SCIENTIFIC PUBLICATIONS OF THE STATE UNIVERSITY OF NOVI PAZAR
SER. A: APPL. MATH. INFORM. AND MECH. vol. 16, 1, 2024, 73–90

Towards Citizen-Centered Smart Services Research Insights from a Project

Ejub Kajan, Emir Ugljanin, Ulfeta Marovac, Aldina Avdić, Adela Ljajić

Dedicated to the memory of Professor Ćemal Dolićanin (1945–2023)

Abstract: This paper presents research on citizen-centered smart services from the perspective of a long-term project. Initially intended for E-Government services in social, economic and biological domains, it is naturally extended to smart city services due to the common goals of citizen-centered services, common data sources, similar technologies, and the like. It is also overlapped with some international cooperation with similar goals. Both, E-Government and Smart Cities, feed with information that originates from their sensors, citizens and IoT, and, in return, provide smart services that aim to improve the quality of life and achieve more efficient management in the different environments in which they operate. The sensing is provided by listening to citizen opinions via social media, like Facebook, X (formerly, Twitter), and the like, and other Web 2.0 methods like crowdsensing and crowdsourcing, etc. The emerging world of IoT allows to sense some environmental phenomena like temperature, air pollutants, traffic congestion, fire, flooding, and to evidence some illegal behavior on the streets, for example. Development and deployment of citizen-centered services are analyzed through the prism of four dimensions, namely (1) modeling and decision-making, (2) sensing and analyzing, (3) willingness and engagement, and (4) openness and transparency. In addition, key technological enablers that drive the full achievement of these dimensions are briefly discussed. To address the challenges several issues are discusses. These include the enormous amount of data with high velocity, value and variety, the complexity of Serbian language and grammar with the absence of linguistic resources, and the limitations of IoT resources, to mention a few. These challenges are discussed by related work and a brief overview of proposed solutions with special emphasis on a framework, design of several specialized social machines, text processing in Serbian language, sentiment analysis, business process management and an excerpt of developed and deployed services. We briefly outline the main contributions of published work,

Manuscript received April 26, 2024; accepted May 22, 2024

Ejub Kajan (ORCID 0000-0002-3753-3535) is with the State University of Novi Pazar, Serbia;
Emir Ugljanin (ORCID 0000-0003-1721-3208) is with the State University of Novi Pazar, Serbia;
Ulfeta Marovac (ORCID 0000-0001-7232-3755) is with the State University of Novi Pazar, Serbia;
Aldina Avdić (ORCID 0000-0003-4312-3839) is with the State University of Novi Pazar, Serbia;
Adela Ljajić (ORCID 0000-0001-7326-059X) is with The Institute for Artificial Intelligence Research and Development of Serbia, Novi Sad, Serbia

https://doi.org/10.46793/SPSUNP2401.073K

including the key and technologies utilized, all supported by several experimental results. We conclude by identifying some unresolved problems and proposing future research directions.

Keywords: open government, smart cities, social machines.

1 Introduction

ICT landscape has changed a lot during the last twenty years. New computing paradigms, like service- and resource-oriented, cloud, edge and social computing with high penetration of IoT into everyday life, bring new opportunities to provide accurate and timely services to citizens. The literature review shows that new topics like open government beyond E-Government (E-Gov) , smart cities (SC) , and data analytic have recently been attracting more research [1].

Here, we adopted E-Gov definition: "*as a system for the management of public services that, based on ICTs, aims to improve the quality of the services provided by the Government to its stakeholders (citizens, companies, employees, other governments, etc.), increase its transparency, make improvements to its operation and achieve more efficient management in the different environments in which it operates*" [26]. Similar goals feature SCs .

In [20], SC is defined from a *usability point-of-view* as a developed urban area that creates sustainable economic development and high quality of life by excelling in multiple key areas: economy, mobility, environment, people, living, and e-government. From a *system point-of-view*, Mitchell defines smart city as a mix of digital communication networks, ubiquitously embedded intelligence, sensors, and software [50]. In the context of this paper SC sensors are both citizens and IoT . While citizens provide and consume information, IoT devices act according to their duties, i.e. sense, actuate and communicate.

The aforementioned goals are not easy to achieve. In the context of this report, we emphasize a few. In most cases, content, context and coverage of sensed data from social media are not known in advance and require additional processing. To be worse, in the case of languages, e.g. Serbian and the like, which are characterized by complex grammar and the modest set of available lexical resources [45], the threshold of machine readability of data taken from social media significantly decreased.

The rest of the paper is organized as follows. Section 2 presents dimensions of E-Gov and SC through which design, development and deployment of their services take place. Section 3 introduces some key terms along with their definitions and the related work. The evolution of the framework and research results are described in section 4. Concluding remarks and future research directions are given in 5.

2 Dimensions of E-government and Smart cities

E-Gov and SC could be seen from four dimensions as shown in Fig. 1, that are in brief:

1. *Modeling and decision making* address the need of E-Gov and SCs to carefully model and refine their operations, on the one hand, and to make decisions according to their sensors, on the other hand. Modeling requires tools that are capable of dealing

with spontaneity, freedom of speech, and privacy of social media users. The lack of maturity of existing software engineering tools, and business process (BP) modeling tools are identified and discussed in the literature, e.g. in [11] and [48].

2. *Sensing and analyzing* relate to listening citizens and putting sensed information into appropriate analysis to enhance decision making at the given time and space. This involves dealing with Big Data phenomenon characterized by so-called 4Vs(Volume, Variety, Velocity, and Value) [14], and with Natural Language Processing (NLP) for opinion mining and sentiment analysis [17], etc., as well. The business impact of social media technologies is widely recognized taking into account various aspects, such as e.g., crowdsourcing and Mashup to E-Business interoperability [21], but also in terms of knowledge engineering [30], rapid application development, collaboration, etc [2].

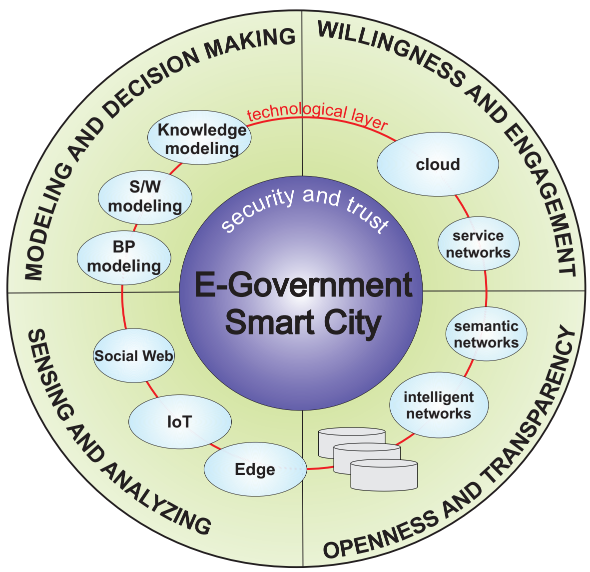

Fig. 1. Common dimensions of E-Government and Smart Cities

3. Citizens participation is the crucial dimension of E-Gov success[15] and could be measured by *willingness and engagement*. The former is discussed by Kollock defining some key factors that would motivate users to community contribution [32]. The later is discussed by Lowndes et al. with CLEAR (Can, Like, Enabled, Asked, Responded) framework for motivating engagement [39].

4. *Openness and transparency* are another E-Gov dimensional features that facilitate its success and appear as essential parts of public governance. Transparency could be defined from two perspectives: information availability and information flow. Whilst

the former points to the kind of information available, the latter describes how stakeholders access various types of government information, for example, inward vs. outward and upward vs. downward. Openness is related to the degree of interoperability between E-Gov BPs and with citizen services, as well [52].

Policies, procedures and technical measures are needed for preventing unauthorized access, data theft, and other incidents in order to establish security of services and foster trust to them.

3 Background

This section gives a concise definition of some key terms and a brief overview of social machine research.

3.1 Key terms and definitions

Business artifact (BA) is a concrete, identifiable, self-describing chunk of information that can be used by a business person to actually run a business [51], e.g., a record in a database or a file, etc.

Business Process (BP) is a set of logically related tasks performed to achieve a particular business outcome [18]. When we think about E-Gov , business outcomes are good services offered to citizens.

Business Process Management (BPM) is " *the art and science of overseeing how work is performed in an organization to ensure consistent outcomes and to take advantage of improvement opportunities*" [22]. BPM is often associated with a 5-stage BP life cycle intended to improve a BP from "*as is*" to "*to be*" that is repeated until a BP becomes mature.

Data, Information, Knowledge and Wisdom (DIKW) is a hierarchy used to contextualize data, information, knowledge, and sometimes wisdom, in respect to each other. It is also used to identify and describe the processes involved in the transformation of an entity at a lower level in the hierarchy (e.g. data) to an entity at a higher level in the hierarchy (e.g. information). [53].

Electronic Health Record (EHR) is a patient report that stores a patient's private information (usually structured) and the physician's notes, diagnoses, laboratory reports, therapies, etc. (usually unstructured and written in free informal text) [6].

Natural language processing (NLP) is an area of research and application that explores how computers can be used to understand and manipulate natural language in the text or speech form to perform useful tasks [16]. Examples of research fields are artificial intelligence, information retrieval, Web mining, social network analysis and similar disciplines, that process information given in natural language by some criteria.

Sentiment analysis is an ongoing field of research that *"aims at determining opinions, emotions, and attitudes reported in source materials like documents, short texts, sentences from reviews, blogs, and news, among other sources"* [17]. It can be applied on a variety of textual sources on different granularity levels (an entire document, phrases, separate words [23] and may include several steps, like tokenization, POS tagging and lemmatization, that are usual in any NLP task, but also may include word disambiguation [19].

Social artifact (SA) is a meaningful piece of information that a Web 2.0 application makes available to users and other applications [40]. Examples include a post on social media, comment on a post, or tagging something, just to mention a few. Each SA consists of social action, e.g. a *post* or a *comment*, and data part included in a social action, e.g. what is posted.

Thing artifact (TA) is a chunk of information that refers to its functionality, captures its lifecycle, and tracks its interactions [41]. Each TA also has action and data part. Actions reflect *thing duties*, i.e. sensing, actuating and communicating, whilst the data part brings the information what is sensed or exchanged.

3.2 Related work

The vision of Social Machines (SMs) was introduced by Tim Berners-Lee who stated "*Real life is and must be full of all kinds of social constraint - the very processes from which society arises. Computers can help if we use them to create abstract social machines on the Web: processes in which the people do the creative work and the machine does the administration* ..."[8].

Since then, several new trends and visions appeared. Examples include, but not limited to the following. In [24], Hendler and Berners-Lee address the challenge of unifying both, creative work done by people and administration executed by machines, in other words, social processes and computation, respectively. In [54], Shadbolt examines blending computation and social contribution with software and dealing with the challenges of building a new generation of social systems, as well.

SMs also changing the way of software developing thanks to the social networking sites that allow their internal capabilities and services to be exposed to external world in the form of open online API platforms[1]. These open APIs allow third-party developers to interact with social-networking sites, access SAs posted by their users, and create other applications and services.

Buregio et al. present SM as a meet-in-the-middle environment that acts as an integrator through specialized APIs. These APIs encapsulate both, the BPs tasks and the SAs actions, allowing the BPs to act over the SAs and vice versa [12]. In line with that, in [10], a semantic SM is proposed as "*knowledge-as-a-service*" provider on the fly that would be able to foster agility and competitiveness of services.

[1] e.g. Twitter API, and Facebook API

Last, but no least, Ugljanin et al. develop a SM that fosters citizens' willingness and engagement in smart city initiatives by drilling into SAs and TAs produced by smart city sensors and connecting them into rule-based chains [57].

An effort introduces Social Miner as a special kind of SMs intended for tracking user actions on a business initiative over social media [29]. In this project, instead of the data that social media bring, the focus of SAs processing is given on their kind, like the post, comment, share, etc., allowing analysis of the interactions that arise in the social world in response to both events triggered and actions taken in the business world and then, to what extent these interactions would impact BAs.

In a recent study [9], Brito et al. identify several challenges in front of research on SMs. These include the lack of empirical data to assess the challenges and benefits of SMs and the lack of their common definition and classification. The same study also emphasizes the need to develop new generation of SMs that will be able to capture the synergy of Web of people, IoT and AI and new software architectures and technologies capable of integrating them in an efficient, scalable, and secure way.

Today, research trends show that everything on the Internet is going to be socialized, if not yet. Examples include socialization of things [3], multi-type artifacts [41], IoT objects [13], just to mention a few. These trends seek for a new generation of SMs based on social relations between Internet objects of the same kind and cooperation and competition between objects of different kinds on the Internet.

4 Project overview

This section provides an overview of the framework development throughout the project duration, highlights research issues along with some published results, and demonstrates key outcomes using appropriate metrics.

4.1 Framework development

The initial architectural concept of the project is shown in Fig. 2. It is based on the specialized SM Advanced Answering Machine (ADVANSE), whose initial functionality is presented in [55]. In essence, it is rule-based reasoning engine that processes a pool of frequently asked questions $\mathfrak{faq} \ni (q_1, q_2, \cdots q_m)$ and a pool of frequent answers $\mathfrak{fa} \ni (a_1, a_2, \cdots a_n)$. Using text processing techniques like tokenization, stop words removal, normalization, and similarity measures ADVANSE acts as a matchmaker that discovers similar words between the sets of $\mathfrak{faqs}$ and $\mathfrak{fas}$ on the same topics.

Human actors communicate and use ADVANSE components including subject matter experts and decision makers (sometimes overlapping) on the E-Gov side, and citizens and/or business people subscribed to some government services. Subject matter experts define and publish problems and/or initiatives, model BPs, and audit citizens' feedback via appropriate dashboards, whilst decision makers adjust E-Gov services according to subscribers' feedback, write new policies, etc.

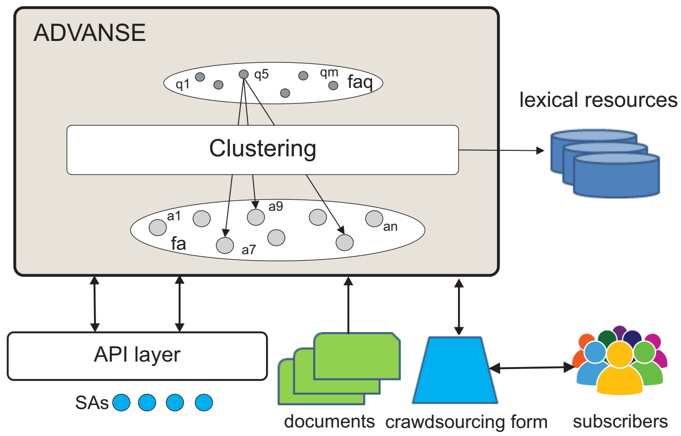

Fig. 2. ADVANSE Framework

ADVANSE is used to develop a service for investigating the local economic environment, i.e., to measure the ability of state and local government bodies to deal with specific problems in creating a stable and favorable business environment, at the city of Niš [28]. Questions are given in the form of a crowdsourcing pool, while answers are obtained from entrepreneurs. Results of service deployment may direct decision-makers to foster education for some professions, change taxes to offload small and medium enterprises, etc.

As a result of efforts made during E-Gov development, especially by building lexical resources in Serbian [45] and through the development and validation of a variety of normalization algorithms [31], [46], [47], a solid base for deeper SA s processing has built.

Thus, a specialized SM intended for sentiment analysis has developed by Ljajić et al. [36], [38], [35]. More specific, this SM focuses on sentiments that involve negations. Two methods are used for sentiment classification: one is based on sentiment lexicon, the other employs machine learning techniques.

Whilst the aforementioned efforts are centered on the content of SAs and/or documents, the focus of this project shifts towards IoT and SC services. Meanwhile, strong international cooperation has been established, a brief overview is given in [27]. As a result of this project and international cooperation, a common framework for building smart city services is defined, as shown in Fig. 3.

The framework consists of one or more specialized SMs intended for SAs and TAs processing to feed a SC service or a bundle of service. These services are consumed by users and/or as a part of a BP services that run in a smart city, e.g. traffic management, health services, special events management, disaster management, etc. Services are usually offered as cloud services. Cloud offers more resources, more reliability and more security, but suffers by high latency. As such, time-sensitive services are often move from cloud to edge, regardless restrictions of edge technologies, like limited resources and poor security. SM should allow DIKW hierarchy establishment due that artifacts fetched from SC sensors are very raw, their content, context and format are not known in advance and should be

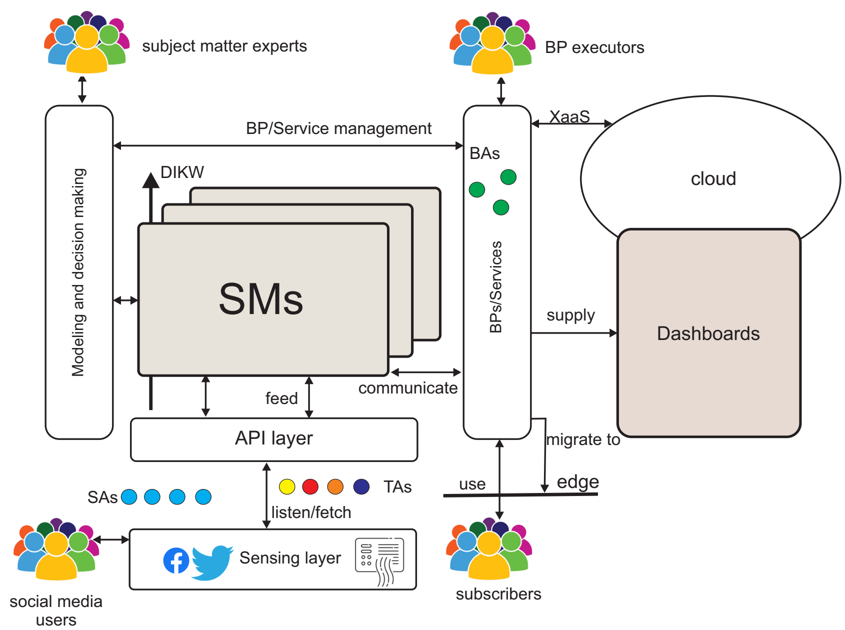

Fig. 3. Common Framework

derived into useful knowledge, as shown in 4.

Avdić et al. [5] use data emitted from crowdsensing and air pollution sensors as inputs to a specialized SM for developing context-aware smart healthcare platform. It serves as a unique middleware that, on the one hand, separates sensor heterogeneity and applications heterogeneity (actually, their BPs), and, on the other hand, establishes a connection between SAs and BAs. Wile the former artifacts feed healthcare platform services with subscribers reactions on air-pollution incidents, BAs carry medical records of patients that have chronic breath disease, like asthma, allergy, etc. On return service feedback appears in the form of warnings, advises, and in some cases with medical assistance call.

The most completed solution that follows key features of the general framework shown in Fig. 3 is B2S4B) [58]. The acronym stands for *Business-to-Social-for-Business*, which reflects the purpose of the framework. In essence, it calls for citizen participation in particular SC initiative and meshes the citizen sensing with related IoT sensing. B2S4B is based on so-called context aware Smart City Observers (SCO)s made by digging into SAs and TAs. These are further organized in the chain according to some rules behind SC BP. SCO chain represents the current stage of a BP lifecycle whose processing allows to answer 5 interrogatives (5w), as shown in Fig. 4, set up by Zachman in the late 70s [61]. Answers to 5w drive reconfiguration of IoT and BP enhancement from *"as-is"* stage to *"to be"* stage.

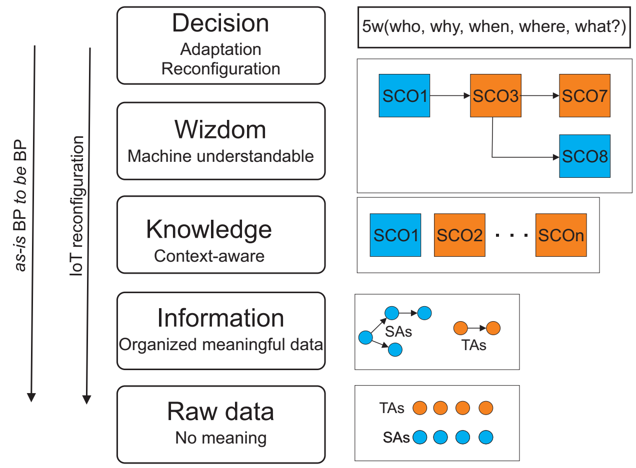

Fig. 4. DIKW (extended with decision) establishment in B2S4B

4.2 Discussion and brief overview of results

During the project, many research issues have been identified and may be classified by *obstacles* to NLP of Serbian and similar languages and *complexity* of the environment where SC services should be deployed. The former include, but no limited to, complex grammar, two alphabets in use, a limited number of available lexical resources, computer, medical, and other jargon in use, etc. The later include the nature of SAs and TAs and IoT, etc.

The complexity of Serbian grammar includes seven cases, multiple verb tenses, palatalization, the ambiguity of words and sometimes even entire sentences, negations, etc. In general, NLP requires several steps to make it machine-readable, to find similar meanings in different text corpora, etc. These steps include normalization and text similarity measures according to some rules. Normalization is a multi-step process consisting of tokenization, stop word removal, and reducing words to their basic form. In the case of Serbian text, it may also be required to settle content of relevant SAs into one of two alphabets in use.

One of the main obstacles in processing documents, SAs, and other text in Serbian is the lack of adequate lexical resources. As such, the project's objectives also included creating some essential resources, like *stop words* dictionary, *sentiment lexicon*, dictionary of *medical terms*, and *negation signals*, as well. Stop words dictionary for the Serbian language contains 1,241 different stop words [44]. It was developed using the Serbian language's grammar as well as comparisons with the collections of stop words which are available for Serbian and related languages. This dictionary's application was demonstrated on three different datasets (tweets, medical reports, and movie reviews). It was demonstrated that stop word removal using a created dictionary improves the performance of the *N-gram* language model.

Research [45, 44] indicates that specific domains also require the creation of domain dictionaries. Therefore, for the problem of sentiment analysis or the problem of processing medical reports, there was a need to adapt existing resources and create additional specific ones. The sentiment dictionary for the Serbian language was developed based on Bing Liu's Opinion Lexicon [25] and evaluated using a dataset of tweets concerning public personalities. These tweets were manually labeled with three sentiment classes [35].

Negation processing also appears as a side problem of sentiment analysis, so resources were created for it. Handling negation is crucial for developing more effective systems, particularly those containing a large portion of negation - like medical reports where negation is commonly used to exclude potential medical diagnoses. Since sentiment analysis is widely used to determine text polarity, handling negation is crucial for this task, as negation can reverse the predicted polarity if not properly addressed.

Health services are on the agenda of any smart city project. Everyone can easily imagine a situation when a person collapses in a public place. In case of serious health problems like heart attack, choking, etc., it is of crucial importance to deliver EHR of a person under attack in machine-readable form, to all services that will be activated in order to save her life. But, EHR processing is not an easy task due to the free-form of a record full of medical, Latins and other terms.

The set of diagnoses specified in the worldwide classification ICD-10 was the basis for the creation of particular lexical resources for medical domains for this study. Three word dictionaries were produced: one for ICD-10 diagnosis codes, another for terms found in Serbian diagnosis names, and a third dictionary for terms found in Latin diagnosis names [4, 43].

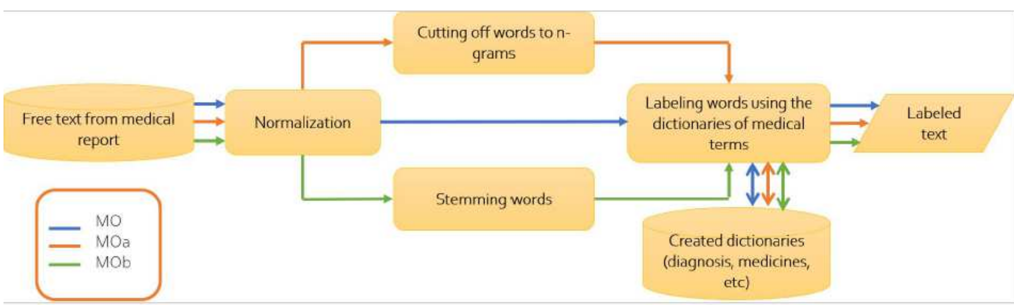

Fig. 5. Processing steps by method M2

Several methods have been developed for labeling medical terms in an informal text. Method M0 is lexicon-based and uses both, medical and non-medical lexicons. The other methods use machine learning (ML), in particular, M1 is based on supervised ML, while M2 uses supervised machine learning and rules [6]. Necessary actions for processing such informal text follow the similar steps as described earlier for document and SAs processing are shown in Fig. 5.

Complexity of the environment where SC services should be deployed is multifold. In addition to previously emphasized attributes of SAs and TAs , unpredictable events, like aforementioned health case, traffic accident with seriously injured people, air-pollution accident, fire, etc. could be happen at any time and any place in a city. SCOs proposed in B2S4B are lightweight and could be encapsulated fast through DIKW hierarchy on the nearest edge node of the accident, and bridge the latency problem, as well. On the other hand IoT devices suffer by limitations, like the diversity and multiplicity of things' development and communication technologies and limited IoT -platform interoperability, and by things restrictions like weak collaboration, poor cognitive capabilities of the current generation of IoT devices, and the lack of proper modeling techniques as identified in [42]. The concept of chained SCOs is a step forward for better cooperation between things.

Table 1. A brief overview of results

Subject	Published	Key tech. and tools	Main contribution
Framework	Simić et al. (2014) [55]	BasicWeb, PHP, UML	Clustering, text similarity measure
	Avdić et al. (2019) [5]	MPAndroidChart, Basic Web, PHP, MYSQL, JSON, Android API 14+	Architectural Framework and Implementation of Air Pollution Dashboard
	Ugljanin et.al (2022) [58]	BPMN2.0, Docker, JSON, JS, MQTT, MongoDB, Node, PHP, Python, REST-API, cloud, edge	Establishment of DIKW hierarchy by handling SCO chains and answering 5w in favor of SC BPs adaptation and refinement during unpredictable events.
NLP & Sentiment analysis	Marovac et al. (2012) [46]	PHP	N-gram based extraction of key terms (n-grams) from documents in Serbian language.
	Ljajić et al. (2018) [38]	Weka, Python	Improving sentiment analysis by handling negation rules.
	Avdić et al. (2020) [6]	Weka, Python	Methods for classification terms in medical texts in Serbian, corpuses created.
	Avdić et al. (2020) [7]	Weka, Python	Methods for normalization terms in medical texts in Serbian.
	Marovac et al. (2021) [44]	PHP, Python	Creating a stop word dictionary in Serbian
	Ljajić et al. (2019) [35]	Weka, Python	ML methods for handling negation in SAs.
	Marovac et al. (2021) [45]	Python	Analysis of the impact of normalization on sentiment word mapping
BPM & Services	Ugljanin et al. (2018) [60]	Basic Web, PHP, MySQL, JSON, Facebook Graph API and Webhooks	Bridging the gap between SC BPs and SC sensors.
	Kajan et al. (2015) [28]	crowdsourcing, UML	service for enhancing local economic environment
	Dolićanin et al (2015) [20]	N/A	Overview of the research worldwide on citizen-centered services
	Ugljanin et al. (2018) [59]	HTML, JavaScript, CSS, Ajax, Bootstrap library, PHP, MySQL, Apache, Facebook Graph API, Facebook SDK, REST, JSON	Provide service that analyzes citizens feedback on SC initiative helping decision makers to do foster or abandon that initiative

4.3 Some key experimental results

Text analysis is challenging because natural languages have many rules and exceptions. Reducing terms to their most basic form requires so-called *normalization*. We study three techniques — *n-gram* analysis, *stemming*, and *lematization* — for reducing Serbian words into their most basic form during multiple research projects. N-gram candidates (q) for keywords are selected by their frequency of occurrence in the text. The most significant n-grams (w) are calculated using the chi-square test (1), where n_w is the total number of words in sentences in which the n-gram(w) appears, p_q is the percentage of the total number of words in sentences in which the key n-gram q appears in the entire document, and $freq(w,q)$ is the number of occurrences of these terms together in sentences [46].

$$\chi^2(w) = \sum_{g \in G} \frac{(freq(w,g) - n_w * p_q)^2}{n_w * p_q} \tag{1}$$

Three normalization methods compared for their effects on the sentiment analysis of the texts [37]: cutting words into n-grams of length 4,5,6, and 7 ($CN4, CN5, CN6, CN7$), stemming using Milosevic stemmer [49] (SN), and lematization using the morphological lexicon of Krstev et al. [33] (LN). We report the outcomes of applying these methods while accounting for the sentiment algorithm's classification time complexity and accuracy. The obtained results show that cutting into n-grams is far less complex, the ratio of the complexity of different methods is: $LN : SN : CN7 : CN4 = 3636245 : 5575 : 4804 : 2272$. The accuracy of sentiment analysis using the machine learning method is the best when using stemmers (Table 2).

Table 2. Correctly classified tweets using machine learning method depending on normalization [37]

Method	**SN**	**LN**	**CN4**	**CN5**	**CN6**	**CN7**
Accuracy	85.27 %	84.25%	83.96%	84.45%	84.47%	84.71%

In [6], the metrics to compare the results of M0, M1, and M2 methods for processing EHR records are presented and include $Precision$, $Recall$, $F1-Score$, and $Accuracy$. For medical terms only they are defined by equations (2-5).

$$Precision = \frac{N_{CLMT}}{N_{LMT}} \tag{2}$$

$$Recall = \frac{N_{CLMT}}{N_{MTL}} \tag{3}$$

$$F1-Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

$$Accuracy = \frac{N_{CLMT}}{N_T} \tag{5}$$

Here, N_{CLMT} is the number of correctly labeled medical terms, N_{LMT} is the number of terms labeled as medical, N_{MTL} is the number of medical terms to be labeled, and N_T is the total number of terms. The best $F1-Score$ and $Accuracy$ were obtained by the M2 method, thus here we emphasize cumulative results obtained by the M2 method on corpora with about 2000 labeled reports, compared with frequently used publicly available tools Tree Tagger and Random Forest Classifier, as shown in Table 3.

Table 3. Comparison of methods for different term categories [6]

Term category	**M2 Method**	**Tree Tagger**	**Random Forest Classifier**
Symptom	0.963895	0,940355	0.933064
Biochemical analysis	0.950119	0,333333	0.625698
Diagnosis/disease	0.934211	0,918103	0.901099
Drugs/therapy	0.791155	0,566327	0.771218
Anatomical organ	0.912621	0,856525	0.937468
Symbol of negation	1	0,99115	0.901763
Description of symptom	0.858908	0,857438	0.935256
Specialty	0.930788	0,904357	0.886364
Term with a negative meaning	0.990476	0,898305	0.978355
Term that emphasizes meaning	1	0,891566	0.991781
Latin word	0.737589	0,268293	0.453258
Stop words	0.975724	0,9977	0.677451

Validation of the proposed solutions for handling negations in in sentiments is done using corpora of 9059 SAs in the form of tweets [38]. Both, sentiment lexicon (SL)-based methods and machine-learning methods, are justified using the determination of $Polarity(SA)$, absolute and relative change in accuracy i.e. $Acuracy_{AbsCh}$ and $Acuracy_{RelCh}$, respectively, as per expressions 6 - 8, where CM is a method to be compared, and B is baseline method based on $Polarity(SA)$ and SL. Results of these metrics is shown in Table 4, Where LBM0, LBM1, and LBM2 relates to equations 6 - 8, respectively

$$Polarity(SA) = \begin{cases} positive & if \quad sumPos > sumNeg \\ neutral & if \quad sumPos = sumNeg \\ negative & if \quad sumPos < sumNeg \end{cases} \tag{6}$$

$$Acuracy_{AbsCh} = CM - B \tag{7}$$

$$Acuracy_{RelCh} = \frac{CM - B}{B} \times 100 \tag{8}$$

Table 4. Applying sentiment analysis on different SAs by lexicon-based methods [38]

		LBM0	**LBM1**	**LBM2**
ALL	Accuracy	48.57%	51.07%	53.73%
	Improvement	N/A	2.50%	5.16%
	Rel Improvement	N/A	5.15%	10.62
OnlyNeg	Accuracy	39.66%	44.76%	50.23%
	Improvement	N/A	5.09%	10.56%
	Rel. Improvement	N/A	12.86%	26.63%
OnlyRuleNeg	Accuracy	38.86%	46.89%	50.97%
	Improvement	N/A	8.03%	12.11%
	Rel. Improvement	N/A	20.66%	31.16%

To assess the efficacy of the B2S4B framework, various performance analysis techniques are deployed to evaluate the effectiveness and scalability [56]. Here we describe an SCO instance, that monitors two sources within a 6-hour time-frame. It is implemented in NodeJS and executed on an isolated Docker container environment. To mitigate external influences on the SCO' performance, Docker is deployed on a low-end Linux Virtual Private Server (VPS). Monitoring encompasses CPU, disk, and network utilization, revealing promising outcomes and low resource utilization, e.g. CPU time used is less than 4.5%.

5 Conclusion

This paper gives an overview of research for developing E-Gov and SC services using dedicated SMs . Many challenging research issues have been identified, like complex grammar, the lack of lexical resources, the lack of proper modeling techniques, and the characteristics of complex environments, where services should be deployed. An excerpt from experimental results validates development methods and algorithms applied in NLP for processing SAs , as well as to handle unexpected events in a SC via IoT. Some of the solutions are applicable also in other domains.

Future research directions will be organized around the socialization of multiple dedicated SMs , that would foster their cooperation in order to coordinate SC services from different ecosystems, e.g. smart health and smart transportation, event-driven on the fly. It will also require developing the semantic representation of the knowledge obtained by processing SAs and TAs , and to develop methods and algorithms to build some intelligence from derived semantics.

Acknowledgments

This paper is written in memory of the late Prof. Ćemal Dolićanin who led the III44007 project over decade.

This paper is partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia under Grant III44007.

It is a great pleasure for the first author to emphasize that three PhD degrees were achieved during this project [4, 34, 56], under the supervision of Prof. Dragan Janković, Prof. Milena Stanković, Prof. Suzana Stojković, and Prof. Dragan Stojanović, respectively, at the University of Nis.

We also would like to thank our international colleagues who contributed to the results of this project with related research we have done together in the last decade, especially Prof. Zakaria Maamar (University of Doha for Science and Technology, Qatar), Dr. Noura Faci (University of Lyon1, France), Dr. Vanilson Buregio (University of Doha for Science and Technology, Qatar), and Dr. Mohamed Sellami (Telecomsud, France).

References

[1] E. ABU-SHANAB, Y. HARB, *E-government research insights: Text mining analysis*, Electronic Commerce Research and Applications, 38 (2019), p. 100892.

[2] S. J. ANDRIOLE, *Business impact of web 2.0 technologies*, Commun. ACM, 53 (2010), pp. 67–79.

[3] L. ATZORI, A. IERA, G. MORABITO, M. NITTI, *The Social Internet of Things (SIoT) - When Social Networks Meet the Internet of Things: Concept, Architecture and Network Characterization*, Computer Networks, 56 (2012).

[4] A. AVDIĆ, *Realizacija servisa pametnog zdravstva i njihova integracija u koncept pametnih gradova*, Doktorska disertacija, (2021).

[5] A. AVDIĆ, E. KAJAN, D. JANKOVIĆ, D. AVDIĆ, *Towards context-aware smart healthcare platform*, International Journal of Electrical Engineering and Computing, 3 (2019), pp. 26–31.

[6] A. AVDIĆ, U. MAROVAC, D. JANKOVIĆ, *Automated labeling of terms in medical reports in serbian*, Turkish Journal of Electrical Engineering and Computer Sciences, 28 (2020), pp. 3285–3303.

[7] A. R. AVDIĆ, U. A. MAROVAC, D. S. JANKOVIĆ, *Normalization of health records in the serbian language with the aim of smart health services realization*, Facta Universitatis, Series: Mathematics and Informatics, (2020), pp. 825–841.

[8] T. BERNERS-LEE, M. FISCHETTI, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*, Harper San Francisco, 1st ed., 1999.

[9] K. BRITO, A. ALVES, S. E. FERREIRA, V. BURÉGIO, G. . CARDOSO, S. ROMERO, *Evolution of the web of social machines: A systematic review and research challenges*, IEEE Transactions on Computational Social Systems, 7 (2020), pp. 373–388.

[10] V. BURÉGIO, N. FACI, E. KAJAN, Z. MAAMAR, M. SELLAMI, *Bringing semantics to the social web*, in 4th International Conference: Contemporary Problems of Mathematics, Mechanics and Informatics, Novi Pazar, 2016, pp. 27–28.

[11] V. A. BURÉGIO, E. KAJAN, M. SELLAMI, N. FACI, Z. MAAMAR, D. BENSLIMANE, *Revisiting software engineering in the social era*, IJSSOE, 6 (2016), pp. 36–46.

[12] V. A. BURÉGIO, Z. MAAMAR, S. L. MEIRA, *An Architecture and Guiding Framework for the Social Enterprise*, IEEE Internet Computing, 19 (January-February 2015).

[13] L. CAMARGO, J. PAULETTI, A. PERNAS, A. YAMIN, *Viso approach: A socialization proposal for the internet of things objects*, Future Generation Computer Systems, 150 (2024), pp. 326–340.

[14] M. CHEN, S. MAO, Y. LIU, *Big data: A survey*, Mobile networks and applications, 19 (2014), pp. 171–209.

[15] H. CHOURABI, T. NAM, S. WALKER, J. GIL-GARCIA, S. MELLOULI, K. NAHON, T. PARDO, H. SCHOLL, *Understanding Smart Cities: An Integrative Framework*, in Proceedings of the 45th Hawaii International Conference on System Sciences (HICSS'2012), Jan 2012, pp. 2289–2297.

[16] G. G. CHOWDHURY, *Natural language processing*, Annual review of information science and technology, 37 (2003), pp. 51–89.

[17] N. F. F. da Silva, E. Hruschka, E. R. Hruschka, *Tweet sentiment analysis with classifier ensembles*, Decision Support Systems, 66 (2014), pp. 170 – 179.

[18] T. H. Davenport J. E. Short, *The New Industrial Engineering: Information Technology and Business Process Redesign*, Sloan Management Review, 31 (1990), pp. 19–41.

[19] C. Diamantini, A. Mircoli, D. Potena, E. Storti, *Social information discovery enhanced by sentiment analysis techniques*, Future Generation Computer Systems, (2018).

[20] Ć. Dolićanin, E. Kajan, D. Randjelovic, B. Stojanovic, *Handbook of research on democratic strategies and citizen-centered E-government services*, Information Science Reference, 2015.

[21] F. Dorloff, E. Kajan, *Balancing of heterogeneity and interoperability in e-business networks: The role of standards and protocols*, IJEBR, 8 (2012), pp. 15–33.

[22] M. Dumas, M. La Rosa, J. Mendling, H. Reijers, *Fundamentals of Business Process Management*, Springer, 978-3-642-33142-8, 2013.

[23] M. Giatsoglou, M. G. Vozalis, K. Diamantaras, A. Vakali, G. Sarigiannidis, K. Chatzisavvas, *Sentiment analysis leveraging emotions and word embeddings*, Expert Systems with Applications, 69 (2017), pp. 214 – 224.

[24] J. Hendler, T. Berners-Lee, *From the semantic web to social machines: A research challenge for ai on the world wide web*, Artificial Intelligence, 174 (2010), pp. 156 – 161.

[25] M. Hu, B. Liu, *Mining and summarizing customer reviews*, in Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, 2004, pp. 168–177.

[26] M. S. Janita, F. J. F. Miranda, *Quality in e-government services: A proposal of dimensions from the perspective of public sector employees*, Telematics and Informatics, 35 (2018), pp. 457 – 469.

[27] E. Kajan, *Software engineering a forty years research retrospective*, Scientific Publications of the State University of Novi Pazar Series A: Applied Mathematics, Informatics and mechanics, 12 (2020), pp. 83–98.

[28] E. Kajan, A. Avdić, U. Marovac, A. Ljajić, G. Šimić, J. Stanković, *Enhancing local economic development using collective intelligence*, in 2015 23rd Telecommunications Forum Telfor (TELFOR), IEEE, 2015, pp. 882–885.

[29] E. Kajan, N. Faci, Z. Maamar, M. Sellami, E. Ugljanin, H. Kheddouci, D. Stojanović, D. Benslimane, *Real-time tracking and mining of social media*, Computer Science and Information Systems, 17 (2020), pp. 1–24.

[30] E. Kajan, Z. Maamar, *Knowledge engineering in the social era*, in 2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), Haikou, Hainan, China, December 20-22, 2021, IEEE, 2021, pp. 2150–2155.

[31] E. Kajan, A. Pljasković, A. Crnišanin, *Normalizacija tekstualnih dokumenata na sprskom jeziku u cilju efikasnijeg pretraživanja u sistemima e-uprave*, Etran, Zlatibor, jun, (2012).

[32] P. Kollock, *The economies of online cooperation: Gifts and public goods in cyberspace*, Communities in cyberspace, 239 (1999).

[33] C. KRSTEV, R. STANKOVIĆ, I. OBRADOVIĆ, D. VITAS, M. UTVIĆ, *Automatic construction of a morphological dictionary of multi-word units*, in Advances in Natural Language Processing: 7th International Conference on NLP, IceTAL 2010, Reykjavik, Iceland, August 16-18, 2010 7, Springer, 2010, pp. 226–237.

[34] A. B. LJAJIĆ, *Obrada negacije u kratkim neformalnim tekstovima u cilju poboljšanja klasifikacije sentimenta*, Doktorska disertacija, (2019).

[35] A. LJAJIC, U. MAROVAC, *Improving sentiment analysis for twitter data by handling negation rules in the serbian language*, Comput. Sci. Inf. Syst., 16 (2019), pp. 289–311.

[36] A. LJAJIĆ, U. MAROVAC, A. AVDIĆ, *Processing of negation in sentiment analysis for the serbian language*, in IcETRAN 2017 Conference proceedingsAt, Serbia, June 2017.

[37] A. LJAJIĆ, U. MAROVAC, M. STANKOVIĆ, *Comparison of the influence of different normalization methods on tweet sentiment analysis in the serbian language*, Facta Universitatis, Series: Mathematics and Informatics, (2019), pp. 683–696.

[38] A. LJAJIĆ, M. STANKOVIĆ, U. MAROVAC, *Detection of negation in the serbian language*, in Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics, WIMS '18, New York, NY, USA, 2018, Association for Computing Machinery.

[39] V. LOWNDES, L. PRATCHETT, G. STOKER, *Diagnosing and remedying the failings of official participation schemes: The clear framework*, Social Policy and Society, 5 (2006), pp. 281–291.

[40] Z. MAAMAR, V. A. BURÉGIO, M. SELLAMI, N. S. ROSA, Z. PENG, Z. SUBIN, N. PRAKASH, D. BENSLIMANE, R. SILVA, *Bridging the gap between the business and social worlds: A data artifact-driven approach*, Trans. Large Scale Data Knowl. Centered Syst., 35 (2017), pp. 27–49.

[41] Z. MAAMAR, E. KAJAN, M. AL-KHAFAJIY, M. DOHAN, A. FAYOUMI, F. YAHYA, *A multi-type artifact framework for cyber-physical, social systems design and development*, Internet Things, 22 (2023), p. 100820.

[42] Z. MAAMAR., E. KAJAN, I. GUIDARA, L. MOCTAR-M'BABA, M.SELLAMI, *Bridging the gap between business processes and iot*, in IDEAS 2020: 24th International Database Engineering & Applications Symposium, Seoul, Republic of Korea, August 12-14, 2020, ACM, 2020, pp. 2:1–2:10.

[43] U. MAROVAC, A. AVDIĆ, D. JANKOVIĆ, S. MAROVAC, *Creating resources for marking diagnoses in electronic health reports in serbian*, (2020).

[44] U. MAROVAC, A. AVDIC, A. LJAJIC, *Creating a stop word dictionary in serbian*, Scientific Publications of the State University of Novi Pazar Series A: Applied Mathematics, Informatics and mechanics, 13 (2021), pp. 17–26.

[45] U. MAROVAC, A. LJAJIC, E. KAJAN, A. AVDIC, *Towards the lexical resources for sentiment-reach informal texts - the serbian language case*, in V CPMMI conference, Novi Pazar, Serbia, 2018, pp. 1–10.

[46] U. MAROVAC, A. PLJASKOVIC, A. CRNISANIN, E. KAJAN, *N-gram analysis of text documents in serbian language*, in In Proceedings of the 20th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2012, pp. 1385–1388.

[47] U. MAROVAC, A. PLJASKOVIĆ, A. LJAJIĆ, E. KAJAN, *N-gram analysis of text documents in serbian language*, in Proceedings of the 20th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2012, pp. 1385–1388.

[48] H. Mili, G. Tremblay, G. B. Jaoude, E. Lefebvre, L. Elabed, G. El-Boussaidi, *Business process modeling languages: Sorting through the alphabet soup*, ACM Comput. Surv., 43 (2010), pp. 4:1–4:56.

[49] N. Milošević, *Stemmer for serbian language*, arXiv preprint arXiv:1209.4471, (2012).

[50] W. J. Mitchell, *Intelligent Cities*. http://www.uoc.edu/uocpapers/5/dt/eng/mitchell.pdf, 2007.

[51] A. Nigam, N. Caswell, *Business Artifacts: An Approach to Operational Specification*, IBM Systems Journal, 42 (2003), pp. 428–445.

[52] M. Pankowska, *National frameworks: survey on standardization of e-government documents and processes for interoperability*, Journal of theoretical and applied electronic commerce research, 3 (2008), pp. 64 – 82.

[53] J. Rowley, *The wisdom hierarchy: representations of the dikw hierarchy*, Journal of information science, 33 (2007), pp. 163–180.

[54] N. Shadbolt, *Knowledge acquisition and the rise of social machines*, International Journal of Human-Computer Studies, 71 (2013), pp. 200 – 205. 25 Years of Knowledge Acquisition.

[55] G. Šimić, Z. Jeremić, E. Kajan, D. Randjelović, A. Presnall, *A framework for delivering e-government support*, Acta Polytechnica Hungarica, 11 (2014), pp. 79–96.

[56] E. Ugljanin, *Platforma za intrakciju sa društvenim mrežama i internetom stvari u pametnim gradovima*, Doktorska disertacija, (2023).

[57] E. Ugljanin, E. Kajan, Z. Maamar, M. Asim, V. Burégio, *Immersing citizens and things into smart cities: a social machine-based and data artifact-driven approach*, Computing, (2020), pp. 1–20.

[58] E. Ugljanin, D. Stojanović, E. Kajan, Z. Maamar, *B2s4b: A platform for smart city business processes management and adaptation*, Studies in Informatics and Control, 31 (2022), pp. 75–86.

[59] E. Ugljanin, D. H. Stojanović, E. Kajan, Z. Maamar, *Initiating and tracking social actions to adapt and improve smart city's business processes*, in 2017 25th Telecommunication Forum (TELFOR), IEEE, 2017, pp. 1–4.

[60] E. Ugljanin, D. Stojanović, E. Kajan, Z. Maamar, *Re-engineering of smart city's business processes based on social networks and internet of things*, Facta Universitatis, Series: Automatic Control and Robotics, 16 (2018), pp. 275–286.

[61] J. A. Zachman, *The information systems management system: A framework for planning*, SIGMIS Database, 9 (1977), pp. 8–13.