasons. The first is the volatility of crop prices [16], which impose high risk on crop inventory holders, unless the inventory prices are properly hedged. Given, the awareness of Polish farms in this matter is not high, i.e. assuming Polish crop producers, especially smaller farms do not hedge crop prices than such companies incur considerable risk levels which might affect their profitability or even an overall performance. The second reason is that crop as inventory is subject to several physical risks such as, for example humidity, mold, or vermin, hence if crop is not stored properly than inventories can get impaired, loosing considerable value.

Based on argumentation provided above and in accordance with the majority of scholars the following hypothesis is developed:

H1: There is a negative and statistically significant relationship between profitability and the inventory levels at Polish crop producers.

Additionally, since the level of inventories should intuitively contribute to the growth of sales, the second hypothesis is formulated as follows:

H2: The higher the inventory the faster the growth of sales of Polish crop producers.

Verifications of relationships between inventory levels and both profitability and the growth of sales should allow to understand studied subject more thoroughly.

3. Methodology

The sample for this study comprise 300 Polish crop producers in a 5-years period between 2013 and 2017. The figures used in the study have been obtained from EMIS database (Emerging Markets Information Service) in October 2018. In particular, financial statements of 300 Polish crop producing companies in studied periods, between 2013 to 2017 have been obtained. Since EMIS database have not comprised financial statements of all studied companies for each of the studied periods only 1,017 observations have been obtained. Furthermore, due to the use of lagged variables, the number of observations used in this study reduced to 661 accordingly.
Table 1 provides the variables, with their descriptions, used in the study. This study uses 4 profitability variables, namely return on sales – in two versions, return on assets and return on equity. Stated 4 variables are widely acknowledged measures of profitability in the literature and so commonly used in various studies. Return on sales is used in two versions, the first version employs profit after tax, whereas the second one EBIT (earnings before interests and tax) and is considered as good proxy of profit margins. Therefore, this variable is of particular interest to both decision makers and scholars [2, 8, 14, 22, 25]. Growth of companies is measured in the literature in various ways [12, 23]. This study measures the growth of sales as a variance of sales revenues between two consecutive years. Finally, the study uses two variables relating to current and prior year inventories, both calculated as a percentage of sales.

Table 1. Variables used in the study

<table>
<thead>
<tr>
<th>Variable</th>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return on sales</td>
<td>ROS</td>
<td>Ratio of profit after tax and sales</td>
</tr>
<tr>
<td>Return on sales 2</td>
<td>ROS2</td>
<td>Ratio of EBIT (Earnings before interests and tax) and sales</td>
</tr>
<tr>
<td>Return on assets</td>
<td>ROA</td>
<td>Ratio of profit after tax and total of assets</td>
</tr>
<tr>
<td>Return on equity</td>
<td>ROE</td>
<td>Ratio of profit after tax and equity</td>
</tr>
<tr>
<td>Growth of sales</td>
<td>GRS</td>
<td>((Sales t-year) – (sales t-1 year)) divided by (sales t-1 year)</td>
</tr>
<tr>
<td>Inventory</td>
<td>INVS</td>
<td>Inventory as a percentage of sales</td>
</tr>
<tr>
<td>Inventory prior year</td>
<td>INVSPY</td>
<td>Prior years’ inventory as a percentage of sales</td>
</tr>
</tbody>
</table>

Pursuant to selection of variables for the study, a descriptive statistic of all variables have been reviewed. As significant portion of statistical test of significance requires normality assumption of both tested variables to be met or nearly met, for example Pearson correlation test [7], normality tests have been undertaken. In particular, Doornik-Hansen and Shapiro-Wilk tests of normality [10, 18] have been calculated. The null hypothesis for both of these tests is the same and states that the variables are normally distributed. Finally, relevant tests of significance have been selected and calculated as to obtain required results, which were next analyzed and discussed.
4. Results and discussion

Table 2 provides descriptive statistics of variables used in this study. On average, crop producers included in the study exhibit a high mean growth of sales revenues of 0.177 over the studied period from 2013 to 2017. Mean value of return on equity of 0.1038 is not especially high as compared to other, more profitable industries, but exceeds considerably the interests on bank deposits. As disclosed in Table 2, inventories of studied companies were high, accounting to around 0.37 of yearly sales levels.

Table 2. Descriptive statistics of variables used in the study

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS</td>
<td>0.0898</td>
<td>0.9141</td>
<td>0.0819</td>
<td>-22</td>
<td>6.2888</td>
<td>-20.9244</td>
<td>517.765</td>
</tr>
<tr>
<td>ROS2</td>
<td>0.1301</td>
<td>0.1787</td>
<td>0.0994</td>
<td>-0.9333</td>
<td>0.9432</td>
<td>0.9721</td>
<td>6.7531</td>
</tr>
<tr>
<td>ROA</td>
<td>0.057</td>
<td>0.2255</td>
<td>0.0446</td>
<td>-3.0251</td>
<td>2.9583</td>
<td>-1.7394</td>
<td>114.049</td>
</tr>
<tr>
<td>ROE</td>
<td>0.1038</td>
<td>0.7184</td>
<td>0.0753</td>
<td>-7.6075</td>
<td>9.1739</td>
<td>0.7045</td>
<td>83.2961</td>
</tr>
<tr>
<td>GRS</td>
<td>0.177</td>
<td>1.4352</td>
<td>-0.0118</td>
<td>-0.8009</td>
<td>27.6154</td>
<td>12.9204</td>
<td>215</td>
</tr>
<tr>
<td>INVS</td>
<td>0.3715</td>
<td>0.2675</td>
<td>0.3406</td>
<td>3.0083</td>
<td>2.3022</td>
<td>15.4869</td>
<td></td>
</tr>
<tr>
<td>INVSPY</td>
<td>0.3636</td>
<td>0.267</td>
<td>0.3228</td>
<td>2.0928</td>
<td>1.7474</td>
<td>6.7132</td>
<td></td>
</tr>
</tbody>
</table>

Source: own elaboration based on data provided in EMIS database

Results of Doornik-Hansen and Shapiro-Wilk tests of normality are being provided in Table 3. Since all p-values are significantly below the threshold of 5%, it is evident that none of the studied variables have a normal distribution. As a consequence, Pearson correlation, which requires normality assumption to be meet should not be used for any conclusions. Hence, Pearson coefficients are disclosed in further works only for indicative purposes.

Table 3. Results of normality tests with p-values (2 sides)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Doornik-Hansen</th>
<th>P (α=5%)</th>
<th>Shapiro-Wilk</th>
<th>P (α=5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS</td>
<td>125501</td>
<td>0</td>
<td>0.11785</td>
<td>9.58E-48</td>
</tr>
<tr>
<td>ROS2</td>
<td>228.874</td>
<td>2.00E-50</td>
<td>0.864917</td>
<td>2.04E-23</td>
</tr>
<tr>
<td>ROA</td>
<td>8626.9</td>
<td>0</td>
<td>0.367553</td>
<td>8.96E-04</td>
</tr>
<tr>
<td>ROE</td>
<td>6614.87</td>
<td>0</td>
<td>0.321961</td>
<td>8.63E-44</td>
</tr>
<tr>
<td>GRS</td>
<td>35654.6</td>
<td>0</td>
<td>0.203138</td>
<td>3.39E-46</td>
</tr>
<tr>
<td>INVS</td>
<td>236.489</td>
<td>4.44E-52</td>
<td>0.870876</td>
<td>6.09E-23</td>
</tr>
<tr>
<td>INVSPY</td>
<td>210.456</td>
<td>2.00E-46</td>
<td>0.889121</td>
<td>2.23E-21</td>
</tr>
</tbody>
</table>

Source: own elaboration based on data provided in EMIS database

As indicated above, since a normality distribution of studied variables cannot be assumed, this study uses Spearman’s rank correlation coefficient and Kendall
rank correlation coefficient. Both of stated tests are non-parametric measures of rank correlation, which are resistant to the effects of outliers and nonnormality [7]. In this case, stated test are superior to Pearson correlation.

The results of selected statistical tests of dependence between profitability and growth of sales and inventory levels (INVS) in the end of studied periods are provided in Table 4.

<table>
<thead>
<tr>
<th>Variable</th>
<th>PEARSON</th>
<th>p</th>
<th>Spearman</th>
<th>p</th>
<th>KENDALL tau-B</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS</td>
<td>0.0641</td>
<td>0.2045</td>
<td>-0.0091</td>
<td>0.4079</td>
<td>-0.019</td>
<td>0.471</td>
</tr>
<tr>
<td>ROS2</td>
<td>-0.0881</td>
<td>0.0117</td>
<td>* 0.0169</td>
<td>0.332</td>
<td>0.019</td>
<td>0.2329</td>
</tr>
<tr>
<td>ROA</td>
<td>-0.0641</td>
<td>0.0499</td>
<td>* -0.1541</td>
<td>0</td>
<td>* -0.0973</td>
<td>0.0001</td>
</tr>
<tr>
<td>ROE</td>
<td>-0.0952</td>
<td>0.0072</td>
<td>* -0.2368</td>
<td>0</td>
<td>* -0.146</td>
<td>0</td>
</tr>
<tr>
<td>GRS</td>
<td>-0.0165</td>
<td>0.3362</td>
<td>-0.1634</td>
<td>0</td>
<td>* -0.1037</td>
<td>0</td>
</tr>
</tbody>
</table>

*Source: own elaboration based on data provided in EMIS database*

The results of selected statistical tests of dependence between profitability and growth of sales variables and inventory levels (INVSPY) on the beginnings of studied periods are provided in Table 5.

<table>
<thead>
<tr>
<th>Variable</th>
<th>PEARSON</th>
<th>p</th>
<th>Spearman</th>
<th>p</th>
<th>KENDALL tau-B</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS</td>
<td>-0.2331</td>
<td>0</td>
<td>* -0.046</td>
<td>0.1189</td>
<td>-0.0239</td>
<td>0.1787</td>
</tr>
<tr>
<td>ROS2</td>
<td>-0.2055</td>
<td>0</td>
<td>* -0.0219</td>
<td>0.2866</td>
<td>-0.0058</td>
<td>0.4124</td>
</tr>
<tr>
<td>ROA</td>
<td>-0.1786</td>
<td>0</td>
<td>* -0.174</td>
<td>0</td>
<td>* -0.1119</td>
<td>0</td>
</tr>
<tr>
<td>ROE</td>
<td>-0.0675</td>
<td>0.0414</td>
<td>* -0.2513</td>
<td>0</td>
<td>* -0.1605</td>
<td>0</td>
</tr>
<tr>
<td>GRS</td>
<td>-0.1643</td>
<td>0</td>
<td>* -0.2657</td>
<td>0</td>
<td>* -0.1736</td>
<td>0</td>
</tr>
</tbody>
</table>

*Source: own elaboration based on data provided in EMIS database*

Based on results of Spearman rho coefficients and KENDALL tau-B presented in Table 4 and Table 5 relating to return on assets (ROA) and return on equity (ROE) the first (H1) hypothesis stating that there is a negative and statistically significant relationship between profitability and the inventory levels at Polish crop producers must be accepted. The results of this study are therefore aligned to the majority of other researches.

Since the Spearman rho and KENDALL tau-B coefficients are negative and statistically significant (with p around zero) the second hypothesis stating that the higher the inventory the faster the growth of sales of Polish crop producers must be rejected.
Additionally, it should be noted that the coefficients for both sales profitability variables are statistically insignificant (Spearman rho and KENDALL tau-B), which shows that higher inventory levels do not increase sales profitability, i.e. crop producers to not enjoy higher margins despite possession of higher inventory levels.

5. Conclusion

The objective of this paper is to verify if the level of inventories at crop producers can be considered as a critical success factor. This objective is being achieved on inventory levels example at Polish crop producers in a 5-years period between 2013 to 2017.

Critical success factors for the purpose of this study are defined as the factors which, if improved increase profitability or growth of sales of studied companies. Based on results obtained in empirical part of the paper, which employs Spearman rho coefficient and Kendall tau-B a negative and statistically significant relationship between inventory levels and both the profitability and the growth of sales is identified. The dependence between inventory levels and growth of sales and profitability of Polish crop producers is being identified with the use of statistical methods, which confirms that the level of inventories at crop producers is the critical success factors.

The practical implication for decision makers of this study is that crop producers with lower inventory levels, in general, enjoy higher profitability and growth of sales revenues. This, in turn, promotes decisions aimed at stock reductions. Findings of this study seem to be aligned to business practice of crop producers, as inventory at stated industry is subject not only to typical inventory risks like obsolesce stock or shortages, but also high price volatility and additional industry specific physical risk factors such as mold or vermin. Hence, crop producers in the end of the business seasons are recommended to consider sales of stocks with reduced margins, which should improve their profitability.

This study has however, several limitations, which are primarily related to the sample of studied companies, which is limited to only one country and one industry. Additionally, longer time period could also be analyzed. Aforementioned limitations of the study are good indications for further research. Accordingly, further studies could focus on different countries or different industries or could study longer time frames.
REFERENCES


MICRO- VS. MACRO SCOPE SIMULATION IN BPM MODELLING

ROMAN PIETRONI

Wroclaw University of Science and Technology, Department of Operations Research, Finance and Computer Applications

The business structure and process development in many modern companies is assisted by architectures, frameworks, and a new generation of computer systems with a broad representation of methods and tools of design and modelling, i.e. object-oriented, event-oriented, activity-oriented, process-oriented, agent-oriented, communication-oriented, behaviour-oriented, etc. are recommended. The aim of this paper is to compare micro- vs. macro scope modelling and to investigate the relevance of micro- and/or macro simulation in the context of business process management (BPM) modelling for organizational design/redesign. The BPM modelling features and some specific issues of simulation modelling are discussed. Combined with BPM modelling methodologies, these issues are used to drive a concept for development of an integrated methodology for BPM models’ building and simulation approach in organizational BPM design/redesign projects.

Keywords: BPM, Simulation modelling

1. Introduction

The modern business activities are being featured by the ongoing rise of a post-industrial network society with its technological ICT improvements and economy shifts. Any information, process and capital can flow almost immediately and competition is not constrained by any national or regional barrier. In order to function in such a changing (also turbulent) and complex environment all organizations need to adapt themselves to new conditions. Encouraged by this need, the business process management (BPM), management of change, continuous improvement and
building excellence of business processes have become issues of importance for business organisations. Business process re-engineering (BPR), continuous process improvement (CPI), total quality management (TQM), and other approaches have been developed to help organizations in managing and implementing change in order to improve productivity and gain competitive advantage. The increasing business and academic interest has resulted in many approaches, methodologies, and techniques to support organizational change, particularly by many business process management (BPM) modelling methodologies and perspectives - see BPM modelling reviews and surveys, e.g. [1], [6]. However, despite this support, the level of failures in BPM modelling projects still remains high. Amongst the various explanations that have been offered is a lack of tools for evaluating proposed changes before implementation. The lack of a theoretical basis for BPM modelling constructs is quite clear when some elements of a theory are compared with the model constructs and frameworks. Some of the lists of model constructs describe all three semiotic dimensions of the subject-model-modeller relationship, and some seem unaware of the independence of the subject-model (i.e. semantic) relationship and the model-modeller (i.e. pragmatic) relationship. Also finally, operability requires that the BPM modelling must be done by formally created models using formal model constructs. Non-formality implies errors during model execution in BPM software that take the form of uncontrolled divergences between the model and the subject being modelled – and it is not due to purposeful abstraction and ambiguity. Formal models have to be written in a formal modelling notation using an ontology-based modelling grammar.

The emergence of the simulation modelling (SM), as an important field of BPM modelling support, creates a new need and interest for BPM research community in the ways in which the SM method or technique can assist the process of modelling and analysing business processes, functions and structures. However, despite the existing application examples, SM in BPM modelling has not to date received the methodological support needed to establish it as a separate research area. Also a micro-scope method of modelling, usually as stochastic discrete modelling, is a dominant form of simulation to be applied in BPM models. And finally, BPM community is not interested in macro-scope method of BPM modelling, even it can bring fundamental answers to some questions about dynamic properties of system and process structure behaviour. The paper addresses this methodological gap by investigating the suitability of micro- and macro SM in the context of process-oriented (PO) organizational analysis and design, and also answers the question how the BPM process paradigm is related to traditional SM paradigm. In fact, the organizational change is primarily a design problem and encourage a use of SM models with team/group communication procedures as a mechanism that can efficiently address process design/redesign requirements.
2. Process paradigm in BPM modelling projects

BPM modelling is obviously a part of business modelling and also business process models differ in the types of decisions they are able to support [4]. The central notion in business model should be the concept of value, not always dominant in BPM models. Basically, a business model does not define how value-adding activities are carried out. Although a business process is the fundamental unit of BPM, there is no clear definition of the term. Some differences can be found in BPM according to three variables: the fundamental (the smallest identifiable and independent) unit of a process analysis, the primary objectives, and the mechanisms that the process uses to transform its inputs into meaningful outputs.

There is a common agreement that in BPM:

- business processes are decomposed into a number of more elementary steps (usually referred to as tasks or activities),
- focus should be on business process with its internal (i.e. within the organization) or external customers,
- a process customer can be a person or another process, and processes should satisfy the expectations of their customers (added value concept),
- there is a low attention to the mechanisms which business processes use to attain their objectives and also diverse views of importance are favoured.

The definitions used in BPM imply a systemic view of organizations, as sets of interacting entities with their external environment in order to achieve specific objectives. This view implies in a natural way that the techniques of system and process analysis and experimenting (by e.g. by simulation/gaming) have the potential to address the problems of BPM design and business change. The PO paradigm implies a way of looking at organizations based on the processes they perform rather than the functional units they are divided into. But despite the changes in economic and social environments, management values, needs, expectations (also wrong archetypes) still determine the BPM in organisations.

The PO analysis advocates the re-unification of separate business tasks so that they constitute a set of activities with clearly added value to their (internal or external) customers. It can be argued that PO analysis for the purpose of business change is primarily a design problem. According to the information processing and decision-making paradigms of organizational design, processes can be viewed as sets of decision models, each of which is identified by a type of decision and contains information processing tasks. These tasks are the smallest identifiable units of analysis and their arrangement is the critical design variable determining the efficiency of the resulting structures. Techniques that allow experimenting with alternative configurations and process layouts, and comparing between diverse proposals, would be useful for organizational design. SM is well suited for this design purpose but a specific methodology for employing it for BPM must be developed.
3. SM as a part of BPM modelling

Computer-based simulation models of BPM issues can help overcome the inherent complexities of analysing businesses and contribute to a higher level of business understanding. The basic idea behind SM is simple: we wish to acquire knowledge and reach decisions for a real-world BPM system. Because the system is not easy to study directly, we proceed indirectly by creating and studying another entity (a simulation model), which is similar to the real-world system that we are confident that some of what we learn about the model will also be true of the system. In the business change and its BPM modelling projects, SM is attractive since it allows the studying of BPM complexity before its implementation and it can help to define early deficiencies in the design process when correction is easily and less expensively made. SM also prefers PO approach defined as a time-ordered set of interrelated activities describing the flow of an entity through a system. Also SM terminology is similar to that used in organizational analysis, allowing for good communication and exchange of ideas between SM experts and decision-makers.

A major advantage of SM over other operational research techniques (OR/MS) is that it allows for replicable experimentation with any element of a business system and process, and answering what if questions. The aspects of tangible (e.g. technical) and intangible (e.g. managerial) elements of a business can be incorporated in a model. Furthermore, SM allows the decision-makers to obtain a system-wide view of the effects of local changes in a system and allows for the identification of implicit dependencies between parts of the system (e.g. by application of SSM soft system and systems thinking approaches). Process deficiencies and bottlenecks than can be identified without destroying the business by TO-BE process experiments, and prototyping or implementation. When combined with visible diagramming (see e.g. [2]), animation and interaction capabilities, e.g. by simulation gaming and GMB - group model building sessions (see e.g. [7]), SM facilitates better understanding of a system's behaviour and of the impact of proposed changes on costs/benefits, and better communication. Finally, SM encourages BPM an organisational learning, and a cultural shift in the way any modelling is perceived by performance measurement.

Five modes of BPM simulation modelling practice (modes) can now be identified: mode 1 - SM as software engineering activity, mode 2 - SM as gaming activity in research and education (training), mode 3 - SM as a process of organisational change by intervention, mode 4 - SM as a facilitation tool in organisational change, and mode 5 - SM as a sub-process of BPM architecture/framework structure (SM in BPM package). The first two modes are to be identified by majority of observations and descriptions of business and BPM simulation modelling projects in the literature. The next three modes can be derived from BPM practitioners reports and proposals/recommendations found in the literature. Actually, at present there is
little evidence of significant BPM practice of the third and fourth mode - it represents a potential future for SM applications in BPM. As already stated, these modes are derived from, and relate to, SM as it is practised in BPM - that is not to say that they are exclusive to this domains, or that other practices do not exist.

SM as software engineering (mode 1) is typified by lengthy projects, possibly taking years to complete, that are performed by teams of modellers. The BPM users of the models are often far removed from the development process, and may get involved only when a specific problem is to be tackled with the completed model. The prime motivation of such projects is the accurate representation of the BPM real world. In some cases, this is to such an extreme form, that a model is developed without having a specific problem to be tackled, leading to models looking for a problem. SM as gaming (mode 2) is based on application of simulation games in BPM, particularly as modelling approach (research tool) to design a business process or experimental (operational) tool to educate and train business process actors in organisation. In fact, there are many commercial simulation games for BPM in many business domains (e.g. in production/manufacturing, logistics, marketing, crisis management). In SM as a process of organisational change by intervention (mode 3), the role of the BPM modeller is as an agent of change, whose task it is to help the user (who may better be described as a customer) perform his/her job better. The modeller works with the customer to help him/her better understand the nature of the organisations problems and to identify actions that may lead to an improvement. The prime motivation for such projects is problem understanding and problem solving. These projects tend to be short, typically a few weeks, are performed by a lone modeller, and require high levels of customer involvement. SM as facilitation (mode 4) in BPM can be seen as a special, but extreme, case of simulation as a process of organisational change where a BPM model is developed and used (in an interactive manner) in a group meeting as a means for understanding the real BPM world and for promoting discussion on potential process improvements. The prime motivation is understanding and provoking a debate about the problem situation. Model accuracy is potentially of little significance as long as it is useful for promoting the group discussion. There is much similarity between this mode of practice and GMB method. Finally, SM as a subprocess of BPM architecture/framework (mode 5) means that particular module of BPM software package, based on process formal representation (usually with micro-scope modelling approach), is running a simulator developed for enactment BPM life-cycle stage, and there is not required a methodological knowledge by the package user.

For the top management purposes, a system dynamics modelling (see e.g. [3, 8, 9]) based on causality and feedback paradigm is particularly recommended. The macro-scope System Dynamics (SD) modelling method in BPM is adopting systems thinking concept and approach. A model that is helpful for understanding
process “global” (holistic and systemic) issues in BPM SD modeling, is the *iceberg model* (Fig. 1), often used in systems thinking. In BPM life-cycle stages, global business process issues can be looked at in some research and analytical layers, allowing successful system and process restructures and improvements.

![Figure 1. The “Iceberg model” in macro-scope SM with systems thinking and SD approach](image)

The iceberg model is the systems thinking tool designed to help an individual or group discovering the patterns of behavior, supporting structures, and mental models, that underlie a particular process event. If we apply the iceberg model to business process global issues, we could say that at the tip, above the water, are *events*, or thing that we see or hear about happening in the whole business process. If we look just below the water line, we often start to see *patterns*, or the recurrence of events. Finally, at the very base of the iceberg are the assumptions and worldviews that have created or sustained the structures that are in place. The important thing to understand is that in problems’ solving, the greatest leverage is in changing the structure. Like the different levels, deep beneath the patterns are the *underlying structures* or root causes that create or drive those patterns.

Simulation models have been used successfully in individual business analysis and change studies in BPM projects. The typical solution is a generation of key performance indicators (KPI) on the basis of the business process discrete-event (DES) simulation models and its included information (Fig. 2). Usually as KPI we use throughput time, capacity utilization, wait time, number of department and system changes. But it is not possible to carry out a model evaluation without human intervention because of strong influence of interpretation made in particular environment conditions of the BPM system. To calculate KPI some static calculations may be also used. SM modules cycle through the process progress chronologically under real time conditions. But still there are some important potentials of an application of macro-scope SM, particularly by System Dynamics (SD) continuous modelling, allowing to focus on dynamic properties (Fig. 3) of systems’ structures driving process behaviours.
Figure 2. An example of micro-scope BPM – DES production model structure and its typical simulation results (*Arena v.15.00*) – entity detail summary and user specified counters.

<table>
<thead>
<tr>
<th>Entity 1</th>
<th>N/A Time</th>
<th>Other Time</th>
<th>Total Time</th>
<th>Transfer Time</th>
<th>Wait Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.00</td>
<td>28.09</td>
<td>0.00</td>
<td>28.09</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>0.00</td>
<td>28.09</td>
<td>0.00</td>
<td>28.09</td>
<td>0.00</td>
</tr>
<tr>
<td>Entity 1</td>
<td>Number In</td>
<td>Number Out</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>93</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>93</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. An example of macro-scope BPM – SD production model structure and its typical simulation results (*Vensim PLE v. 6.4E*) – dynamics of state variables.

There are many differences in micro-scope BPM SM (DES simulation) and macro-scope BPM SM (SD simulation). Basic differences come from modelling interests and orientations, which leads to different set of concepts, constructs, formal representation, dynamic properties, and goals (Tab. 1). However, SM seems to be also not practical due to cost-effectiveness, and static procedures are preferable giving rough, and nearly similar analytical results. As a conclusion, there is still a need for SM tools that are easily accessible and understood by BPM community, not always proficient in SM. Despite of many successful cases, a comprehensive methodology that integrates SM in business process change projects to evaluate design/redesign scenarios has not yet been successfully developed and some new initiatives and concepts must be developed.
<table>
<thead>
<tr>
<th>BPM modelling aspects</th>
<th>Micro-scope BPM modelling</th>
<th>Macro-scope BPM modelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modelling construct</td>
<td>Items in flows. Unique, and identifiable dynamic items flowing in a activity network of activities. Basic decision variables: static structure of activities and allocation of resources.</td>
<td>Only flows. Unique, non-identifiable items in homogenous dynamic flows of material, energy, and information. Basic decision variables: decision making rules and policies for flows.</td>
</tr>
<tr>
<td>System event as a system and process state change</td>
<td>Mathematical discrete function.</td>
<td>Mathematical continuous function.</td>
</tr>
<tr>
<td>Time advance mechanism in simulation – reviewing system and process states</td>
<td>Interval between system observation points is dependent when events occur (event-oriented time mechanism) or could be also constant – quasi-continuous ($dt=const$). Model recalculations could be: event-oriented, activity-oriented or process-oriented.</td>
<td>Interval between system observation points is usually constant – quasi-continuous ($dt=const$). Model recalculations are sequential and time dependent.</td>
</tr>
<tr>
<td>System and process input data of simulation</td>
<td>Attributes’ representation of dynamic objects (decision rules, join and split rules, routing, generation and flow delay times); Description of static objects (capacities, availabilities).</td>
<td>BPM inputs: decision making rules and policies for flows' control (management), system and process description (flows, stocks, cause-effect relations and feedbacks, decision making and flows delays, decision rules amplifications, capacity constraints).</td>
</tr>
<tr>
<td>System and process output data of simulation</td>
<td>General statistics about system and process (average cycle time, average efficiency, average utilisation of resources, average queue length) and additionally each static or dynamic item can be individually tracked (waiting time, queue length, counter, resource availability, resource efficiency, activity time, utilisation of resource).</td>
<td>System and process dynamic properties (patterns of behaviour, dynamic stability, linearity, steady-state behaviour, equilibrium state conditions, feedback loop dominance) and only general statistics about system and process operational performance (inventory levels, intensiveness of flows, efficiency, transit time, resource utilisation).</td>
</tr>
<tr>
<td>Ordering and regulation in flows</td>
<td>Dynamic (tokens) items can flow in FIFO, LIFO, SPT, by priority, by random, or any customised order. Also static (resources) entities can operate with any order regulation.</td>
<td>Flows are always in FIFO order.</td>
</tr>
<tr>
<td>Routing in network</td>
<td>Dynamic items are automatically routed to the first available branch (items can be only in one place at a time). Cloning is possible but it is made by copying items.</td>
<td>Flows must be explicitly routed by being turned off one branch and turned on at the other. It means that flows can be split.</td>
</tr>
<tr>
<td>Common application goals and domains</td>
<td>Analysis and design in business practice: BPM, BPR, production and manufacturing, logistics and SCM, customer service, systems engineering, software engineering, computer engineering and telecommunication.</td>
<td>Research, analysis and design in science and business practice: systems thinking, systems engineering, economics and finance, ecology, production and manufacturing, logistics and SCM, BPM, BPR, HRM, marketing, sociology, psychology, politics.</td>
</tr>
</tbody>
</table>
4. A new concept of SM in BPM

SM model development within BPM projects can be a complex challenge and set of activities. Some important issues in SM require the attention of business modellers. SM models do not constitute an end in themselves. In the context of organizational design and change management, they should be viewed as means to support informed decision making by management. Therefore, it is of paramount importance that the model boundaries and level of detail are well defined and remain constantly aligned with the objectives of the wider business analysis project. The basic problems for SM as a part of BPM relate to classical SM paradigms and a need for a specific and multi-view BPM. There are problems with finding appropriate views, structures, data selections and modelling procedures for BPM purposes. For example, there are some typical modelling conflicts: generic view vs. specific view, clear definition and scoping vs. polymorphic modelling, aggregated data collection vs. modular model design and analysis, model decomposition vs. model integration (aggregation), verification/validation vs. multilevel analysis and experimental design. Also output analysis of SM models in BPM is treated as complicated task in computer programming rather than tools to support the job of business analysts (a need to use objective data). This is particularly difficult when the real-world system does not yet exist (e.g. a new business). The modeller has then to proceed indirectly (e.g. by collecting data from similar systems) and make additional assumptions that add to the complexity of the SM process.

Decisions have to be made regarding the type of model to be developed (discrete, continuous or hybrid), the SM worldview to adopt (e.g. event, activity or process oriented), the choice of the modelling platform (general purpose, simulation purpose), etc. Although there are special-purpose techniques available to assist and guide SM output analysis, the correct interpretation of results and derivation of useful recommendations, will necessarily depend on the ability of the modellers and decision-makers to make judgments. Output analysis of BPM models should be studied in transient and maximum statistics behaviour, as opposed to traditional SM steady-state (or average) behaviour analyses. A business change may impact the organization on many levels, e.g. processes, the people, the organizational structure, the resources, the information requirements. A well-designed organizational study should present a full business case identifying all possible implications of changes. A natural way is to adopt five-view concept of ARIS (see e.g. [5]) architecture for BPM, distinguishing the two dimensions of organisation modelling: horizontal views (organisation, function, data, control-process, and product/service), and vertical life-cycle (requirements specification, technical design and implementation details). For example, by linking only view perspectives of modelling, the ARIS approach allows to develop approximately 100 models, even most of these models are static representations in a diagram form. As a result, the
implication for SM is that models need to be polymorphic, i.e. they should allow for different organizational perspectives to be modelled and analysed in an integrated and consistent manner. There is also a need to practise mental and generic views, supported by communication sessions.

A simulation model for BPM requires synchronization of the process activities of organization parts. SM should allow for defining sub-models that indicate the decision range of organizational parts to enable them assess the impact of changes on their performance. Therefore, we need SM methods to allow for modular model implementation and for experimentation with selected sub-models. Although SM can theoretically be used for modular model development and use, the majority of existing SM packages does not generally include such characteristics. Modularity should be achievable even if this is not the initial target of the SM task. To enhance model reusability, individual models should be easy to link and to modify. A single model might need to be decomposed to sub-models, when e.g. departments of an organization need to assess their individual performance. A way to achieve problem-free model decomposition and integration is by defining standard model interfaces, but there is a lack of industrial standards to define the interconnectivity issues between simulation model components. A special issue related to SM in BPM concerns the multiplicity of analysis levels within a study.

The development of methodology frame is the first step towards establishing the field of SM in BPM modelling and combining generic SM methodologies with generic approaches to business change management comes first as a methodology concept. This methodology should explicitly address the issues of BPM purposes (BPR, CPI, TQM) to ensure that they are properly managed within an organizational design project. Regarding the SM stages, a generic approach consists usually of 10 steps and focuses on aspects of model development, validation, and statistical control of SM experiments. However, it does not provide guidance how the steps identified can be carried out and supported by techniques. Also some specific methodologies for incorporating SM in BPM modelling for a process change study, e.g. iterative, spiral framework with four main phases (Initiate, Simulate, Experiment, Conclude) and more detailed sub-steps, cannot be a good solution for BPM projects because of a need to build models not only in two-dimension frame: polymorphic (many views), AS-IS/TO-BE representation. The proposed concept for SM integration within BPM modelling projects (Fig. 4) considers an idea of sequential top-down organisation of modelling stages supported by concepts and methods of: soft system modelling (SSM) of non-structured problem situations to reach consensus in BPM, system dynamics (SD) macro-scope SM to analyse dynamic properties of BPM system and processes, DES micro-scope SM to study KPI in BPM, gaming SM to analyse actors interactions, group model building (GMB) intervention SM to solve BPM problems, and organizational learning (OL) to collect knowledge.
4. Conclusions

The analysis and design of organizational processes can be assisted by the development of simulation models. SM can provide a valuable mechanism for addressing the problem of quantitative and qualitative evaluation of prospective designs of business processes. However, there are still many problems to successfully use SM in BPM and frame for modelling, data collection, experimental design, and multi-perspective model analysis, are only some of the issues that have been considered in this paper. One aim could be to develop a comprehensive theory of SM in BPM modelling supported by analysis of special requirements of modelling human behaviour as part of the business environment. It could drive the development of special-purpose SM software to assist organizations in employing SM in organizational studies with classical tools, like IDEF, ARIS, iGrafx. Particularly we can start with development of an organizational group design frame that will build upon SM environment and complement it with a number of other computer-supported tools to assist the various phases of modelling, analysis and design iden-
tified within the BPM methodology. Such tools may: capture data from the actual processes to facilitate design of SM models (data collectors and analysers), support the development of modular, polymorphic business SM models consistent with the SM theoretical propositions (BPM simulators/games), support experimentation with and analysis of SM output results in order to assist decision-making (e.g. special-purpose DSS and KBS tools), support discussions, implementation and monitoring of designed business processes and collect data about performance. However, since the quality of the BPM models themselves affects the quality of process engineering and reengineering outcomes, that are guided by these models, it is important to reflect upon the process of model construction in order to understand its basic macro-scope characteristics (preferred by process owners and managers because of a better comprehension and communication), and gain insight into how these contributions may successfully be employed in practice.

REFERENCES

CUSTOMER’S TRUST MANAGEMENT IN POLISH E-COMMERCE – CASE STUDY

PAULINA RUTECKA \textsuperscript{a)}, KARINA CICHA \textsuperscript{b)}

\textsuperscript{a)} Marketing Manager, sklepEstetyka.pl
\textsuperscript{b)} Department of Communication Design and Analysis, University of Economics, Katowice

E-commerce is one of the most important and fast-growing way of trading. The obstacles common in traditional market seem to lose their meaning here. More and more often the benefits given to the customers are remarkably alike among wide range of stores providing their merchandise via e-commerce service. Those features determinate the question in what way the customer’s loyalty to the particular online store can be achieved and how the trust can be created and maintained. The main aim of the paper is to analyze the case of an online store and its trust-building strategies not related only with price of offered products, as well as presentation of the results of adopted actions which were to ensure increase in sales and to build strong market position of a multi-brand e-commerce retailer.

Key words: E-commerce, management, IT, trust

1. Introduction

E-commerce is still very debatable area. The possibility of implementation changes and improvements, either in IT sector or customer service, makes the dynamics of competition in this market particularly high.

The difference, however, between e-commerce and traditional trading is that not every change requires huge financial outlays. Most of changes implemented into e-commerce platforms are IT based, introduced in stages and tested on selected groups of users in real time.
What is more, factors typical for traditional market seem to be irrelevant in the process of customer loyalty building. Nevertheless, experience shows that online stores are focused on acquiring customers rather than building their trust or their own brand or image, what should be obtained. Implementing IT solutions, as well as IT-based marketing solutions into e-commerce ensures reaching this goal, what can be observed in managing practice.

The main objective of the paper, as well as the main goal for managing practice in e-commerce sector is not only to show the increase in sales of offered products but also to describe the process of brand recognition building among potential customers and steps required to achieve strong market position as a retailer. In order to reach those goals, trust-building strategies are implemented.

On the detailed level the article presents fields in which particular improvements, correlated with the main goal, were undertaken. The starting point was the diagnosis of purchase process in order to build clear customer’s journey path. The next step was to define the moment of transformation customer into client so the particular marketing strategies as well as IT solutions could have been adopted into managing practice. Last, but not least detailed objective was to follow all implementations in real-market conditions in order to analyze their efficiency in terms of consumer’s trust.

2. Research methodology

In 2004 American Marketing Association expanded marketing definition, which had been published in 1985. The currently used definition states that “marketing is an organizational function and a set of processes for creating, communicating and delivering value to customers and for managing customer relationship in ways that benefit organization and its stakeholders” [1].

In order to organize the process accordingly to its definition, what is required the most is any information that a company can obtain about its customers, their satisfaction level, their needs, their brand awareness, as well as their motivation and behavior related to shopping (or in case of e-commerce online purchases).

Gathered information allows to build not only knowledge about firm’s clients but, what is more important from organizational point of view, to predict, to plan and to prepare the trade offer for the customers, which will respond their needs. The marketing process organized around those action may be considered as both effective and efficient.

Information used further in managing process is mostly obtain from researches. Research is collecting specified data oriented on identifying (ex. market share research, market potential research, sales analysis research etc.) or solving (ex. market segmentation research, product research, pricing research etc.) a certain problem [2].
Scientific marketing research process includes several stages such as:
indentifying the problem or opportunity; conducting exploratory research; development of hypothesis; conducting conclusive research; building conclusion [1]. What is more, the subject of the research requires proper research methodology in order to collect precise data.

The research process should start by defying what type of data is needed to be collected. In order to focus on quality of research matter, qualitative research are prepared. The purpose of qualitative research is to explore people’s behaviors as well as their motivation. Quantitative research is used to quantify the problem by way of generating numerical data or data that can be transformed into usable statistics [3].

The marketing research tools can base on primary as well as secondary data. Depending on their type those data are obtained from many sources.

Figure 1. Marketing research tools

In case study, presented in the article, data were gathered by: UX research of a working online store, client’s purchase process observation, telephone interviews conducted by Customer Service Centre and desk research including general overview of business competitors. Collected data were qualitative and referred to the customers’ behavior during purchase process as well as any obstacles which occurred and interrupted it.
The main purpose for this paper was to show how important for any e-commerce retailer is to build trust-based relationship with its customer. First research hypothesis for such an objective can refer to the process of trust formation or, in other words, to the steps that can be undertaken by the retailer in order to gain such a trust as a result of careful client’s behavior observation. The second hypothesis refers to the particular tools that can be used and implemented to form and then sustain long-lasting relationship with the client.

3. Electronic commerce - definition outlines

The Central Statistical Office in Poland defines e-commerce as “transactions carried out by networks, based on the IP protocol and other computer networks” and emphasizes that “the goods and services are ordered by these networks, but the payment and final delivery can be made in both ways on-line or off-line” [4]. However, according to quoted definition, orders received by phone, fax or e-mail are not part of e-commerce. Among enterprises operating in this area these forms are a supplement for electronic sales. They are considered as a communication channels.

The term “multi-channel” refers to the usage of at least two communication channels by the seller in order to conduct transaction [5]. Marketing channel is considered as a "point of contact with the customer or medium through which the company interacts with the consumer" [6].

In accordance to the business point of view and for the purposes of the following considerations those two channels are treated as a part of the e-commerce sales strategy.

The relation-building aspect of e-commerce clarifies when to analyze “E-commerce in Poland 2017” report prepared by Gemius for e-Commerce Poland and Polish Chamber of Digital Economy. In 2017, 54% of Polish Internet users shopped online, which is a 4% increase by the previous year and shows an unflagging upward trend in this market [7].

Among factors, which determine the choice of electronic purchases over other forms of shopping, mentioned by the participants of the research, were: 24/7 availability, comfortability, simplicity, taking less time and saving money.

Crucial information for building loyalty in e-commerce is the opinion of examined group towards the security of online shopping. In 2016, 43% of research participants claimed that shopping online is risky, while in 2017 such an opinion decreased down to 38%.

The results of these studies indicate growing consumers interest in online shopping and willingness to use e-commerce services.
4. Customer’s trust

Marketing defines trust as building consumer relationship through trustworthy dialogue and unbiased information [8]. In case of out-of-date or questionable information the result may be in abruption of shopping process.

4.1. Process of purchase

Entrepreneurs in e-commerce prepare marketing strategy based on possible and diagnosed consumer behavior, which can be shown on a funnel diagram.

![Figure 2. Sales funnel](image)

*Stage 0 - Awareness:* The consumer is aware of the need of a commodity.

*Stage 1: Interest/Visitor:* The network user initiates the product purchase most often starting from entering general key phrases into search engines.

Building brand relationships with external stakeholders at this stage is mainly limited to creating convincing marketing messages.

*Stage 2: Consideration/Sales Lead (Subscriber):* The network user begins the initial recognition of the Internet as a market, selects online stores with products which meet his needs and visits websites of selected online stores.

In this stage, the website visitor has several options for further action, which will finally transform him into Subscriber: to narrow the search or to leave contact details in order to use store’s advisory offer.

In response for visitors action and to build trust and consumer loyalty e-commerce store may optimize web page content quality through providing with useful and reliable information (both: product description and trade offer), access
to product opinion and UX optimalization (ex. CTA), as well as contact customer via phone or e-mail or implement marketing automation solutions.

Stage 3: Purchase/Client: Subscriber (sales lead), deciding to finalize the transaction and stepping into purchase stage, becomes a client. A client is understood as a user who made a single purchase in any online store.

In order to induce safety an online store provides client with UX check-out process optimalization, security of payment system and e-mail post-transactional information.

Stage 4: Returnability/Returning client: Satisfied by purchase and transaction service client is willing to make another order. By doing that may be considered as returning client.

The store may support client’s returnability by marketing actions (ex. cross-selling offers, loyalty programs, extended “Thank you page”) or sharing discount coupons for next order.

Stage 5: Buzz/ Fan: The client, who makes many purchases and who is very satisfied by any of them as well as the customer service and order fulfillment identifies with store brand and recommends it along with the products (buzz marketing).

4.2. Trust management as a marketing strategy

The importance of trust in e-commerce is meaningful. We define trust management as “a set of actions to create systems and methods that allow individuals to make assessments and decisions relating to the reliability of potential operations involving risks, related to other units (assessment of the credibility of other entities), and enabling the participants and owners of these systems to grow and proper represent their own credibility” [9].

Building one’s own credibility is a complex and multi-stage process, however building trust will in the future attract loyal customers, as well as expanding the range of commercial activity to consumers currently unconvinced to this form of shopping. This expansion may be considered as a form of reaching an important market segment defined as “a group of buyers within the market, created by clients with similar needs” [10]. The main need of this segment is to trust the entity responsible for the online orders service.

Lee-Turban model of consumer confidence for online purchases emphasizes that trust depends on factors such as; trust in the e-seller treated as an element of its credibility; confidence in technology; trust in transaction security [11].

Therefore, an e-commerce company, building a trust management strategy, should pay special attention to those aspects that satisfy consumer’s need for trust by: selecting appropriate logistic partners; selecting reliable and product-safe packaging service; offering safe and proven systems for handling online payments; providing reliable information about products availability; providing authentic
photos of offered products; an objective and comprehensive description of the products proposed; professional customer service (allowing for easy contact by phone, e-mail or personal with the store); clear and intuitive interface of the website; ensuring participation in credible opinion-forming programs; transparent shop regulations, returns, complaints, privacy policy of the user.

The elements mentioned above are necessary to build trust. In order to implement them it is crucial to involve all the staff of the online store, at every level.

The value of trust and its existence in business is unquestionable, the issue of measuring it becomes more problematic, though. There are, however, factors which can be observed and measured that can indicate the level of customer’s trust. For such indicators one can consider: external programs for researching the level of client’s satisfaction; cyclicity of orders; engagement in social media; indicator of emails openness; loyalty programs statistics; usage of discount coupons.

5. IT solutions for e-commerce

E-commerce is characterized by measurability of its clients operations. Modern e-commerce platforms to measure customers’ activity use their own analytic tools, as well as external systems. Common practice is to summarize statistics from different sources in order to create extended reports.

5.1. Tools for customer’s trust analysis

The most popular and accurate tool is Google Analytics, which in its basic free version indicates the traffic on the site divided into user’s sessions and unique visits. Calling the session several times suggests that the user is considering purchase, analyzing products or information about the store itself.

An important information that can be obtained from the same tool is the user’s journey map. It indicates, for instance whether the customer is familiar with store regulations, or looking for detailed information about the product.

Google Analytics also provides with special e-commerce reports that indicates turnover within a particular brand (in the case of multibrand stores) or within a particular product. Analyzing products with a similar features and price, of which only few are characterized by significantly higher sales, one can suspect that the description or photos for the rest of the products are not convincing and reliable enough. One can also assume that the the customer may be aware of a negative products’ or brands’ opinion. In this particular case the trust carefully built by the store towards its offer may be undermined not because of its own faulty strategy, but rather due to external factors.
An interesting tool for trust level measurement are programs recording the user's path. The installed script allows to register mouse movements, page scrolling, highlighted and user-clicked elements on the page. Analysis of the sessions recorded in this way perfectly shows on which elements of the page the user focused longer.

Specialized tools for analyzing various aspects of the website will indicate elements necessary to examine the level of customer satisfaction and trust, as well as verify the effectiveness of the implemented strategy of building client’s trust.

5.2. CRM and marketing automation as a key tools for long-lasting relation building

Customer Relationship Management (CRM) in the most general definition is understood as “a set of procedures and tools important in managing customer contacts”[12]. CRM is thus understood in two ways, as a business strategy or as a definition of software used to manage customer relations.

CRM as a kind of software enables collecting customer information in the form of an index. Customers can be segmented based on any criteria. Most often, they are also placed on a specific stage of the mentioned sales funnel.

Using the CRM system the information about the client can be entered into database both manually or automatically. The first situation refers to all information obtained by the sales consultants while communicating with the client (ex. through phone). The second case involves scripts operating on the website, which track clients movements. What is particularly important is that the rules for the scripts are configured individually by the marketer, according to the needs. Obtained information is stored, but not every single one is used within the process of offer personalization.

Modern, technologically powerful CRM systems coupled with marketing automation systems allow to direct client's movements in an automated way. As a result of performing specific actions on the website (ex. purchase, entry to the product card), the customer receives subsequent tags or in response to a given event the system performs assumed action (ex. sending an email). The actions performed by the system are strictly correlated with client’s purchase map or with estimated preferences based on previous behavior online.

Using meticulous annotations about customers, their preferences, previously used products, and even life situations mentioned in the conversation, strengthens the relationship between the online store and the customer. That suggests to the customer lack of anonymity and has a real impact on the client's sense of security.

Each time building a trust management strategy including CRM tools and marketing automation must be individually constructed. The use of IT tools in the communication process, conducted analytical activities and constant improvement of online store’s UX are the key factors necessary to manage the trust of customers.
Online stores implementing these elements and effectively managing customer trust can significantly differentiate their offer from the competition, depart from the patterns of activities, practiced for years, thus saving significantly the advertising budget. At the same time, focusing their attention on maintaining customer relationships, they can engage much less funds than those necessary to acquire new customers.

6. Case study

The following case study was prepared based on the example of the online store, operating in the dermocosmetics sector. A characteristic feature of online stores in this industry is having the same range of products offered at the very similar prices (suggested by the producers for retail distribution). Those factors need to be taken into consideration while building the customer’s trust strategy.

In the initial phase, analyzed store has adapted to the common methods of operation among competitors. Product prices were reduced to the lowest. In response to the prices offered by competitors on the market many times the price was lower by one grosz. Lowering prices gave the opportunity to obtain the first position in price comparison websites, which consumers are willing to use, and which can sort by prices. However, it turned out that such an action causes a never ending price fight until it falls below the break-even point.

The next step aimed at increasing the number of additional benefits that the customer will achieve by choosing the store's offer. The store focused on adding more and more attractive gifts to each order. Mostly those costly giveaways were cheaper full-size products, samples or gadgets. Competitive store used the same model of operation at the same time. The addition of freebies, however, started to increase costs of every transaction and the margin of offered products once again fell below the break-even point.

The next strategy was to implement a free delivery to the order, which also increased the transaction costs for the store. As in the previous strategies, in this case the value added effect was quickly compensated by the competitor. Each of these activities was aimed only at increasing the attractiveness of the offer in order to attract a customer, not to gain his loyalty.

The described activities were undertaken on the basis of analysis that the store had been conducting. It was observed that the client was not willing to return to the store to buy another product even though in case of dermocosmetics the effects of they usage are visible only after using several packages of certain treatment. Conducted analysis also showed that customers who bought the first product, did not buy another one, or bought a random product from, for instance, five-step treatment. The conclusion of this analysis was that the client, despite the need to continue the treatment, was not loyal to the store in which he was recently buying.
but he started the purchase process from the beginning, searching for stores offering the product and choosing the most attractive offer at that time.

Due to the fact that store and its competitor had reached the same point, in which the same products were offered, at the same prices, with free delivery and the same range of free gifts, and hence, there were no factors that could distinguish any of the stores in a way that attracts customers, special attention was paid to building loyalty and customer service, creating a separate strategy for these elements.

As part of its activities, described store began to pay special attention to the dialogue with customers in two areas: interpersonal communication maintained by the Customer Service Center (CSC) and automated communication via e-mail and SMS.

In order to improve interpersonal contact maintained by the CSC, quality monitoring standards of conversations were introduced. The procedure obliged the CSC employee to create annotations about the client based on the information obtained during the interview, as well as periodic contact with the client in order to present a new offer or inform about promotions.

As part of automated communication improvement, apart from regular mailing of newsletters, the store introduced sending segmented newsletters and post-trade messages encouraging participation in external programs for opinion polling.

None of the introduced changes was directly related to the purchase stage. The changes focused on building client’s trust in the consideration stage as well as sustain it in the returnability stage.

7. Conclusion

The customer’s trust towards the entity performing the transaction is particularly important in e-commerce. The implementation of appropriate IT tools, meticulous analysis and improvement of systems in terms of user experience has a real impact on the level of trust. The implementation of the trust management strategy influences the attractiveness of the offer and the competitiveness of the store among others, operating in the same market sector.

All mentioned and described actions may result in the increase in number of customers, returnability of customers, and undoubtedly clients’ opinion about the store and its offer. Those factors are the background for considering the store as a reliable and trustworthy business entity, as well as are the key for its positive image building.

What is more, careful trust management strategy leads to advertising costs reduction aimed at attracting new customers. The returning customer is known to the store that answers accurately his needs and maintains a long-term relationship
with him. A motivated customer whose loyalty was acquired by the store in a well-thought-out management process can become a reliable source of store’s income, as well as an informal marketing tool, by sharing thoughts about the store, offered products or customer service with other potential clients.

REFERENCES

ENHANCEMENT OF SCRUM-BASED PROJECT MANAGEMENT LEARNING EXPERIENCE BY USING WEB APPLICATION

WALDEMAR SOBIECKI, ANNA KURZYDŁOWSKA

Faculty of Mathematics and Natural Sciences, Cardinal Stefan Wyszyński University in Warsaw

In recent years project management has become a complex and diverse field of scientific research. It embodies a number of specific tools and techniques, choice of which to particular project is far from being obvious, especially for the beginners. Such situation rationalizes efforts aiming at the development of methods of training, which are more efficient than the “trial-and-error” approach. In this context the paper presents a web application that simulates real life scenarios and allows for developing skills, among others, in identifying the needed/available resources against the dynamically changing goals. The platform teaches the agile approach to project management with Scrum and gives an opportunity to “materialize” the progress made by users with an original piece of software. The prototype of the web application has been successfully tested at Cardinal Stefan Wysynski University.

Key words: Traditional Project Management (TPM), Agile Project Management (APM), Scrum, Project Management Life Cycle (PMLC), Software Development Life Cycle (SDLC), E-learning

1. Introduction

Today’s society is often described as “Information Society”. We are surrounded by unprecedented and ever increasing variety of information and technology. This leads to the question of how to utilize available resources in the most efficient way.
This article gives a glimpse of how to combine theoretical knowledge with implementing real solutions in IT projects. In order to better visualize the idea, author created an application that has two main objectives – to teach and to provide an environment for developing working it solutions. The application was specifically crafted for arguably one of the most popular agile framework called Scrum.

1.1. Definition of a project

There are multiple definitions of a project, some can be found in international standards such as: Projects in Controlled Environments (PRINCE2), Association for Project Management Body of Knowledge (PMBOK) or ISO standards (i.e. ISO 21500:2012 standard). Some definitions have been thoroughly formulated by researches, while some by practitioners. Nevertheless most of them define project as a temporary endeavor undertaken to create a unique product, service, or result.

In addition to this, projects should meet additional criteria and take place under some specific circumstances. The most important criterion is the goal of a project, usually best reflected by the scope of project. Another important feature is the interplay between the needs to be served, the time and cost of implementation, often referenced as a project triangle. On top of that, projects are executed in limited multi-resources environments (people, knowledge, technology). As a result they carry specific risks. According to the Project Management Body of Knowledge Guide all these aspects render projects their temporary nature [22].

1.2. Project management

Over the years, it has turned out that multiple factors influence projects; thus, initial expectations of the customers and the outcome of the projects finally accepted by them differ significantly. Projects can be finished with satisfaction of the customers only by applying some specific rules that make successful completion of the projects possible. The first and foremost important concept in this context is concept of project management.

Depending on circumstances different project management strategies can be used. The goal is to impose processes and procedures that do not limit creativity of the people involved, by providing them with the tools to efficiently meet project requirements.

A person that is responsible for selection of proper tools and techniques that will be used during the project is called a project manager. Since each project is different, the essence of project management is not in the routinely repeated solutions that have worked in the past. The inherent nature of the project implies that project managers
need to be adaptive, creative and flexible. At the same time, however, there are some general rules that help them identify the right paths.

PMBOK defines project management as an application of knowledge, skills, tools, and techniques to project activities to meet the project requirements [22]. It involves client to meet sponsor’s needs and deliver a new business value. PMBOK marks out five process groups which can serve as checkpoints during project realization:

Initiating - This is a phase when project manager is assigned to the project. His/her task is to evaluate the current situation. He/she learns the business needs that are presented by the sponsor, who represents the customer. As a result a vision of the project is created. More detailed requirements are documented by the project manager, as they will be needed later.

Planning - In this phase a general vision is described in detail (defining a scope of the project) and project plan is created. This plan covers topics such as budget, milestones and risks.

Executing - Project manager executes project according to the plan. In traditional approaches to project management, unforeseen changes that are reported by stakeholders should be first agreed with the central board. When approved, they can be added to the scope of the project.

Monitoring and Controlling - This is actually the only phase that is not linear. It is a perpetual exercise that lasts as long as the project is not finished. It consists of tracking, reviewing and regulating applied to all actions that are taking place in the project.

Closing - It is a time of a formal closure of the project. Before project is officially declared as “done”, few things need to take place. First, the outcome should be accepted by the customer (sponsor). It is also a good practice to have some sessions on lessons learned during the project. Such sessions allow to formulate conclusions and make improvements in future.

There are also different approaches to project management life cycle (PMLC). The choice of PMLC should be done based on analysis of project goal vs. solution. The solution describes ways in which the project goal can be achieved. When both the goal and solution have been analyzed, then one of the four approaches to PMLC can be adopted [3].

Traditional Project Management (TPM) - used when both goal and the solution are known. It is a plan-driven approach which means it needs a detailed plan preceded by WBS (Work Breakdown Structure). Traditional Project Management works well for
all the well-known project subjects, such as infrastructure projects. This approach should be taken into consideration if there is very little chance of unforeseen situations.

Agile Project Management (APM) - used in the case of the projects that are “change-driven”. Such projects usually have a general goal, but the way how to achieve it is not pre-determined. Thus actual solutions are discovered as the project continues. An immanent attribute of agile projects is that the scope and resources might change over time, which makes it difficult to forecast the end result.

Extreme Project Management (XPM) - Extreme approach suits best projects that are high-paced and with high failure rates. Such projects often have neither known solution nor specified goal. A typical example are Research & Development (R&D) projects.

Emertxe Project Management (MPX) - the name comes from “E-x-t-r-e-m-e” read from right to left. MPX projects are defined as reversed R&D projects. The difference is that in R&D projects although the solution still needs to be figured out, there is usually some desired end state. In Emertxe the goal is not defined. In such projects there are only tools that are available and can be utilized, and it is a major challenge for the project manager to find good use for these tools and create an added value.

1.3. Software Development Process

In projects that are strictly set on developing a software, the concept is very similar and is called a Software Development Life Cycle (SDLC). The difference between them is that PMLC is more suitable for establishing life cycle cognizance first and SDLC is more for adapting development models for delivering the technical content.

All above concepts focus on the common problem: how to minimize the difference between the wants and the needs. The traditional approach is not always sufficient to meet the constantly changing business needs. Requirements are in the center of the client interests and are often redefined. That imposes additional responsibility on project managers as to the proper choice of SDLC model.

There are two main approaches to SDLCs. The first one is called a traditional project management (TPM). It has a long history of successful implementations. It mostly relies on linear and incremental build-up of a software. Some of the most popular systems used in this context have been listed below:

- Waterfall
- V-model
- RUP (Rational Unified Process)
The other group is the Agile (APM) approaches. Despite their diversity, they all share an idea of working in an iterative (incremental) way. Such methodology allows for a quicker response to customers’ needs. The Agile methods are relatively new and are gaining on popularity at the expense of the traditional methods. There are numerous implementations of the Agile approach, some of them have been listed below:

- Crystal
- Dynamic Software Development
- Feature Driven Development
- Lean Software Development
- Extreme Programming
- Scrum

2. Scrum

Uncertainty and constant change of business requirements are inseparable parts of IT projects. In order to achieve goals one needs to be flexible and react quickly. Scrum has been used since 1990s for actions such as: developing, enhancing, releasing and sustaining new products [13]. Over the years it has proven over and over again in complex projects being run in an uneasy market conditions. These successes have built its reputation of a flexible and reliable tool for solving complex problems. Nowadays, Scrum is one of the most commonly used agile approaches when delivering innovative products or services [11, p. 34]. Official Scrum guide highlights that it is neither a process nor a definitive method, but rather a process framework, that delivers a structure and allows to employ various processes and techniques [13, p. 3]. It distinguishes three major components (dimensions of the projects), that are: people, artifacts and ceremonies. These three, if combined in the right order, allow for reliable and efficient problem solving.

2.1. People in Scrum

The core of each Scrum endeavor is a team. The teams involved into project implementation were originally divided into two funny sounding groups, based on metaphor of pigs & chickens. However, controversy built around that segmentation has been solved by replacing the original classification in the official Scrum Guide between 2010 and 2011 [14]. Today Scrum teams are defined in a more structured way, i.e. below.
Developers - it is a group of people that are directly involved in creating a final product. Teams should be self-organizing and cross-functional to ensure comprehensive, interdisciplinary approach to the problems arising. Each member of the team should possess a wide range of skills and be flexible enough to replace his/her colleague if necessary (e.g. in case of illness). A team should not exceed nine people, as experience shows that the most effective teams have from three to six members [2, p. 90]. It does not imply that Scrum can only be used for solving “small problems”. For more complex projects, there are usually multiple developer teams.

Product Owner - one of the most common needs during project realization is to manage the changes, in particular changes of the business requirements. In order to minimize turbulence, client in Scrum is represented by one person only, who plays the role of a Product Owner and is responsible for negotiating possible project adjustments. Assigning just one person to the role of Product Owner reduces the time required for negotiating requests for changes. Having multiple Product Owners would make the whole process practically impossible to manage. Obviously, the Product Owner might look for advice of his/her colleagues.

Scrum Master - the primary goal for the Master is to help the team remove any obstacles that arise. Scrum Master manages internal relations between members of a team and takes care of organizing meetings and controlling deadlines.

2.2. Artifacts

Artifacts are all the widely understood “tools”, that Scrum delivers to people that are engaged in the project. The most crucial ones have been briefly described below:

- **tasks** - elementary units of work that a developer can produce. Tasks are combined into stories, which in turn, create a final product. In practice, tasks are usually defined by decomposing stories into smaller chunks (method known as “principle from general to specific”). Product owner should take care of setting priorities and updating sets of stories, so that they match changing needs of client.

- **story points** – points are assigned to each story. They should reflect time needed for the story to be completed.

- **scrum board (task board)** - stores stories, this is usually a physical board with post-its stuck to it. It should be situated in a central point of office so that each team member has an access to it. The simplest board consists of categories: to do, in progress, completed [17, p. 19].

- **product backlog** - that is list, that contains all stories. Stories should be prioritized (from the most important to the least ones). The most important
stories should be placed at the top of the board. As the stories go up on a board, they should be explained more in detail (extended) so that they are understood by each of the team developers. Here it is important to define list of conditions that have to be met to declare that the story is “done”. This term is used to indicate that the current piece of project as potentially shippable. When the definition of done is met, then story can be moved to completed section on the board [11, p. 422].

- sprint backlog - these are stories chosen from product backlog that will be created during current sprint. The time estimated for the stories should be aligned with the length of the sprint. In other words, stories should be realized during an ongoing sprint (without a need to move them into next one).
- burndown charts - are showing how fast the team is moving forward. X-axis shows time (or next sprints) and y-axis shows number of story points that still have to be completed. Burndown charts are also taking into account abnormal situations or pauses such as employee illnesses or appearance of new stories.

2.3. Ceremonies

Ceremonies are the events that occur during project development. In order to perform these events, teams need to use previously presented artifacts in a very precise order. They combine theory of Scrum with all the tools it delivers into subprocesses that in the end lead to the final product.

Sprint planning - takes part before each sprint [11, p. 359]. Its purpose is to prioritize stories from the product backlog, so that the progress in product development is possibly the highest [2, p. 97]. It is a time when team agrees on what will be built in upcoming sprint. Product Owner does not have to take part in the whole ceremony. However, if he/her wants to point out some additional functionalities/features he/she is allowed to do so. It is worth noting though, that it is up to the developers to decide if and which stories can be successfully realized during the sprint. Once stories are chosen they are usually split into smaller tasks. The outcome of sprint planning should be a sprint board with all the stories nicely divided into tasks.
- daily Scrum - short meeting (up to 15 minutes) of Scrum Master and the developers, there is no need for Product Owner or stakeholders to participate in it. It is a brief communication about actual status and the progress made so far [2, p. 97]. There are three main questions that every participant of this meeting should ask: “What have I done yesterday?”,”What am I planning to do Today?”,”What are the obstacles that are blocking me?” Role of Scrum Master is to actively eliminate obstacles and solve the issues that developers are facing.

275
• story time - it is used to upgrade stories in product backlog, also might involve changing criteria of acceptance [17, p. 29]. This sometimes is considered as additional ceremony, but during this time many important subjects can be discussed such as: criteria of acceptance of stories, conditions it has to meet and how different developers understand those rules - usually criteria might differ depending on point of view of each developer. It is good to have the same understanding of tasks that will be implemented as it will improve the final outcome.

• story sizing estimation - it is an action of time estimation needed for story (tasks) to be completed. Usually determined in time scale (in rare cases it can also by represented by symbols or words). Breaking down the estimation time into points is more useful as it can be later used for sprint planning, where stories need to be selected in such a way that there is a high probability that all the stories will be finished within one sprint. The crucial thing is to relate the estimated time with other stories. One of the most commonly used techniques for story sizing is Planning Poker. Each developer gets special cards (with story points) and shows them at the same time. If the difference between minimum and maximum value is substantial, the “poker players” explain their estimations. After that voting goes through next iteration until all developers agree on timing for discussed story [19, p. 98].

• sprint review - it is a meeting that takes place on each last day of sprint. The main goal is to demonstrate a potentially shippable product (or the outcome of the sprint) It is a good practice to plan it on early morning hours and inviting all stakeholders. Invitations should be sent sufficiently early, so that all can come. Scrum Master and the developers team collate all comments so that they can be taken into account in future sprints. They present stories that have been completed. Ideally all stories that were chosen should be finished [12, p. 98]. End of sprint usually means a beginning of the next one. Usually at the end of the sprint review there is a session called “preplanning” which already involves planning the next sprint.

• sprint retrospective - this is an internal meeting of a developers’ team (without stakeholders). The purpose is to summarize the efforts that have already been made. It is the time to rethink current actions and present new learnings to all team members. Good practice tells that retrospective should take from one to two hours for every week of work on the project [17, p. 33].

• release planning - sometimes the Scrum process is enriched by the process of long term planning (taking into account more than one sprint). In this type of
meeting the product owner should join the team of developers to discuss stories from the backlog, setting schedules, analysis of the workforce capacity and spotting new opportunities as well as identifying potential risks. It is a good practice to set a timeline and time limit for this kind of meetings [20, p. 22].

![Scrum process](image)

**Figure 1.** Scrum process

A Fig. 1. shows in a nutshell how Scrum proceeds through next iterations (sprints) in order to deliver a piece of final product. Each of such pieces is later shown during a sprint review and can serve as a reference point in identifying the current status of the project.

### 3. E-learning system for the Scrum methodology in IT project management

#### 3.1. Existing solutions and challenges for E-learning

There are many available e-learning systems. One of the most complete approaches to E-learning can be realized with the help of computer software in form of Learning Management Systems (LMS). Main goal of LMS is not only to deliver an environment for learning, but also to make process of management easier and more structured. Despite the fact that LMS are usually focused on the administration processes, they are also open for other forms of learning processes such as: blended learning (form of learning that combines traditional approach with distance learning) [8, p. 60]. LMS provide tools that allow to track students’ achievements as they progress with the course [9, p. 9-12]. They provide functionalities that can be used while performing daily activities such as: management of educational materials, schedules, reports and individual assessments of participants.

The European LMS Market Dynamics research shows that one of the most commonly used LMS in Europe is Moodle [1]. Moodle stands for Modular Object-
Oriented Dynamic Learning Environment and is distributed on an open source license. It has a lot of features that allow for networking. As creators of Moodle say, the biggest strength of this solution is the switch from just storing educational materials on a server to building a supportive community in the process of gathering knowledge [10, p. 19-20].

The growth of e-learning over the years has led to a concept of MOOC (Massive open online course). It is a concept of open courses that are widely available for a mass receiver. It is called massive because the number of listeners greatly exceeds numbers of participants of normal e-learning courses. They are usually free of charge and make use of the most up-to-date tools and technologies. Companies like: edX, Harvard Open Courses or Coursera often offer courses held by prestigious universities, such as: Harvard, University, Berkley University of California or Massachusetts Institute of Technology. In the United States of America, institutions running such courses are accredited by the American Council of Education. This makes it possible to acclaim results of the students in the traditional learning process [21, p. 58].

One of the biggest difficulties when learning is to find the linkage between academic knowledge and real life. As an example one can refer to the graduates, who have very strong theoretical background but may not yet have professional experience. For such users it might be worth trying a more customizable approach. Something that would be flexible enough to adjust to the skills and needs of participants. A solution that takes into account needs of novice programmers and project managers learning Scrum methodology will be presented in the next section.

3.2. Main idea of the application

The application presented here is a web application written in Python (back-end) and JavaScript (front-end). It offers an environment for a student to place him/her in an imaginary scenario, where he/she is a part of a programming team charged with the goal of manufacturing a piece of software that is cut according to the vision of course creator or Product Owner.

The intent was not to create a fully complete solution (as that would require much more time), but rather to familiarize and stimulate participants to use Scrum approach with their further real life development projects.

The presented solution was inspired by the idea of gamification, which is based on simulating different scenarios that require from participants taking good and accurate decisions [5].

Since accessibility is a key feature of an e-learning idea, a presented system is a web application. Before use, each user is obliged to create an individual account,
which can later be registered to one of the existing course templates. Each course starts with a description and some predefined settings and tasks, which are open for modifications. Once user joins the course, he/she has two Scrum sprints to create all the tasks (number of sprints has been limited for the sake of simplicity). As a result the team should create a potentially shippable product that should fulfill the requirements set up in the course. At the end of the course students not only should have a better understanding of agile methodology such as Scrum, but also they should be able to continue their project outside the application. The main idea is to familiarize users with the concept of Scrum, project management and the opportunities of “organized” cooperation.

One of the biggest challenges when creating an application like this is to replace these Scrum elements that need a “true” interactions, such as scrum meetings or the concept of stickers used on a scrum board, by a virtual solutions. Solutions to that problems have been presented in the latter part of the article.

3.3. Who takes part in a process

There are several roles defined in the presented application. The first role is a role of Product Owner, which symbolizes a vision of the final product. Product owner might invent changes or change existing system requirements. He/she is responsible for all actions that are coming from outside of the team. He/she focuses on the product being done on time. His/her vision might change over time. Next, there is a role of Scrum Master - one person responsible for organizing the team. Scrum Master resolves conflicts, makes decisions when needed and tries to keep the whole team motivated to deliver a final product on time. Each sprint ends by showing a potentially shippable product that is created by a team of developers. Although developer role is the most basic one, it is essential for the whole course - this is where most of the job is done. The primary job of developers is to write code that corresponds to what has been agreed on user stories. They need to cooperate to efficiently deliver the final product. Apart from the users, that are directly involved in a realization of the course, there is also a role of an administrator which allows to supervise each course and make system modifications if necessary. Due to simplicity of a current application, some tasks requiring both changes in the application or specified course can be done by administrator.

The software presented here is not a fully complete solution, some simplifications were used, so roles are limited. In real projects there should be a much greater variety of roles, such as: testers, UX designers, graphic designers, software engineers etc.
3.4. Course workflow overview

The goal of this e-learning system was not only to guide students through the course, but also to allow for as much of their own invention as possible.

![Scrum Learning application workflow](image)

**Figure 2.** Scrum Learning application workflow

Fig. 2. presents main stages of the application. The general direction of work goes from left to right and from top to bottom. The subsequent horizontal stages are only available when Scrum Master enables them (they are locked by default). Sections indicated in Fig. 2 are presented below, sorted according to their order of occurrence in a project.

3.5. Learning stage

The learning stage begins with introduction to the application and describes the goal and mechanics of the system as well as provides some general information about Scrum, team management and programming techniques.

Application allows to upload some of the most commonly used formats such as .pdf, .mp3 or .mp4. Course creator role is to put all the files that should be necessary in the process of learning Scrum (such as Manifesto for Agile Software Development) as well as materials that concern a specific programming language (JavaScript is used for solving tasks).

Apart from that, learning stage covers topics that are specific for the project and would be necessary to understand the topic of the course. For instance, if the project was to develop software for a chess game, it would be nice if the rules of the chess were somehow explained in this section.

The last part of this stage is a test, which has to be passed in order to proceed to the next stages of the course. It ensures that every participant of the course has acquired the required knowledge. It resembles a little bit an entrance exams to the universities and is essential for assuring that everyone is “on the same page” with the
course reading. Without passing the quiz it is not possible for the students to proceed further.

3.6. Project details

First step for starting a project is to assign roles in a team (forming a team in Scrum is one of the most important tasks). In the “current progress” section users can see details of the course such as: deadlines, participants and their roles, burndown charts, number of user stories etc. Each user can access this site at any given moment of the course.

At the beginning this section remains empty but as the team makes progress with their tasks more and more information is presented. Product backlog stage is one of the most important ones. It is here, where the Scrum team creates stories that will be realized. Some stories were predefined by the course creator. Other might be created by the users itself, if they want to create something extra. This gives the team an independence and makes each course very different from others.

General discussion section is a place where every participant can share their opinion on a public chat. Communication in Scrum is very important. It would be best to communicate in person, but that is one of many simplifications of the Scrum rules in the application, to fit into e-learning criteria.

Scrum poker is a place where team members estimate the time needed for stories to be completed. It is a very effective technique that allows to reach an agreement, even if first estimations are far apart.

3.7. Sprints

Once stories have been estimated and moved to the product backlog the actual work can be done. Each sprint starts with selecting stories from product backlog to the sprint backlog. Users together come to agreement on how many stories can be delivered during the sprint. Once selection is made, developers may proceed to sprint board (as shown in Fig. 3). Each developer can choose a story that he will be working on and move it to Work in progress section on the board.
Once the developer decides the story is ready he/she can move the story to “To be reviewed” section. In this stage, automatic tests of the code can be carried out in addition to any required manual test. After the story passed tests it can be moved to Completed section and story points are accounted to the current project status. The developer can pick another one from Not started section and start working on it. If story has not passed the tests or the person that was doing a manual inspection did not accept it, then such story needs to be rewritten to match the requirements.

At the end of each sprint there is a retrospection. This mechanism of presented application checks which stories have been finished and have met the criteria defined in the tests.

The goal of each sprint is to develop a shippable product increment. When sprint is finished, the system gathers all the stories from Completed section, combines them and presents results on a screen. This allows product owner and each member of the team to see the current progress of the project.

When the course is finished students can keep working on a project on their own by uploading files with code to the version control system of their choice or keep the results as a proof of their competence. They could also create another course and refine the project.

3.8 Project summary

The essence of Write reviews stage is a sincere opinion about each member of the team. Each participant has to fill a review form about other participants. The form is
simple and has only two fields. A brief description of a user goes into the first field. The next field stores an overall grade (value from 1 to 5). Reviews are anonymous.

In Read reviews section each user can read his/her reviews without knowing their authors. This encourages all the participants to create honest reviews.

It is important to note that that Project summary should not be the only stage that allows for assessment as it should also take place during solving tasks and handling inter team communication during Scrum Poker or General Discussion phase. Such an approach to assessment is called formative assessment and is oriented toward individual student. Its effectiveness improves with the number of assessments. For that reason they should be performed throughout for the whole duration of the course and not just at the beginning or at the end. Systematic assessing stimulates students’ involvement and systematic work that can boost their motivation and engagement. Finally it results in more thorough understanding of the material [6].

In contrary, summative assessment usually checks degree of subject absorption [7]. The tests are a valuable source of information for teachers as they can identify what parts of material have been well understood by students and which need a more in-depth explanation or revision. However, this approach is often criticized for leaving student on his own. If tests are taken at the end of course, it might be difficult to realize that some students are having problems with specific part of material in time.

4. Summary

The main objective of this paper was to present a new tool for training project managers and developers by offering to them the learning environment imitating the major aspects of the Agile projects. Since the project management skills greatly improve with the growing experience, the hands-on experimenting with various scenarios offered by the web-based system described here is likely to provide a helpful insight into field of project management.

It should be noted that the application at this stage is a prototype open for modifications. There are multiple ways that it could be further developed with one idea being to increase the number of parameters to current Scrum settings e.g. defining length of the sprint stage. Another idea is to build a tool, that would actually not only serve as a learning platform, but could also be used in a real-life projects. Such change would require by-passing the currently existing restrictions on number of iterations and project roles and increasing the role of product owner. These are, however, only technical issues, which can be solved with relative ease.
REFERENCES


CONTEXT-SOCIAL MODEL OF PROSUMPTION IN E-COMMERCE - ANALYSIS OF A PROTOTYPE SOLUTION FOR CLOTHING INDUSTRY

TOMASZ WOŹNIAKOWSKI, MAGDALENA NOWAKOWSKA, RAFIK NAFKHA

Warsaw University of Life Sciences (SGGW)

The aim of this paper is to look at cooperation practices between producers and prosumers in e-commerce and to explore future possibilities. The paper presents project of an innovative solution utilizing habits of social media users to provide valuable suggestion to clothes’ purchasers. Described system will enable not only acquiring opinions on specific products by also its correlations with other items. Recommendation system could be also utilized by clothes manufactures and sellers to deliver valuable data about consumers real opinions and preferences. The last decade has seen a considerable increase of online shops for fashion goods. Technological advancements, improvements in logistics, and changes in buyer behavior have led to a dissemination of apparel goods and respective data on the Web.

Key words: Prosumption, rating system, fashion industry

1. Prosumption

The modern consumer creates his identity through the ownership of goods, because he believes that quantity and quality of items possessed transfers to his status, social prestige and recognition [1]. This conviction is supported by the producers and politicians who describe consumption as necessary condition of economic growth in capitalist economy. To satisfy his need for unique, tailored
products customer has to take part in production process by specifying design requirements or by partially customize product himself.

The term ‘prosumption’ originates from combination of two words: ‘production’ and ‘consumption’. Prosumption means integration of production and consumption processes, until the boundaries between them becomes blurred [2]. Consumer and producer function are combined, however in different way than in traditional agricultural societies, when prosumers were independent and self-sufficient [3]. Prosumer nowadays is an active and conscious consumer who wants to be co-creator, not just a passive recipient. Prosumers are more aware of their rights as well as theirs position on the market. That’s why companies are forced to change their strategies and treat clients as equal and establish new kind of relationship with them. This is enforced by new technologies which give consumer quick access to information of product and services, opportunity compare prices and ask other whether merchandises are worth buying [1]. Prosumers are not afraid to experiment and they think for themselves. They are opinion leaders; other people value their point of view and advice given. Moreover, prosumers are characterized by skepticism, since, although they appreciate the media and advertising as a source of information or entertainment, they do not take the media uncritically [6].

Web 2.0 prosumption

Web 2.0 differs from Web 1.0 not only in technical aspects enabling richer user experience. The main difference is user’s participation in creation of content. Users are no longer passive viewers. Instead they are forming online communities, they blog, add content and comments to network discussions, contribute reviews, share experiences and opinions. For manufacturers these users are valuable source of feedback because they provide new ideas for products’ improvements. As a result a large part of product know-how is generated by actual end-users of the products. These prosumers often substitute manufacturer’s product support, answering beginner’s questions and delivering quicker and more adequate solution to common problems. Honest opinions about product, demonstrating not only advantages but also limitation sometimes even including links to alternative products may be not enthusiastically welcome by manufactures.

Manufactures gladly receive consumers feedback in early stage of business, but as it grows to large, steady client base, they may restrict comment and search opinions available for consumers, resulting in most active prosumers disappointment, expressed on producers forums. As an example recent changes in comment system on Allegro, large auction site, were broadly criticized by users and resulted in temporary drop in number of transactions amount. Users who built their credibility through years, as a result of the changes were deprived of history
and now its status is equal to users who join the service year ago [7], what results in decrease of buyers trust in sellers’ rating.

There are indications that capitalism is having a difficult time gaining control over at least some of the prosumers on Web 2.0 This means that, at the minimum, companies will have a different relationship with such prosumers than it has with producers, consumers, or more traditional prosumers [8]

2. Analysis of a current feedback solutions

To reveal how prosumers are acting in current online communities we have browsed shopping sites and social networking sites, extracting knowledge from consumers’ statements. We also compared how biggest successful retailers incorporate prosumers voices and pinions on their sites. Our research shows that major shopping sites allow customers to express their opinion about product, and buyers willingly share their opinion.

Table 1. Comparison of feedback option available on biggest e-shopping sites

<table>
<thead>
<tr>
<th>Shopping site</th>
<th>Text review</th>
<th>Rating</th>
<th>Real photo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aliexpress</td>
<td>Yes</td>
<td>1-5 stars</td>
<td>Yes</td>
</tr>
<tr>
<td>Ebay</td>
<td>Yes</td>
<td>1-5 stars</td>
<td>No</td>
</tr>
<tr>
<td>Amazon</td>
<td>Yes (t/purchase over 50$)</td>
<td>1-5 stars</td>
<td>Yes (also video)</td>
</tr>
<tr>
<td>Target</td>
<td>Yes</td>
<td>1-5 stars (in few categories)</td>
<td>Yes</td>
</tr>
<tr>
<td>Walmart</td>
<td>Yes</td>
<td>1-5 stars</td>
<td>Yes</td>
</tr>
<tr>
<td>Rakuten</td>
<td>Yes</td>
<td>1-5 stars</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Source: own preparation on the basis of retailer portals

3. Rating systems

Prosumers activity manifests itself in participation in estimation of entities such as products, web-sites, companies or other users. They express opinions by the mean of rating (or reputation) systems, which can be independent services or integrated part of shop or auction site. Rating systems provide valuable advice to users, based on the collective experience of other users. These comments can not only convince new users to buy products, they can also provide information which size or color options one should choose. Users who complies with these recommendations will have more satisfying products experience.
Authenticity of users’ opinion is related to cost of participation, where the cost is low (when user only has to click recommend/not recommend links) user doesn’t have to thoroughly think over his opinion, when cost of participation is high (e.g. user has to write review) user opinion will be more valuable.

Anonymity of many online communities has advantages and disadvantages. It’s hard to identify participants and trace their history, they can often change pseudonyms, create multiple accounts, sometimes their options are sponsored by manufactures. Reader doesn’t know if persons whose post he reads are expressing honest opinion about product they have acquired and used for a while, or are working on commission of clothing companies, and theirs duty is to write a good reviews of product (or maybe even to critique competitors). On the other side pseudo-anonymity of the web also allows to express opinions which are controversial, but valuable for others consumers. User only want to know if the person giving recommendation applies the same criteria, has the same expectations from a product, knowing real name is not important. Rating sites enable prosumers to expose information about themselves, which may be essential for other users to decide if particular opinion is valuable. One of these sites is Makeup Alley, a Web site that focuses on user-generated reviews, the unvarnished truth about beauty products online. Makeup Alley has a library of well-cataloged, super-specific criticism that allows women to do their prepurchase research among peers they trust, despite or perhaps because of their pseudonymity [9].

![Figure 1. MakeupAlley product review screen](https://www.makeupalley.com/product/qu.asp?itemid=5977)
4. Social context solution for e-market in fashion industry

The exploding use of social media plays an important role in consumers’ growing desire to utilize fashion choices to express themselves, showing style, self-image and values.

Inspiration for this prosumers driven rating systems’ idea was a combination of several phenomena occurring on the edge between technology, society and economy. Development of mobile, portable and wearable devices, as well as redefining the role of the media, consumption of art and entertainment and many other examples of socio-economic and technological changes had great impact on e-market. Important role play social media such as Facebook, Twitter, YouTube as well as the phenomenon of active inclusion consumer in the process of delivering goods and services to the market.

The contextual-social model in e-commerce is based on several assumptions:
1. Location of goods and services in context of the user, interaction with user, user’s environment and use-cases is commonly accepted, trusted and effective method of communicating and adopting the offer of goods and services to the market.
2. Use of commonly accepted and adapted tools and methods related to social media, such as sharing one’s image, sharing information about products and services as a part of raising oneself prestige and well-being, sharing information about methods and context of use, evaluation of product and services, free exchange of ideas and opinions placed places in the context of e-commerce business model will as a result stimulate spontaneous collaboration of the users and select a new group of producers and suppliers who can benefit from this process.

At the same time we are observing strong tendency to adapt a new generation of semantic network, a solution which promotes standards for describing content on the Internet in a way that allows machines and programs (e.g. agents) to process information in a manner appropriate to their meaning. Semantic web uses specifically designed languages for publishing data: OWL (Ontology Web Language), RDF (Resource Description Framework) and XML (eXtensible Markup Language). The importance of information resources is determined by the so-called ontology. The technology of the semantic network, solving the problem of semantic chaos and acts as an intelligent provider of information fragments selected according to interest of the recipients. It is assumed that web network is extension and development of the traditional role of the media in the construction of consumer identity [4].

The concept of contextual and social model in e-commerce can be showed on the example of fashion industry. Process of purchasing garments is well-known to all consumers. There is a distance created by technology between the product and
the consumer. Making purchase traditional way consumer can try on clothes before buying them. Electronic commerce doesn’t give such opportunity. Consumer has to rely on size charts provided by manufactures. Exact measurements of so-called standard sizes: S,M,L,XL may vary between brands, or between different lines of clothes for the same brand. Even if retailer provides size chart specifically for given product it still can be misleading, not taking into consideration many factors such as elasticity of material. Perception of colour depends of lighting. Important information such as exact fiber composition, which should be placed on labels are often lacking on retailers sites.

Recommendation from prosumers

To choose clothes wisely one could have utilize not only information provided by retailers but also data generated by other prosumers. They can write subjective review, value quality, impression that product makes. They can attach real photo of the product, showing how this garment looks on a particular person, how it matches theirs hair colour and skin tone, how it fits person’s shape. Prosumers can also link a YouTube video, showing clothes in motion, making a illusion of three dimensional view. Clothes can be presented in matching sets, showing how to combine different pieces in a ingenious full outfit. Other consumers will be not only buying the same piece of cloth but copying the whole look.User acquires advice about fashion items, which includes real life photos, short review, links to s sites where he can see the photo of person wearing outfit in context or a particular social event. All these information consist of willingly and explicit delivered feedback.

Traditional recommendation systems rely on implicit, less accurate data, and techniques used by e-commerce sites in order to find a most appropriate items are based on finding associative rules between different user or different items bought by particular user.

Content-base filtering relies on recommending items similar to those user have bought or rated in the past. Similarity between items in the same category is calculated based on their features, which may include exact subcategory, size measures, material, colour, designer, style, occasion.

These features falls into three categories: continuous, boolean and discrete features For all these features we can calculate similarity of two items. If size of domain of particular attribute far exceeds size of remaining one, this attribute outweigh effect of remaining attributes.[10] Simple solution for a problem of scalability and mutability of different attributes is assigning weighs, so total similarity of two items is sum of weighted similarities for individual attributes.

\[
D(x_i, x_j) = \sum_{k=1}^{p} w_k \cdot d_k(x_{ik}, x_{jk}) \quad \sum_{k=1}^{p} w_k = 1, \tag{1}
\]
where \( w_k \) is weight for \( k \)-attribute, \( d_k(x_{ik}, x_{jk}) \) similarity for \( k \)-attribute of object \( x \) and \( y \) [10]. Instead of recalculating weights values for different number of attributes so they always sum up to 1, we can simply divide sum of weighted similarities by sum of weights.

\[
D(x_i, x_j) = \frac{\sum_{k=1}^{p} w_k \cdot d_k(x_{ik}, x_{jk})}{\sum_{k=1}^{p} w_k} \quad (2)
\]

 Recommending clothes that are comparable to clothes bought before may be not very useful, because when client wants to buy the same product (maybe in other colour or size), he will be aware of that fact, and simply enter in search engine desired product features. More challenging is to recommend product the user haven’t bought or try before.

Collaborative filtering relies on finding groups of users who share the same preferences. Algorithm decides which group user belongs to by taking into account his purchases and ratings. Collaborative techniques are efficient in providing valuable recommendation but also have some serious drawbacks. First is cold-start or sparsity problem, when initial data about user’s preferences are too small to locate neighbourhood group. Another difficulty is scalability problem, with more users and items computation time grows. User’s preferences changes, recommend system have to adapt for the changes, so calculation should be repeated. Relying on user’s past purchases may be deceiving. Both recommendation techniques are computation heavy and not always guarantee useful recommendation.

Proposed system used a different approach. It relies on explicit entered data about body type, height, weight, detailed measures. Algorithm searches for items from specified category recommended by users who have similar body type and present clothes for analogous occasions. Consumer has possibility to look at real photograph of desired item wore by a person with comparable body type. He can read a review written by a someone he can relate to.

On a retailer site you can only see a photo of a one person, one body type, hairstyle. Community created image gallery can show how particular cloth looks on a different people. They are not professional perfect models, so user can get an idea how these clothes are going to look on him. Moreover he can see whole outfit, clothes recommended by others as a complementary to searched item. In these approach there is no need for collaborative filtering algorithm, based on previous purchases, which may be misleading about real user’s taste. Consumer will be presented with selection of items, which other prosumers willingly want to recommend. Prosumers whose opinions will be most valuable for a specific person
are those who looks alike, have similar body measurements, colouring., but also those who have similar taste in fashion, or are presenting garments for analogous occasion. Not all these criteria can be easily declared by choosing search options. Users should also have opportunity to tag other users as ‘trusted’, which means that these users opinions are more valuable, and recommended by him clothes should be high in search results.

We can’t rely only on prosumers’ voluntary workforce to collect all the information about available products. To acquire broad range of products for a system we could utilize information provided by manufactures and retailers. This data can be completed by information supplied by actual users. Hybrid approach addresses a cold-start problem, where there is no enough users-collected data to find recommendations for search with a number of specific criteria.

5. Rating system components

Data collected from retailers about a particular fashion item include information defining product type (shirt, jacket, skirt), attributes (e.g. sleeve length) and availability of different colour variants and sizes, if these is woman or/and man cloth. Additional information may include style, information about material (wool, cotton) and its percentage in an item. To assign items to categories we have to define mapping of categories for different retailers. Taxonomy of categories should be hierarchical, contain most obvious and easy to assign to categories. More specific categories, which differs between retailers can be marked as tags (e.g V-neck T-shirts falls into category “T-Shirts”, with a tag “V-Neck”)

The proposed system will enable two basic activities for users: search and rating. Search enables user to find desired clothes by giving theirs measures and preferences. Measures and criteria can be entered as allowed range or as exact values. Search can be conducted in retailer data, and also in prosumers’ entered information, where information provided by user’s ‘trusted users’ with whom one can relate, will be most valuable.
Application have to measure which items from database fulfill the search criteria with most accuracy. Product data will be acquired from different websites, so number of available features may vary. Application can’t search only for exact matches, but have to find clothes most similar, just like in case of content-base filtering, but instead of comparing two items, one item is compared to user’s defined criteria.

![Diagram of clothes' categories taxonomy](image)

**Figure 2.** Example of clothes’ categories taxonomy, own work based on existing e-commerce taxonomies

Important aspect in searching is accuracy of data. User wants to find accessible product. Data provided by retailers sometimes lack the information about accessibility of specific size and colour, so prosumers should have option to easily check ‘not available’. Search could only include items which was entered or reviewed in a specific time-span (e.g. last year).
Rating enables to express opinion about items found in database or add a new item (with link to retailer’s site) and rate it. Rating should be as descriptive as possible, but also provide some discrete values to enabled easy search on user’s inputted information. Additional data could include for example subjective quality of material, fiber composition, adjustment to measurement or size information entered by producer (bigger, smaller). User can also add real photos of the product which could differ from idolized image provided by manufacturer. Information can be supplemented by links to videos showing particular piece of cloth in motion.
Consumers may present the whole outfits, recommending to join jacket with particular trousers or skirt, shoes and accessories. This possibility forces to enable option of adding links to others items in the same system.

Rating should be as users friendly and intuitive as possible. We want to encourage to utilize the system not only tech-savvy users but a broad range of ordinary consumers.

6. Conclusion

Comparison of feedback options available on large e-commerce sites shows that companies want to receive feedback form prosumers and present their opinions to other consumers. However these solutions are limited.

Presented system may be used as standalone site but extracted information could be also incorporated in sellers’ and manufacturers’ portals. Because of its autonomy from sellers and producer would be viewed by consumers as more reliable and trustworthy than opinions submitted on selling portal, where negative opinions may be “moderated” and deleted by company’s marketing.

Prosumers want to actively participate in production, by entering valuable feedback, which agitates producers to create more sought after products. Such partnership results in more tailored offer, meaning less unsold items, lack of necessity of return purchased goods Manufacturing of items tailored to consumers’ taste lessens environmental waste and pollution connected with production. Prosumers readily give feedback, but want to be sure it won't be tampered with, that negative opinions won't be erased. Presented system may be used as standalone site but extracted information could be also incorporated in sellers’ and manufacturers’ portals. Because of its autonomy from sellers and producers system will be viewed by consumers as more reliable and trustworthy than opinions submitted on selling portal, where negative opinions may be “moderated” and deleted by company marketing.

REFERENCES


[6] Zmierzch konsumeryzmu era prosumentów
http://www.wirtualnemedia.pl/artykul/zmierzch-konsumeryzmu-era-prosumentow


Reviewers

Adrian Bilski, KZI SGGW
Agnieszka Parlińska, WNE SGGW
Aleksandra Górecka, WNE SGGW
Andrzej Jakubiec, KI SGGW
Andrzej Łodziński, KEiS SGGW
Andrzej Zembrzuski, KI SGGW
Anna Kaczerowska, UŁ, Łódź
Anna Sołtysik-Piorunkiewicz, UE, Katowice
Artur Wiliński, KZI SGGW
Bartłomiej Kubica, KZI SGGW
Bartosz Świderski, KI SGGW
Bartosz Wachnik, PW, Warszawa
Beata Zdniak, KI SGGW
Dariusz Strzęciwilk, KZI SGGW
Grzegorz Koszela, KEiS SGGW
Gustaw Konopacki, WAT, Warszawa
Hanna Dudek, KEiS SGGW
Jarosław Kurek, KZI SGGW
Joanna Baran, WNE SGGW
Krzysztof Gajowniczek, KI SGGW
Krzysztof Karpio, KI SGGW
Krzysztof Siwek, PW, Warszawa
Krzysztof Zmarzlowski, KI SGGW

Luiza Ochnio, KEiS SGGW
Maciej Janowicz, KI SGGW
Maciej Pankiewicz, KI SGGW
Marcin Hernes, UE, Wrocław
Marcin Owczarczuk, SGH, Warszawa
Marek Karwański, KI SGGW
Marek Kocinki, KZM SGGW
Maria Parlińska, WNE SGGW
Marianna Zduniak, KI SGGW
Mateusz Smoliński, PL, Łódź
Michał Kruk, KZI SGGW
Michał Rudowski, PW, Warszawa
Paweł Hoser, KZI SGGW
Piotr Gołasa, WNE SGGW
Piotr Łukasiewicz, KI SGGW
Piotr Wrzeciono, KI SGGW
Rafik Nafkha, KI SGGW
Ryszard Szupiluk, SGH, Warszawa
Tomasz Rokicki, WNE SGGW
Tomasz Woźniakowski, KI SGGW
Victor Shevchuk, PK, Kraków
Waldemar Karwowski, KI SGGW
Wiesław Szczesny, KI SGGW