HUSO 2021

The Seventh International Conference on Human and Social Analytics

ISBN: 978-1-61208-884-6

July 18 – 22, 2021

Nice, France

HUSO 2021 Editors

Nitin Agarwal, University of Arkansas – Little Rock, USA
Bourret Christian, University Gustave Eiffel /Paris East Marne la Vallée, France

HUSO 2021

Forward

The Seventh International Conference on Human and Social Analytics (HUSO 2021) continued a series of events bridging the concepts and the communities dealing with emotion-driven systems, sentiment analysis, personalized analytics, social human analytics, and social computing.

The recent development of social networks, numerous ad hoc interest-based formed virtual communities, and citizen-driven institutional initiatives raise a series of new challenges in considering human behavior, both on personal and collective contexts.

There is a great possibility to capture particular and general public opinions, allowing individual or collective behavioral predictions. This also raises many challenges, on capturing, interpreting and representing such behavioral aspects. While scientific communities face now new paradigms, such as designing emotion-driven systems, dynamicity of social networks, and integrating personalized data with public knowledge bases, the business world looks for marketing and financial prediction.

We take here the opportunity to warmly thank all the members of the HUSO 2021 technical program committee, as well as all the reviewers. We also kindly thank all the authors who dedicated much of their time and effort to contribute to HUSO 2021. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions. We also thank the members of the HUSO 2021 organizing committee for their help in handling the logistics of this event.

HUSO 2021 Chairs

HUSO 2021 Steering Committee
Christian Bourret, University of Paris East - Marne la Vallée (UPEM), France
Els Lefever, Ghent University, Belgium
Dennis J. Folds, Lowell Scientific Enterprises (LSE), USA
Nitin Agarwal, University of Arkansas at Little Rock, USA
Baden Hughes, Cognitiviti, Australia

HUSO 2021 Industry/Research Advisory Committee
Yi Shan, SDE II, Electronic Arts, Seattle, USA
Fan Yang, eBay Inc., USA
Massimiliano Zanin, The Innaxis Foundation & Research Institute, Madrid, Spain
Zhiwen Fang, Microsoft, USA
Xiaolong Jin, 靳小龙, Chinese Academy of Sciences, China

HUSO 2021 Publicity Chair
Mar Parra, Universitat Politecnica de Valencia, Spain
Alvaro Liebana, Universitat Politecnica de Valencia, Spain

HUSO 2021
Committee

HUSO 2021 Steering Committee
Christian Bourret, University of Paris East - Marne la Vallée (UPEM), France
Els Lefever, Ghent University, Belgium
Dennis J. Folds, Lowell Scientific Enterprises (LSE), USA
Nitin Agarwal, University of Arkansas at Little Rock, USA
Baden Hughes, Cognitiviti, Australia

HUSO 2021 Industry/Research Advisory Committee
Yi Shan, SDE II, Electronic Arts, Seattle, USA
Fan Yang, eBay Inc., USA
Massimiliano Zanin, The Innaxis Foundation & Research Institute, Madrid, Spain
Zhiwen Fang, Microsoft, USA
Xiaolong Jin, 靳小龙, Chinese Academy of Sciences, China

HUSO 2021 Publicity Chair
Mar Parra, Universitat Politecnica de Valencia, Spain
Alvaro Liebana, Universitat Politecnica de Valencia, Spain

HUSO 2021 Technical Program Committee
Paul Abbiati, Founding Fellow of the EUROPEAN LAW INSTITUTE, Austria
Nitin Agarwal, University of Arkansas at Little Rock, USA
Rodrigo Agerri, University of the Basque Country UPV/EHU, Spain
Harry Agius, Brunel University London, UK
Hafizi Muhamad Ali, Yanbu University College, Saudi Arabia
Balbir Barn, Middlesex University, London, UK
Christian Bourret, University of Paris East - Marne la Vallée (UPEM), France
Anita Chandra, IIT Patna, India
Guanghua Chi, Facebook, USA
Dickson Chiu, The University of Hong Kong, Hong Kong
Alexandra I. Cristea, University of Durham, UK
Manuel De Stefano, University of Salerno, Italy
Chen Ding, Ryerson University, Canada
Birgitta Dresp-Langley, Centre National de la Recherche Scientifique, France
Thierry Edoh, RFW-Universität Bonn, Germany
Austin Erickson, University of Central Florida, USA
Silvia Florea, Lucian Blaga University of Sibiu,Romania
Dennis J. Folds, Lowell Scientific Enterprises (LSE), USA
Pierre Fournié, University of Paris Est Marne La Vallée, France
Luca Giraldi, EMOJ srl, Ancona, Italy
Denis Gracanin, Virginia Tech, USA
Shatha Ali A. Hakami, University of Birmingham, UK
Tzung-Pei Hong, National University of Kaohsiung, Taiwan
Baden Hughes, Cognitiviti, Australia

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Innovations in Designing Territorial Platforms for Elderly Homecare Services in France

Bourret Christian
Dicen-IdF Research Team
University of Paris East Marne-la-Vallée (UPEM)
Val d'Europe, France
e-mail: christian.bourret@univ-eiffel.fr

Depeyrot-Ficatier Thérèse
Dicen-IdF Research Team
University of Paris East Marne-la-Vallée (UPEM)
Val d'Europe, France
e-mail: t.depeyrot@ndv-consulting.com

Abstract - **With the evolution of the demographic situation and the increase of the life expectancy, elderly people represent a large portion of the population. Their healthcare has become a real concern. As a result of important changes in the way of life, many of the elderly people live away from their children, so that they have to resort to getting services from local institutions when they become dependent. There are not enough accommodation establishments for the elderly people, and many of them prefer to stay at home. Our work takes into account the widespread wish of the elderly to stay at home longer, event when they become dependent. In this paper, we propose a new model regarding establishments for the elderly called "outside the walls" which may offer a solution, providing an appropriate and full array of services for homecare. The observation of some experiments for this model enables to consider that it may help to provide homecare services on a larger scale and open the way to the necessary transformations.**

Keywords - Territories; Healthcare; Digital Transformation; innovations; France; elderly people.

I. INTRODUCTION

With the demographical transition, the healthcare system has to face the problem of the increasing number of the elderly patients. The current frequency of chronic diseases requires to provide adequate care and monitoring for the elderly. This situation demands a transformation in the healthcare organization in order to cope with the lack of appropriate structures [1].

Until now, the usual model of structure offering accommodation and healthcare for the elderly in France has been the accommodation establishment for dependent elderly, EHPAD. An EHPAD is an establishment where dependent elderly people can get both accommodation and healthcare services; a doctor coordinates a team of professionals with some specialists; the staff is mainly composed of nursing auxiliaries, nursing assistants and nurses. It is managed according to a three-party agreement with the Healthcare Regional Agencies (ARS), the Departments Councils and the establishments.

This model is now often reconsidered because of real problems of quality in the services provided, but also of the inadequacy between the needs and the wishes of the elderly and their families and the life in such establishments. Due in great part to limitations on healthcare expenditures, the present situation risks getting worse in the near future with both the demographical evolution and the increased life expectancy: actually, many diseases and disabilities aggravate with age.

An important effort is now dedicated to the improvement of the quality of services. However, the question remains on the number of establishments that need to be planned for. EHPADs have also been strongly impacted by the health crisis linked to the Covid pandemic, in particular during the first wave of spring 2020, when people over 75 years of age, particularly those living in EHPADs, accounted for a large proportion of deaths.

Meanwhile, many elderly people are expressing their wish to stay at home longer, even when advancing in age and in spite of diseases. "The solution is certainly in a universal offer, under common law, which would allow the elderly population to stay as long as possible at home (…)." [2].

Different models for implementing and managing appropriate services at home have been investigated during the workshops in the consultation process for the report about "Concertation old age and autonomy" managed by D. Libault [3]. The EHPAD and the multi services organization for support and healthcare at home may have the same legal structure, in a model where EHPAD plays the leading role. The stakeholders may also assemble their services through a gerontological web platform without forming a legal structure.

The concept "outside the walls" refers to the extension of the services provided to beneficiaries at home on the territorial area of the EHPAD. It emerged from the idea of providing services from skilled professionals working in those institutions to the elderly at home. Two main models of EHPAD may be distinguished: the traditional one and an innovative "EHPAD outside the walls".

After this introduction (Section 1), we propose to examine the background of the EHPAD "outside the walls" (Section 2). We present our scientific positioning and the methodology used in Section 3. Then, in Section 4, we present the concept of territorial EHPAD and we intend to observe it through some case examples in order to highlight the main results of present experimentations. We aim at

 ISBN: 978-1-61208-884-6

analyzing how they can satisfy the needs from different viewing angles: for the elderly and their families, but also for the healthcare professionals who are involved and finally for the healthcare system. After a discussion around the main challenges of this new model (Section 5), we plan to address the questions of transformation of the healthcare system through such territorial innovation (Section 6).

II. THE BACKGROUND

The combined effects of the post-World War II increase of births and the extension of the life expectancy due to scientific progress led to important changes in the demographic situation. The proportion of elderly people in the total of the population was in France in 2018 nearly at the European level, as shown on the chart from Figure 1.

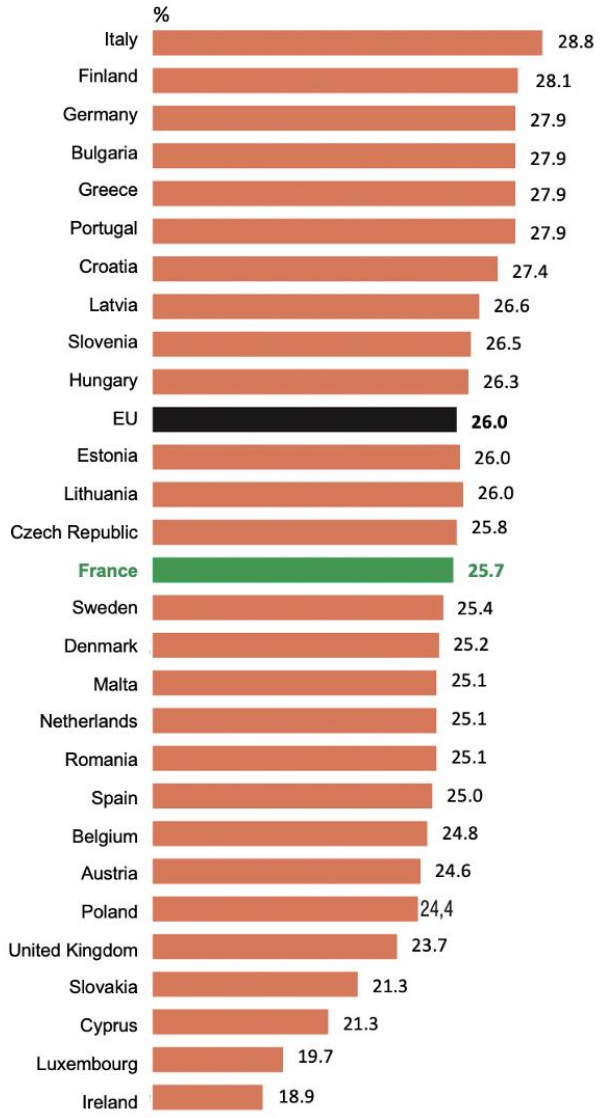


Figure 1. Proportion of people aged 60 or older in the European Union in 2018 - Insee – Extracts in 2019 from Eurostat [7].

The increase in the number of rooms in EHPADs in France was only 2% annually between 2009 and 2018 [7]. With this progression, the total capacity was less than 595 000 rooms in 2018. As for homecare for the elderly, an indicator may be the number authorized for the homecare nursing services, which reached 127 000 in 2018.

Meanwhile, as in other countries, France has to face the issue of the increase in the number of elderly people with loss of independence or autonomy. According to the projections by the French national institute of statistical and economic studies (Insee), 4 million senior citizens would be in loss of autonomy in 2050 [8], as specified in Table I:

TABLE I.
NUMBER PEOPLE AGED 60 OR OLDER IN LOSS OF AUTONOMY
Source – Insee, Omphale projections [8].

Year	2015	2027	2050
Number of elderly people	16 235 900	19 933 500	24 274 500
Number of elderly people in loss of autonomy	2 488 900	2 958 300	3 989 200
At home	1 948 700	2 347 400	3 160 200
In an institution	540 200	610 900	828 900

As a consequence, the care for the elderly people has become a real concern with changes in the way of life: they are often away from their families because of the jobs. They need care as soon as they become dependent.

III. SCIENTIFIC POSITIONNING AND METHODOLOGY

This proposal associates a teacher-researcher in information and communication sciences and an associate researcher, strongly involved in field actions. We place ourselves in an interdisciplinary perspective of action research by also mobilizing management sciences to produce knowledge for action. We are in a position that can be described as "observational participation".

We link the two aspects of information (production and use of data) and communication (cooperation to build new practices together). According to F. Bernard, we focus on four converging aspects: link (interactions), meaning (in situation), production of knowledge, for action [4]. We insist on different notions such as situations, in particular situations of activity or information communication, interactions and ecosystems (systemic approaches in their global context), according to A. Mucchielli approach of “situational and interactionist semiotics” [5], with all the importance of the emotions or feelings and the consideration of human body as a media as for F. Martin-Juchat [6].

Uses of new tools are also essential in a context of global digital transformation: namely, they should enhance data analysis at the territorial level for a better knowledge of the social and healthcare needs; for the follow-up care, they should also help a better coordination between all the stakeholders, whether professionals or relatives. Our approach is also positioned in the main challenge of

 ISBN: 978-1-61208-884-6

improving resilience, not only in the sector of elderly people houses but of the whole healthcare system.

IV. The Concept of Territorial EHPAD

In an opening process and a territorial approach, more EHPADs conduct proactive actions for extending their services outdoors allowing access to other persons than their own residents: it might be for the meals, with sometimes the creation of a specific place for conviviality, a concert performance, or some consultations, for instance when they have organized kinesitherapy, dentistry or ophthalmology on site, etc. Some actions aim at avoiding a negative image and the feelings of a closed space; they contribute to diversify the activities for the residents; other actions are developed for economic purposes, trying to make profitable some services that satisfy the residents' needs. In this orientation, a recent measure consists in the installation of rooms dedicated to temporary stays.

This ongoing process of the territorial opening is part of the current transformation of those structures and the change from a facility-centered logic to a services-oriented logic with the opening on its territorial ecosystem. Another step might be reached with providing services outdoors and build a new patient at home-centered model. Recent call for applications or for projects initiated by some Regional Healthcare Agencies unveil the national orientations for encouraging the EHPADs to their transformation.

In the region Pays de la Loire, the reflection led since 2018 results in the innovative concept of territorial gerontologic services cluster where the EHPADs have to take a central place for providing their expertise. The creation of such clusters requires to define the skills needed and the role of each services provider for organizing the complete homecare support to the elderly people with loss of autonomy. The objective is not merely the coordination between different stakeholders, but above all the creation of an overall basket of elderly' homecare-centered services: the elderly patients' needs and their evolution have to be covered by graduated and quality services, and moreover the continuity and fluidity of their healthcare pathways have to be guaranteed. The reflection points out some missing services within the existing ones provided by homecare stakeholders, due in particular to the hours for their interventions: there is for instance to organize complementary healthcare services for the night.

In a recent call for applications, the ARS Ile-de-France defined the concept of territorial EHPAD in order to support projects for the transformation of the existing structures already settled in their territories and to accompany them aiming to obtain a complete range of services for all the situations of the elderly at home on their geographical proximity. The territorial EHPAD has to position itself both "within its walls" and "outside the walls", in an evolution towards a resource platform as a network in coordination with the hospitals, the ambulatory field, the social stakeholders and the voluntary associations.

Thus, the EHPAD is considered as the appropriate structure for innovation in order to diversify its services delivery and lead experiments based on both its team expertise and new technologies as sensors or telemedicine. In a territorial approach, it has to play a role in developing prevention of the loss of cognitive and physical autonomy or undernourishment. This positioning on the territory requires at the same time to strengthen actions "within the walls" creating new collective spaces and opened events, to develop solutions for providing respite to the elderly at home' families and caregivers, and to organize partnerships with other stakeholders for maintaining the elderly patients at home.

There are different stages in this evolution: first, it may be offering the services provided inside the establishment to the elderly in the territory. Another positioning consists in providing temporary reception when needed for the elderly; finally, services may be organized "out of the doors". Figure 2 formalizes those new trends.

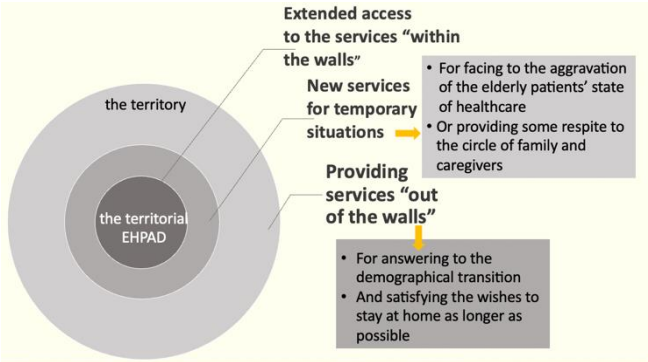


Figure 2. The different axes of the transformation in direction to the territorial EHPAD

V. The Services Provided for Homecare Support

With the increase of the life expectancy, the EHPAD model is challenged from different angles: the very old people are more dependent and have sometimes cognitive or behavioral disorders and several chronical diseases; so, they need more interventions for their healthcare. The problem of recruitment and keeping skilled staff in spite of difficult working conditions is added to this situation.

Moreover, the current demographical transition brings with it a visible social evolution in the worsen isolation for the elderly and their strong wishes to stay at home as long as possible. Building a sufficient number of EHPADs to satisfy the needs does not seem realistic, mainly for financial reasons.

The three different fields of services offered by the EHPAD are: accommodation, vulnerability and healthcare with similar situations to those at home. For the elderly at home, the services have to be provided in answer to their needs in their life pathways, which are both healthcare and dependence.

 ISBN: 978-1-61208-884-6

Among the main breaking points in those pathways, the support for healthcare during the nights and the weekends is identified as causing problems for keeping the patients at home. Security surveillance devices are now more widely used but there are not sufficient. Very often, the patients' conditions demand the intervention of a healthcare provider at their bedsides.

Furthermore, the elderly' s transfers to the emergency department are too frequent. Such situations must be avoided because they present risks of aggravation of the patients 'state of health.

For providing services efficiently, the EHPAD have to work in relation with all the different stakeholders in the patients' healthcare and monitoring, in an intensified partnership to operate on a network with all the different structures that intervene at home in order to complete the services delivery for providing the effective management of the support to the elderly homecare. Figure 3 highlights the positioning of the EHPAD among the homecare stakeholders acting in a patient-centered network:

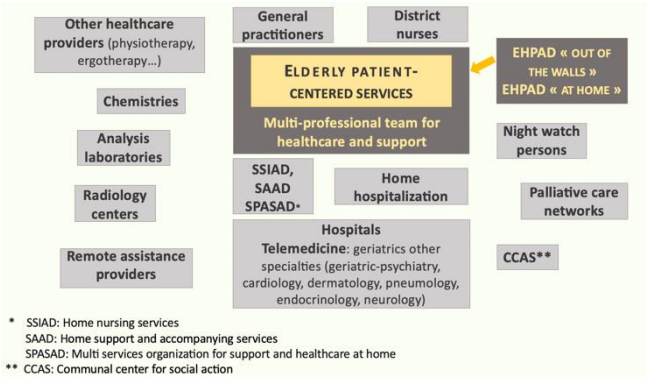


Figure 3. Configuration for assembling graduated services for the elderly' homecare

The objective is to provide a wide range of services for appropriate support at home in the assistance and the graduated patients' care. There is to outsource some services in order to combine them with different home interventions aiming at securing the home support for keeping the patients at home: it consists mainly in strengthening the actions of the existing homecare structures for the patients' healthcare. Thus, the EHPAD is placed at the center for organizing the care of the elderly who wish to stay at home and guaranteeing the fluidity of the patients' pathways.

The EHPAD have to identify how to outsource their services organized indoors in the objective to maintain the elderly patients at home, which skills or expertise are required and which resources assign to this new activity. Moving to the patients' homes takes time besides the interventions. The partnerships have to guarantee a permanent monitoring of the elderly patients in a coordinated response to all the changes. It is important to be coordinated with the hospital in the proximity for managing any aggravation and preparing any potential transfer to the hospital. EHPADs need to re-define their services to ensure they are meeting the needs for elderly patients, as described in Figure 4.

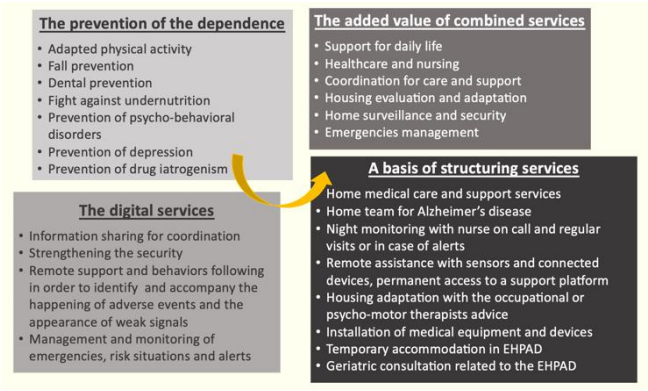


Figure 4. Process of identification of complementary services for the elderly' homecare

Thanks to their expertise and daily practice in geriatrics, the EHPADs may create value in the elderly' homecare, especially for the patients' healthcare continuous monitoring, and also with specific actions aimed at stimulating the patients' autonomy. The option of night home visits of coordinators and nurses from the EHPAD to the elderly' bedside is very important in emergency situations.

VI. The Innovation in the Elderly Homecare

With the demographic transition, it is becoming compulsory at the same time to optimize the organization in order to mutualize the resources and to provide appropriate healthcare to older patients for enabling them to stay at home. Providing additional interventions by specialized professionals is the way to avoid breaches in the elderly' homecare pathways.

A. The integration of services in a territory-scale gerontological pole

The implementation of gerontological poles around the EHPAD at the scale of the territories is a way to fill the gap between the services delivery and the elderly' homecare needs without creating a new structure. Homecare must be considered in a global approach including the medical, psychological and social scopes, which leads to the integration of different services towards a global offer: an integrated social and service network platform [9]. The major breakthroughs are set in the reduction of the compartmentalization within the various structures belonging to different healthcare, medico-social and social fields.

B. Strengthening the older patients' healthcare services

With an increased life expectancy, the elderly' healthcare monitoring is more demanding and requires to

 ISBN: 978-1-61208-884-6

manage some complex cases, due to the interaction of multiple pathologies including several chronic ones. Furthermore, the older patients often have both diseases and disabilities; they may be afflicted by neurodegenerative diseases or psychiatric disorders or become highly dependent. So, the EHPADs have to organize a network in relation with different mobile stakeholders and using telemedicine [10] in order to strengthen the healthcare services provided and develop physical, sensory and psycho-cognitive stimulations. A large range of services has to be provided in relation with other structures for covering all the elderly' needs. Figure 5 illustrates the management of services delivered by different organizations.

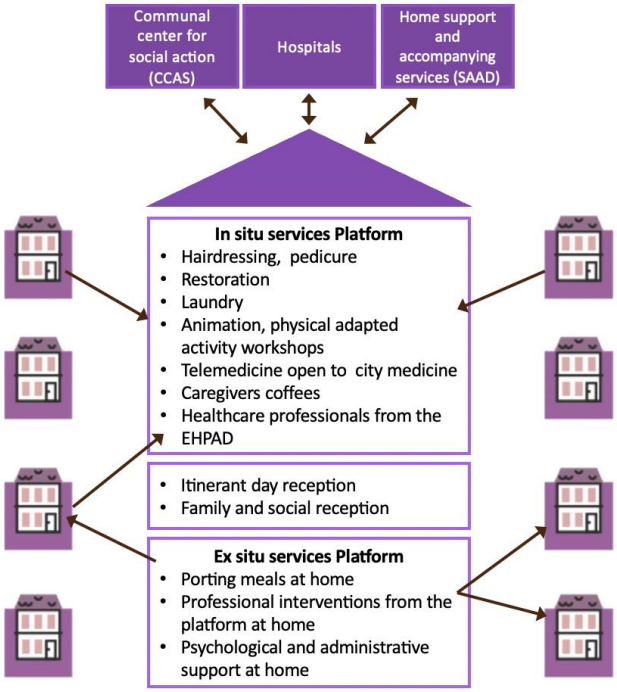


Figure 5. The example of home services coordination platforms [11].

For implementing such an organization, it is necessary to identify the territorial needs and to enhance the coordination between the numerous stakeholders.

C. *Perspectives for* the *human resources management*

For the nursing and support staff in the EHPAD as for homecare, the work is arduous and solutions are sought in order to motivate people to accept and keep those jobs. The possibility for the staff in the EHPAD to diversify their activity and to move "out of the doors" may represent an opportunity for them finding different roles among the homecare teams in a rewarding way and enriching their knowledge. Different formations may be organized. Some tasks like the relation with the other stakeholders and the caregivers may be different for the patients at home. In terms of human resources management, it might bring new perspectives for the evolution of the professional careers.

D. *An innovative approach in the digital processes*

As in the model of Healthcare Networks, the different interventions have to be coordinated and a digital platform may work as an aggregation point for centralizing and articulating them. The aim is to exploit the potential of each local resource and articulate the elderly-centered information. As different shareholders interact, there is not to implement an Information System (IS) designed as for a structure; the creation of value may result from the use of the Information and Communication Technologies (ICT) in an innovative way for improving the information flow and the sharing of the right information with the right professional at the right moment.

As for the organization of a team as the healthcare circle in patients' pathways, all the structures and the different professionals involved in this elderly-centered approach have to be identified. Their actions have to be registered in a structured description. The schedule of the interventions has to be shared in a consolidated patient-centered view.

The aim is to enable an effective transmission of the information, and the access to the relevant data to all the stakeholders according to their roles in the homecare support and their authorizations. Some types of information like incidents and adverse events require traceability through registered reports; other sorts have to be registered as targeted transmissions. An efficient operating mode consists in getting the information automatically pushed to the right professional. Moreover, due to the importance of sharing instantly immediate information, some media like text messages or discussion groups may be essential means.

VII. The Observation of Some Experiments

We tried to list the existing experiments the most similar to the model described rather than to make a complete inventory of the EHPADs that are now opened towards their territory. As the model of healthcare homecare services delivery is recent for the EHPAD, not only there are still few experiments "outside the walls" up to now but also their evaluations are not achieved.

We analyze the main experiments in the chronological order towards the more recent ones in order to point out some potential elements of evolution. We intend to highlight both their main and distinctive features.

M@do (Corrèze) [12]

This project was launched in 2015 by the *Partage & Vie*, a fund that has many facilities and services in the medico-social field both for the elderly and for the disabled people. The project is developed near Tulle in the department of Corrèze for testing the model legally and economically. It aimed mainly at the elderly at home in loss of independence, except those suffering from serious behavioral disorders. Around forty patients are registered in the project. The case manager and the coordinating doctor are both in charge of the coordination. As a structure, M@do employs the professionals and the team is formed

 ISBN: 978-1-61208-884-6

with the staff members. The major objective is to deliver a complete range of services for supporting the elderly patients at home. For securing their monitoring, a night and day permanent assistance service is provided with the intervention of a professional moving to the site if alerts require it.

As well as for the services within the EHPAD, the elderly homecare support services provided are funded by the department council of Corrèze. The main feature in this experiment is the financing of the healthcare services by the Regional Healthcare Agency, which is for an outdoors home patient twice more than the healthcare budget for a patient inside the EHPAD.

The complexity of designing this model is pointed out among the results of this experiment. Moreover, the exceptional derogatory funding method for healthcare at home is the most important factor that may explain why this experiment is not replicated in other EHPADs of the structure or in others, as initially intended. Therefore, the economical dimension of the model is an important issue to deal with: it might represent the major difficulty and requires to be thought in an innovative way.

Ehpad@dom (Yvelines) [13]

The project was set up in 2017 at Sartrouville in the department of Yvelines by an EHPAD in relation with two structures for home services: a Service for aid and accompaniment at home and a Service for nursing healthcare and aid at home. The different structures are managed by the Red Cross. Unlike the project M@do, there is no creation of a specific structure, the organization is based on the EHPAD and the constitution of an integrated mobile team is quite innovative.

The services offer includes the daily aid and homecare, the coordination of all the interventions: those dispensed by liberal practitioners, the home delivery of meals, the aid for administrative procedures, some little layout and renovation work for the housing, a nurse's intervention during the night, a permanent home tele-assistance platform with connected devices for detecting the falls and a special room kept for emergency cases in the EHPAD. The team can follow 24 elderly people, and gathers nurses, nursing assistants, psychomotor, psychologist and maintenance staff.

The main challenge in this experiment is to organize the coordination between the EHPAD and the structures providing services at home. For each new elderly patient, after taking the general practitioner's opinion, the coordinator in the EHPAD goes at home for evaluating the needs and organizing the healthcare support in relation to the families and caregivers: the duration of the different services is defined for all the staff belonging to the EHPAD or the home structures.

The coordination also deals with the medical monitoring, the reception of prescriptions sent by the general practitioner, the order and reception of medicine, etc. Another coordinator in the SSIAD (Home Nursing Services) is in charge of bringing the medicine at home in pill boxes for the week, informing the EHPAD, the general practitioner or the families of possible new needs and potentially contacting the different healthcare stakeholders.

In this way, the experiment deals with the complex issue of the coordination of all the stakeholders for homecare and integration of the different services for their delivery. As for the main factor to highlight, we may point out that the experiment aims at providing services equivalent to those in an EHPAD, although less expensive. This is very important for considering the extension of the experiment with the high budgetary constraints due to the increasing needs.

Diapason 92 (Hauts de Seine) [14]

Another experiment began in 2018 at Asnières in the Hauts-de-Seine Department on the initiative of the Fund called Fondation Aulagnier. It is named Diapason 92. Initially, the project was selected at a call for projects of the ARS in the region Ile-de-France, which gathers Paris and seven departments around, and the departmental council of Hauts-de-Seine for the territory of the EHPAD. The aim is to experiment during five years the coordination of the homecare services for thirty elderly patients requiring to enter the EHPAD.

The unique pathway referent and manager is in relation with all the different professionals involved. For the healthcare, the regular general practitioner remains at the center of the team organization around the patient. Some volunteers are an integral part of the network set up. After the evaluation of the patients' state of healthcare, in order to meet their needs, the personalized healthcare and support plan, the expected stakeholders and their respective work time are defined and a specific team is set up for each patient thanks to the partnerships with homecare structures.

There are three distinct levels of services: the basic one is designed for every elderly person and includes the creation of a specific team, the delivery of nursing and medical homecare by the SSIAD and essential support services (as tele assistance). Another level of services is adjustable and consists mainly in more working time for support, the home meal supply, the support for moving and the itinerant night guard. The optional level has to be paid by the elderly; it includes hairdressing and socio-aesthetic.

Diapason 92 is currently in charge of thirteen elderly patients at home with very different situations. There are many pending files due to the duration for the administrative process. The average out of pocket expenses are 1 254 euros, which is considered quite acceptable.

The general practitioners involved in the project express that they appreciate the organization. It is noted that the activities like the fall prevention as provided in the EHPAD do not arise great interest among the elderly patients.

The organization of the support during the evening is identified at the same time as creating value and requiring improvement. The elderly patients' needs are to be put to bed according to their own habits and wishes. However, it is difficult to make the working hours evolve for the staff in the SSIAD or to find nursing assistants who accept to work at the end of the day. For satisfying such needs, part time contracts are concluded with nursing students for itinerary

 ISBN: 978-1-61208-884-6

guards. Moreover, another observation focuses on the time required for organizing the partnership. For evaluating the experiment, the ARS will consider in particular its effects on the reduction of the length of stay in hospital as returning at home.

VIII. DISCUSSION

The common feature in the experiments presented so far is the global approach in delivering a complete range of support services so that the dependent elderly may stay at home instead of being transferred to an EHPAD. Comparing them, some differences are noticed for the type of elderly patients in the active file. For instance, the projects M@do and Ehpad@dom do not seem to include patients suffering from behavioral disorders; unlike the other projects, Diapason 92 includes very dependent elderly patients (corresponding to the first level in the dependency classification: "GIR 1"). Other results of those experiments and the finalized design of the model will emerge from the work group formed by the National solidarity fund for autonomy with the three structures.

The director of the ARS in Ile-de-France insists on the medico economic relevance of the solutions that have to be set up in a reflection to avoid the compartmentalization. A call for projects was hold in 2019 by the ARS in Provence-Alpes-Côte d'Azur (PACA) Region. Six EHPADs were selected for the experiment which aims at checking how the model "out of the doors" meet the territorial needs with personalized and global support for homecare.

Another experiment is held in Brittany Region by an EHPAD in Rennes in order to face increasing needs that require innovation: the requests for places in the EHPAD reach 400 while there are no more than 87 places. The project, funded by the ARS Bretagne, began at the end of 2019 with twenty elderly patients.

Some homecare services structures begin to experiment a very innovative and complementary model of alternative to the EHPAD from the homecare structure. This model offers very interesting perspectives but is still difficult to implement because of the average dimension of homecare structures and frequent problems of organization for them.

The Think tank *Matières Grises* launched in November 2020 a survey about "the EHPAD for the future" that draws the attention, with a reflection on a new model called "There is no alternative" (TINA), considering the EHPAD as the unique structure for the elderly, but insisting on compulsory evolutions for an acceptable future model. Among the different workshops, one is called "Making possible the EHPAD platform", with both dimensions: "in" and "out".

IX. THE EFFECTS ON THE TRANSFORMATION OF THE HEALTHCARE SYSTEM

In its contribution to the strategy of transformation of the Healthcare System, the High Council on the Future of Health Insurance considers that the priority for the transformation of the system is not to organize the offer on a territory but to facilitate "the access of its population to relevant and quality services concretely defined". The concept "out of walls" is an important step to take towards the transformation for the EHPAD, due to a complete change in its model and in its own perception of its territorial positioning with this new paradigm. Setting up a network for the elderly' homecare on the territory requires a new governance for managing the cooperation between all the stakeholders: would the EHPAD take this role?

In order to organize a shared resources platform opened on its territorial environment, the inventory of the existing structures on the territory has to be done. The different services must be complementary in an efficient way for avoiding any duplication: they have to be specialized for their consolidation in a complete range for meeting the needs of the elderly at home, including a formalized partnership with the hospital. The evolution of the life expectancy induces the requirement of more medicalization for maintaining longer the elderly patients at home; the EHPAD has to specialize, in particular for giving support to the cognitive dependence.

Furthermore, this transformation demands to create an attractive working environment for the gerontology staff, especially for the nursing assistants. The evaluation of the experiments should address the issue of improvement in the coordination of the elderly homecare-centered services and the added value of the EHPAD in the services delivery.

X. CONCLUSION

In the context of the demographical transition, the lack of facilities like the EHPADs in spite of their current expansion and the shortage of nursing assistants or other healthcare providers, the issue of the increasing number of elderly patients must be tackled through innovative ways [15]. The model "outside the walls" may represent the first steps to a change of paradigm for organizing homecare on a large scale. The observation of the experiments points out both the breakthroughs of such new models and the problems remaining to solve. The transformation may only be possible by breaking down the barriers between different sectors in the French Healthcare system, namely medical, medico social and social sectors. Innovation could be supported by new scopes for the coordination between all the different stakeholders at home and the exchanges of data for improving and securing homecare.

According to P. Simon, a paradigm shift seems necessary for a complete transformation to gear the healthcare system from hospital-centrism towards "homespital" [10]. New disruptive healthcare organizations have to be designed and implemented, developing interprofessional cooperation; the medico-social field has to move from the division in very specialized structures to the logic of services delivery for improving the healthcare and life pathways, organizing resource centers for extended homecare.

Moreover, the innovative model of EHPAD "outside the walls" must provide a level of healthcare and monitoring services at home similar to those "within the walls" of an institution, with the transposition at home of the same usual

 ISBN: 978-1-61208-884-6

services delivery as in the EHPAD. It implies the coordination and the quality control of all the services provided at the elderly' bedside [1].

The digital disruption is also essential to support this transformation: beyond the widespread development of the use of devices and sensors for home monitoring, it becomes fundamental to consolidate the distinct services provided by different stakeholders integrating them in a shared digital space. Such digital platforms may contribute to the mutualization of resources. They have to be patients-centered and to manage personalized healthcare pathways with interactive process for sharing information at any time. They must deal with the interoperability issue, which is exacerbated by the lack of common glossaries in the medico-social field.

All these trends and new ways have been greatly accelerated by the health crisis related to the Covid pandemic; the situation emphasizes the importance of a better organization for the follow-up in healthcare as the patients cannot all be treated at the hospital. Furthermore, a part of the elderly does not stay in establishments but at home; the professionals are very often overloaded, which requires to implement all the means that could help them.

This work in progress is a first step for a more long-term work on the transformation of the EHPAD in France with the evolution of the needs and expectations of the users and their families, in a perspective of territorialization of new services and articulation as services platforms for patients who want, for their most part, to be able to continue to live in their own home. For us, it will be a question of integrating not only the needs of users and their families in the whole operational and concrete dimension but also taking into account their feelings and emotions [6].

The notion of crisis is translated into Chinese by two ideograms, meaning both risk and opportunity. For us, this is the definition of the notion of "resilience", which is particularly important in the very serious health crisis we are currently experiencing and which is disrupting our entire health system and, more broadly, our entire society, particularly relations between generations. A new and innovative approach to EHPADs, particularly as "outside the walls" and as a platform of services for home care, is for us an essential way of inventing new models, in a dual context of health crisis and digital transformation with the development of new tools based on new data uses.

J.-P. Le Goff has shown how our "sick society" must face a health crisis of unprecedented magnitude [16]. According to S. Paugam, the main challenge is to rebuild the social link [17] for a society that would be both more supportive and more resilient, and that better integrates its elderly. First of all, in a society that deeply doubts itself and its future, it is necessary to rebuild trust: trust being essential to any social life [18]. The reorganization of the health and solidarity system and therefore of our social protection system (Welfare state) should be a major element in rebuilding society around a project that unites all generations.

REFERENCES

[1] M. Iborra and C. Fiat, Rapport d'information sur les EHPAD / Information report on the EHPAD, Paris, Assemblée nationale, March 2018.

[2] F. Baudier, Crise à l'hôpital et dans les EHPAD, même combat : abattons les murs ! / Crisis at the hospital and the hospice, same combat; break down the walls! Santé Publique /2 (Vol. 31), pp. 293-295 , 2019, https://doi.org/10.3917/spub.192.0293.

[3] D. Libault, Rapport Concertation Grand Age et Autonomie. Available on: https://solidarites-sante.gouv.fr/affaires-sociales/personnes-agees/concertation-grand-age-et-autonomie/article/rapport-de-la-concertation-grand-age-et-autonomie, March 2019.

[4] F. Bernard, "Les SIC une discipline de l'ouverture et du décloisonnement" / "SIC a discipline of openness and decompartmentalization", in Bouzon, A. dir., La communication organisationnelle en débat. Champs, concepts, perspectives / Organizational communication in debate. Fields, concepts, perspectives, Paris, L'Harmattan, 2006, pp. 33-46, 2006.

[5] A. Mucchielli, "Situation et Communication"/ "Situation and Communication", Nice, Less éditions Ovadia, 2010.

[6] F. Martin Juchat, L'aventure du corps. La communication corporelle une voie vers l'émancipation / The adventure of the body. Body communication as a way to emancipation , Presses Universitaires de Grenoble, 2020.

[7] Personnes âgées dépendantes / The dependent elderly TEF, édition 2020 – Insee Références, 2020.

[8] K. Larbi and D. Roy, 4 millions de seniors seraient en perte d'autonomie en 2050 / 4 million of senior citizens would be in loss of autonomy in 2050, Insee Première, July 2019 .

[9] F. Boll and P. Brune, Online Support for the Elderly – Why Service and Social Network Platforms should be Integrated. Elsevier Procedia Computer Science Volume 98, pp. 395-400, 2016, https://doi.org/10.1016/j.procs.2016.09.060.

[10] P. Simon, La Télémédecine, le Télésoin et l'IA Médicale contribuent-ils à transformer au XXIème siècle notre système de Santé ? / Can telemedicine, telecare and medical AI contribute to transform the healthcare system in the XXIth century? http\\www.ManagerSante.com N°24, September 2019.

[11] Fédération Nationale de la Mutualité Française Les enjeux du vieillissement de la population - EHPAD hors les murs et domicile connecté, November 2018.

[12] https://www.fondationpartageetvie.org/jcms/fpv_7992/fr/m-do

[13] https://pourvous.croix-rouge.fr/etablissement-d-hebergement-pour-personnes-agees-dependantes-residence-stephanie

[14] https://agabc.asso.fr/sites/default/files/adessadomicile/page/fichiers_joints/diapason92.pdf

[15] N. Rapegno and H. Bertillot, Transformer l'offre médico-sociale ? Habitats "inclusifs" et établissements "hors les murs": l'émergence d'accompagnements alternatifs pour les personnes âgées et les personnes handicapées. Transforming the socio medical offer ? / "Inclusive" housing and structures "outside the walls": the emergence of alternative support for the elderly and the disabled. [Rapport de recherche] CNSA; EHESP. hal-02074025, April 2018.

[16] J.-P. Le Goff, La société malade / The sick society, Paris, Stock, March 2021.

[17] S. Paugam, Le lien social / The social ink, Paris, PUF, 2010.

[18] C. Senik, dir. Crises de confiance ? / Trust crisis?, Paris, La Découverte, 2020.

 ISBN: 978-1-61208-884-6

An Investigation of When Japanese Twitter Users Deleted Their Tweets Disclosing Their Personal Information

Yasuhiko Watanabe, Leo Mashimo, Toshiki Nakano, Hiromu Nishimura, and Yoshihiro Okada
Ryukoku University
Seta, Otsu, Shiga, Japan
Email: watanabe@rins.ryukoku.ac.jp, t170529@mail.ryukoku.ac.jp, t180450@mail.ryukoku.ac.jp,
t160405@mail.ryukoku.ac.jp, okada@rins.ryukoku.ac.jp

***Abstract*—Nowadays, many people use a Social Networking Service (SNS). Most SNS users are careful in protecting the privacy of personal information: name, age, gender, address, telephone number, birthday, etc. However, some SNS users disclose their personal information that can threaten their privacy and security even if they use unreal name accounts. In this study, we investigated how these users treated their tweets that potentially threatened their privacy and security, in other words, whether they deleted them or not. We collected 233 cases where Twitter users submitted tweets promising to disclose their personal information and they did so honestly. Then, we investigated when they submitted and deleted their tweets disclosing their personal information. The results of our three-month survey showed that 40 out of the 233 cases were deleted and 50% of them were deleted within three weeks after they were submitted.**

Keywords–personal information; Twitter; SNS; privacy risk; unreal name account user.

I. INTRODUCTION

Nowadays, many people use a Social Networking Service (SNS) to communicate with each other and try to enlarge their circle of friends. SNS users are generally concerned about potential privacy risks [1]. To be specific, they are afraid that unwanted audiences will obtain information about them or their families, such as where they live, work, and play. As a result, SNS users are generally careful in disclosing their personal information. They disclose their personal information only when they think the benefits of doing so are greater than the potential privacy risks. However, some SNS users, especially young users, disclose their personal information on their profiles, for example, real full name, gender, hometown and full date of birth, which can potentially be used to identify details of their real life, such as their social security numbers. In order to discuss this phenomenon, many researchers investigated how much and which type of information are disclosed in SNSs, especially on Facebook [2] [3]. Researchers might think that personal information disclosed on Facebook is reliable, or it is possible to check whether personal information disclosed on Facebook is true. This is because

- Facebook users are required to register and disclose their real names when they first start using Facebook.
- Facebook users would be criticized by their friends if they disclose their personal information dishonestly.

On the other hand, a small number of researchers investigated how much and which type of information is disclosed by unreal name account users, such as Twitter users. Researchers might think that personal information disclosed by unreal name account users is unreliable. This is because

- nobody criticizes unreal name account users when they disclose their personal information dishonestly.
- true personal information can threaten their privacy and security even if they use unreal name accounts.

As a result, many of us think that it is natural for unreal name account users not to disclose their personal information honestly. However, Watanabe, Nishimura, Chikuki, Nakajima, and Okada reported that many unreal name Twitter users seemingly disclosed their personal information honestly [4]. It shows that we do not understand well what unreal name account users think about disclosing their personal information. To discuss this problem, it is important to investigate when and how they submit and delete their messages disclosing their personal information.

The rest of this paper is organized as follows: In Section II, we survey the related works. In Section III, we show how to collect tweets where submitters seemingly disclosed their personal information honestly. In Section IV, we analyze when and how these tweets were submitted and deleted, and discuss how long submitters interacted with audiences by disclosing their personal information and when they could not overlook their potential privacy risks. Finally, in Section V, we present our conclusions.

II. RELATED WORK

Personally identifiable information is defined as information which can be used to distinguish or trace an individual's identity such as social security number, biometric records, etc. alone, or when combined with other information that is linkable to a specific individual, such as date and place of birth, mother's maiden name, etc. [5] [6]. Internet users are generally concerned about unwanted audiences obtaining personal information. Fox et al. reported that 86% of Internet users are concerned that unwanted audiences will obtain information about them or their families [1]. Also, Acquisti and Gross reported that students expressed high levels of concern for general privacy issues on Facebook, such as a stranger finding out where they live and the location and schedule of their classes, and a stranger learning their sexual orientation, name of their current partner, and their political affiliations [2]. However, Internet users, especially young users, tend to disclose personal information on their profiles, for example, real full name, gender, hometown and full date of birth, which can potentially be used to identify details of their real life, such as their social security numbers. As a result, many researchers discussed the reasons why young users willingly disclose personal information on their SNS profiles. Dwyer

 ISBN: 978-1-61208-884-6

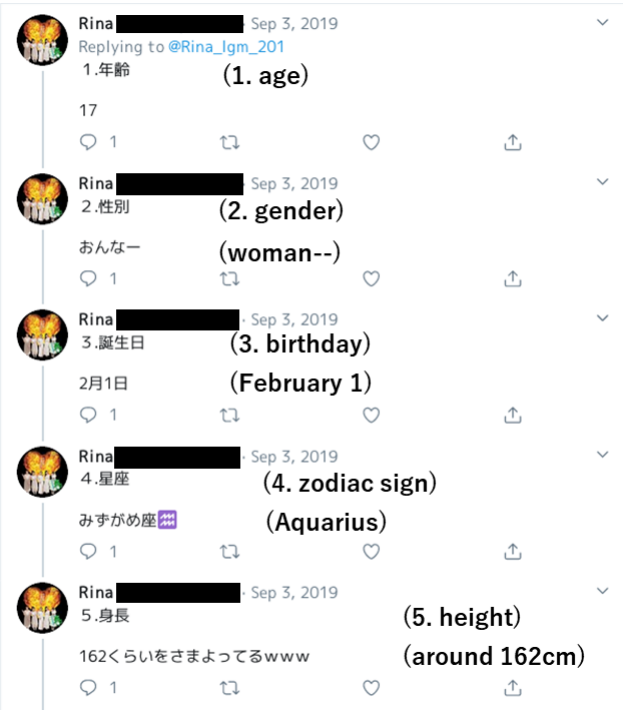


Figure 1. An unreal name account user, *Rina*, disclosed her personal profile items in her tweets.

concluded in her research that privacy is often not expected or undefined in SNSs [7]. Barnes argues that Internet users, especially teenagers, are not aware of the nature of the Internet and SNSs [3]. Hirai reported that many users had troubles in SNSs because they never thought that strangers observed their communication with their friends [8]. Viseu et al. reported that many online users believe the benefits of disclosing personal information in order to use an Internet site is greater than the potential privacy risks [9]. On the other hand, Acquisti and Gross explain this phenomenon as a disconnection between the users' desire to protect their privacy and their actual behavior [2]. Also, Livingstone points out that teenagers' conception of privacy does not match the privacy settings of most SNSs [10]. Joinson et al. reported that trust and perceived privacy had a strong affect on individuals' willingness to disclose personal information to a website [11]. Also, Tufekci found that concern about unwanted audiences had an impact on whether or not students revealed their real names and religious affiliation on MySpace and Facebook [12]. The authors also think that most students are seriously concerned about their privacy and security. However, they often underestimate the risk of their online messages and submit them. For example, Watanabe, Onishi, Nishimura, and Okada reported that many students submit tweets concerning school events and these tweets may give a chance to other people, including unwanted audiences, to distinguish which schools students go to [13]. Watanabe, Nishimura, Chikuki, Nakajima, and Okada also focused on unreal name Twitter users who promised to disclose their personal profile items, analyzed the details of their personal profile items disclosed by themselves, especially their ages, genders, and heights, and showed that most of the submitters disclosed their ages, genders, and heights honestly [4].

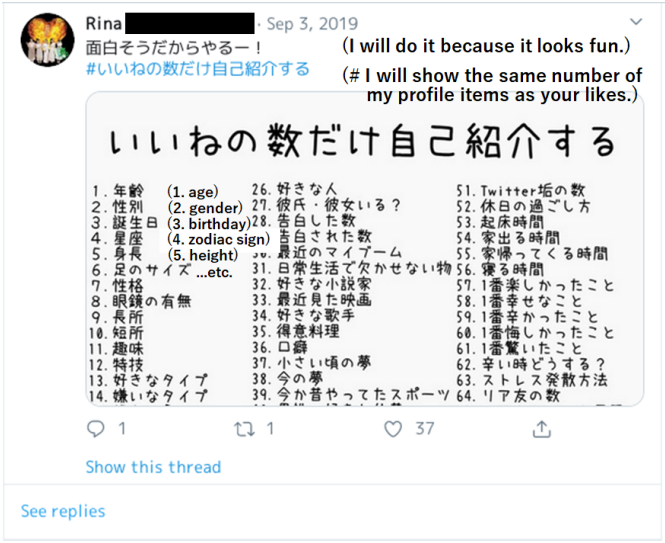


Figure 2. A tweet promising to disclose the same number of submitters' personal profile items as likes to it.

III. A COLLECTION OF TWEETS DISCLOSING SUBMITTERS' PERSONAL INFORMATION

It is difficult to collect tweets disclosing submitters' personal information, such as tweets in Figure 1, directly. To solve this problem, we focused on tweets where submitters promised their followers to disclose the same number of their own personal profile items as likes to their tweets. Figure 2 shows a tweet submitted by *Rina* on September 3, 2019. Both in Figure 1 and Figure 2, her screen name is redacted for privacy. Figure 2 shows that *Rina* promised her followers to disclose the same number of her personal profile items as likes to her tweet. Actually, *Rina* submitted 35 replies disclosing her personal profile items to her tweet shown in Figure 2 from September 3 to 9, 2019. The five tweets shown in Figure 1 were the first five replies submitted by *Rina* to her tweets shown in Figure 2. As of November 20, 2019, we confirmed that 37 likes were given to her tweet shown in Figure 2. Watanabe, Nishimura, Chikuki, Nakajima, and Okada reported that Twitter users seemingly disclosed their personal information honestly when they promised to do it, such as *Rina*'s tweet in Figure 2 [4]. As a result, it is easy to collect tweets disclosing submitters' personal profile items when we collect tweets promising to disclose submitters' personal profile items. The reasons why many Twitter users submitted tweets promising to disclose their personal profile items might be

- they thought they looked fun,
- they wanted to draw attention, and
- they wanted to know how much attention was paid to their tweets.

In other words, they seemingly felt like they were taking part in a game. As a result, most of them kept their promises, like a game rule, and disclosed the same number of their own personal profile items as likes to their tweets. Also, they often used the same sentence in their tweets, like a game password, as shown in Figure 2, *I will show the same number of my profile items as your likes*. In order to collect tweets promising

 ISBN: 978-1-61208-884-6

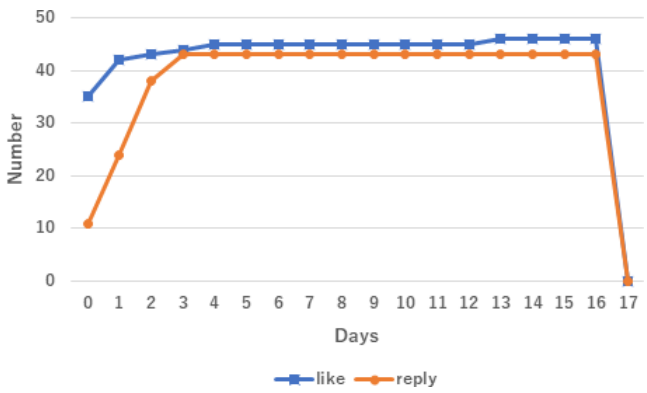


Figure 3. The changes of likes and replies to *okazu_a5*'s December 18, 2020 tweet promising to disclose her personal information.

to disclose submitters' personal profile items, we used the shared sentence as key to collect them. To be specific, we collected these tweets by using twport [14]. Twport helps us to collect tweets where the given sentence is used. By using twport, we collected 233 Japanese tweets promising to disclose submitters' personal information. These tweets were submitted from October 3, 2020 to December 20, 2020. We observed the obtained tweets, likes to them, and their replies once a day at midnight from October 3, 2020 to January 7, 2021.

IV. An Analysis of Tweets Disclosing Submitters' Personal Information

Most of us might think that it is difficult to determine whether unreal name account users disclosed their personal information honestly, and so, it is useless to investigate them. However, Watanabe, Nishimura, Chikuki, Nakajima, and Okada found many unreal name Twitter users who seemingly disclosed their personal information honestly [4]. As a result, we have a chance to investigate unreal name account users who disclosed their personal information on Twitter. In this section, we investigate when and how unreal name Twitter users submitted and deleted their tweets seemingly disclosing their personal information honestly. We think the result shows what they thought about disclosing their personal information honestly. To be specific, we survey Twitter users who promised to disclose their personal information and investigate

- periods from promising to disclose their personal information to deleting their tweets,
- periods from promising to disclose their personal information to submitting their last replies disclosing them, and
- periods from submitting their last replies to deleting their tweets.

Let us consider one example. A Twitter user, *okazu_a5*, submitted a tweet promising her followers to disclose the same number of her own personal profile items as likes on December 18, 2020. Figure 3 shows when *okazu_a5* received likes from audiences and submitted her replies disclosing her personal information. We detected her tweet promising to disclose her personal information at midnight on December 19, 2020, and recorded that she received 35 likes and submitted 10 replies on December 18, 2020 (Day 0 in Figure 3). By December 20, 2020 (Day 2), she received 43 likes and submitted 43 replies disclosing her personal information. Her last reply was submitted on December 21, 2020 (Day 3). After December 21, 2020, she received three more likes, however, submitted no more replies. Figure 3 also shows when *okazu_a5* deleted her tweets: She deleted them on January 4, 2021 (Day 17). She also deleted her tweet promising to disclose her personal information, and so, the number of likes was reduced to zero. In this case, the period from promising to disclose her personal information to deleting her tweets was 17 days. Also, the period from promising to submitting her last reply was three days. The period from submitting her last reply to deleting her tweets was 14 days.

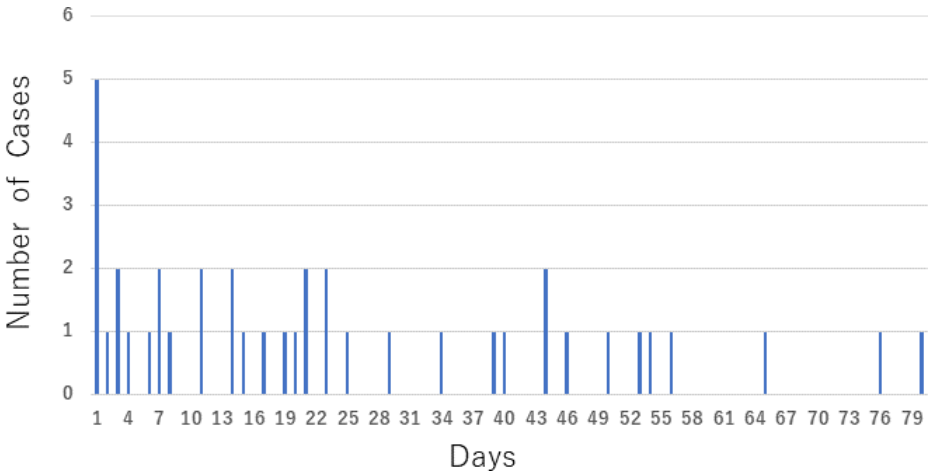
A. Periods from promising to deleting tweets

At first, we discuss the periods from promising to disclose submitters' personal information to deleting their tweets. We found 40 cases where submitters deleted their tweets disclosing their personal information. These 40 cases accounted for 17% of all the cases in the survey. Figure 4 shows the histograms of the periods from promising to disclose submitters' personal information to deleting their tweets. As shown in Figure 4 (a), the most popular day to delete tweets disclosing submitters' personal information was Day 1, the next day when they promised to do it. As shown in Figure 4 (b), 25%, 50%, and 80% of the 40 cases were deleted within four days, three weeks, and six weeks, respectively. One thing to note is that we observed tweets once a day at midnight. As a result, we could not collect cases where submitters deleted their tweets before the end of the day when they submitted tweets promising to disclose their personal information. As mentioned, the most popular day to delete tweets disclosing submitters' personal information was Day 1, the next day when they promised to do them. We think there were many cases where submitters deleted their tweets on Day 0, in other words, before the end of the day when they submitted tweets promising to disclose their personal information.

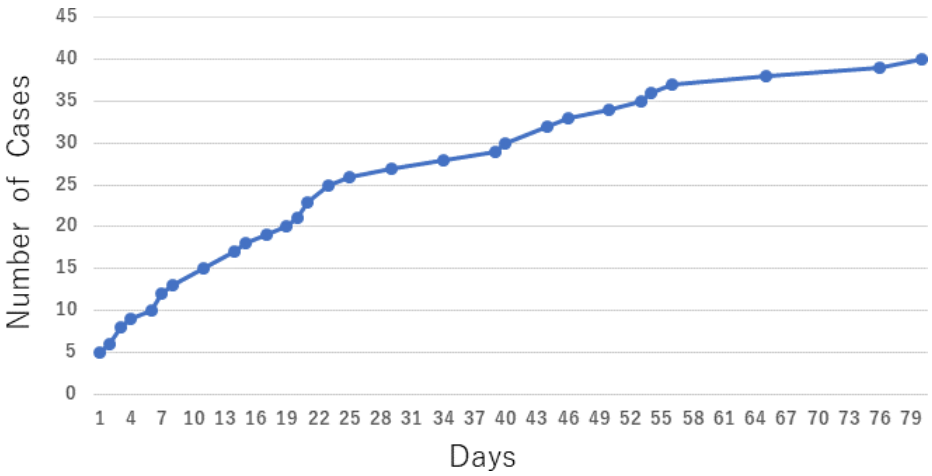
B. Periods from promising to submitting last replies

Next, we discuss the periods from promising to disclose submitters' personal information to submitting their last replies disclosing them. We think that they regarded these periods as times to interact with their audiences. We found 206 cases where submitters submitted replies disclosing their personal information. These 206 cases accounted for 88% of all the cases in the survey. Figure 5 shows the histograms of the periods from promising to submitting their last replies. As shown in Figure 5 (a), the most popular day to submitting their last replies was Day 0, the day when they promised to disclose their personal information. Furthermore, as shown in Figure 5 (b), in 75 % of the 206 cases, last replies were submitted in Day 0 or Day 1. In more than 90 % of the cases, the periods from promising to submitting last replies, in other words, the periods to interact with their audiences, were within four days. On the other hand, there were 27 cases where we could not find replies, however, it does not mean that submitters submitted no replies in these cases. As mentioned, we observed tweets once a day at midnight. We could not detect their replies submitted on the day when they deleted their tweets.

 ISBN: 978-1-61208-884-6



(a) histogram



(b) cumulative histogram

Figure 4. The periods from promising to disclose submitters' personal information to deleting their tweets.

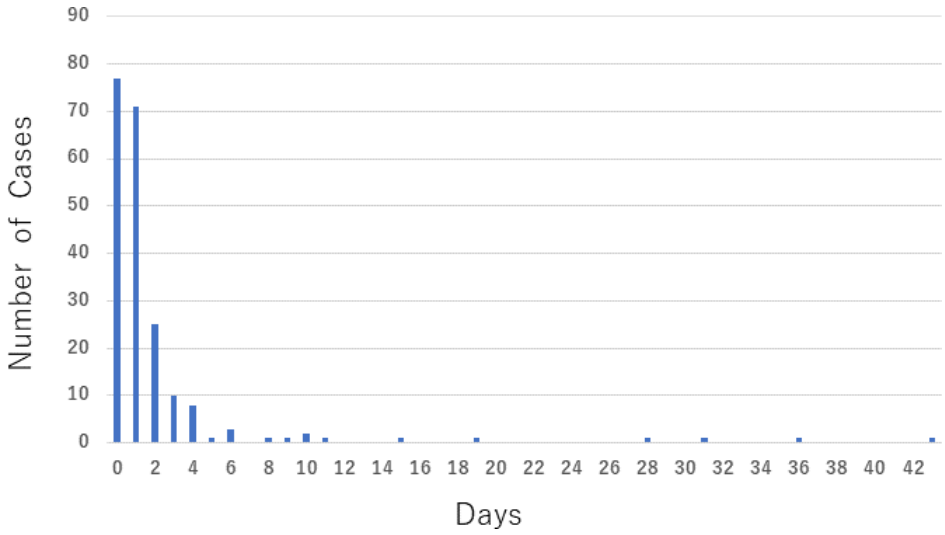
C. *Periods from submitting last replies to deleting tweets*

Finally, we discuss the periods from submitting last replies to deleting tweets disclosing submitters' personal information. We think these periods show how long it took for submitters to think that they could not overlook their potential privacy risks. We found 31 cases where submitters submitted replies disclosing their personal information and deleted them. These 31 cases accounted for 13% of all the cases in the survey. Figure 6 shows the histograms of the periods from submitting last replies to deleting tweets disclosing submitters' personal information. As shown in Figure 6 (a), the most popular day to deleting their tweets after submitting last replies was Day 1, the next day when they submitted last replies. As shown in Figure 6 (b), in 50 % and 80 % of the 31 cases, their tweets were deleted within 18 days and 42 days after submitting last replies, respectively. In this survey, there were nine cases where we found no replies before submitters deleted their tweets promising to disclose their personal information. However, it does not mean that submitters submitted no replies in these cases. This is because, as mentioned, we could not detect replies submitted on the day when submitters deleted their tweets.

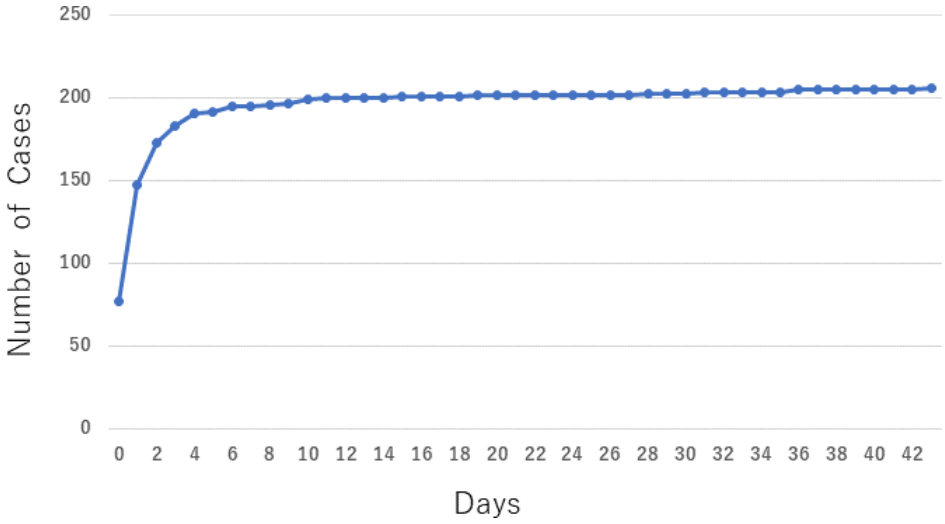
V. Conclusion

It is important to investigate how unreal name account users treat their SNS messages that potentially threaten their privacy and security. In this paper, we investigated how Twitter users treated their tweets seemingly disclosing their personal information. To be specific, we investigated when submitters deleted their tweets disclosing their personal information after they promised to do it. The results of our three-month survey show that 18 % of the surveyed cases were deleted, and 25 %, 50 %, and 80 % of the cases were deleted within four days, three weeks, and six weeks, respectively. The most popular day to delete tweets disclosing submitters' personal information was the next day when they promised to do it. Furthermore, in 90 % of the surveyed cases, submitters interacted with audiences within four days. We intend to survey tweets disclosing submitters' personal information many times

 ISBN: 978-1-61208-884-6



(a) histogram



(b) cumulative histogram

Figure 5. The periods from promising to disclose submitters' personal information to submitting their last replies disclosing them.

in a day. This is because we do not want to miss tweets submitted on the day when submitters deleted them.

REFERENCES

[1] S. Fox et al., Trust and Privacy Online: Why Americans Want to Rewrite the Rules, The Pew Internet & American Life Project, 2000. [Online]. Available: http://www.pewinternet.org/2000/08/20/trust-and-privacy-online/ [accessed: 2021-06-10]

[2] A. Acquisti and R. Gross, "Imagined communities: Awareness, information sharing, and privacy on the facebook," in Proceedings of the 6th International Conference on Privacy Enhancing Technologies, ser. PET'06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 36–58. [Online]. Available: https://doi.org/10.1007/11957454_3 [accessed: 2021-06-10]

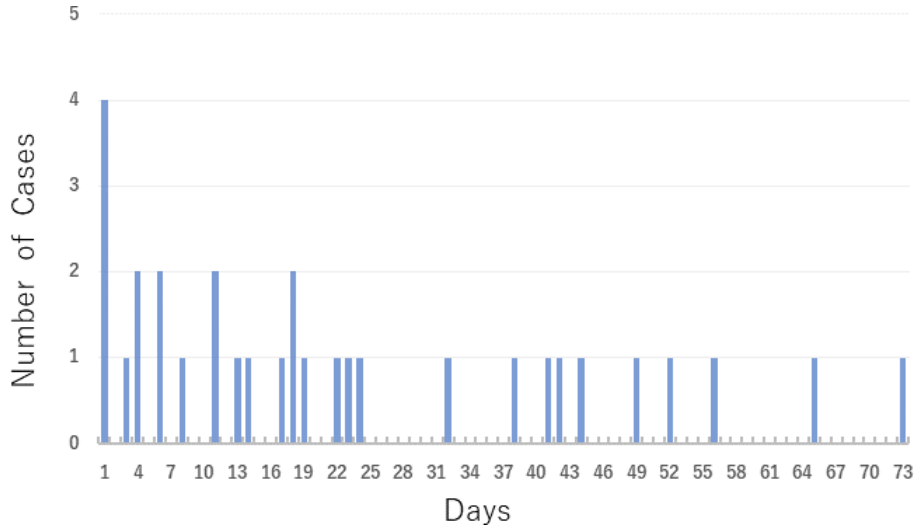
[3] S. B. Barnes, "A privacy paradox: Social networking in the United States." First Monday, vol. 11, no. 9, 2006. [Online]. Available: http://firstmonday.org/article/view/1394/1312 [accessed: 2021-06-10]

[4] Y. Watanabe, H. Nishimura, Y. Chikuki, K. Nakajima, and Y. Okada, "An investigation of twitter users who disclosed their personal profile items in their tweets honestly," in Proceedings of the Sixth International Conference on Human and Social Analytics (HUSO 2020), Oct 2020, pp. 20–25. [Online]. Available: http://www.thinkmind.org/index.php?view=article&articleid=huso_2020_1_40_80035 [accessed: 2021-06-10]

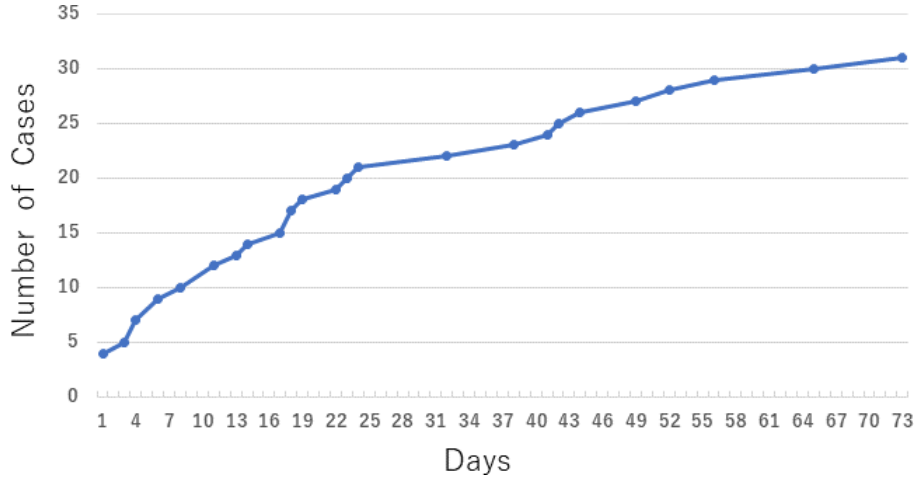
[5] C. Johnson III, Safeguarding against and responding to the breach of personally identifiable information, Office of Management and Budget Memorandum, 2007. [Online]. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/memoranda/fy2007/m07-16.pdf [accessed: 2021-06-10]

[6] B. Krishnamurthy and C. E. Wills, "On the leakage of personally identifiable information via online social networks," Computer Communication Review, vol. 40, no. 1, 2010, pp. 112–117. [Online]. Available: https://doi.org/10.1145/1672308.1672328 [accessed: 2021-06-10]

[7] C. Dwyer, "Digital relationships in the "myspace" generation: Results



(a) histogram



(b) cumulative histogram

Figure 6. The periods from submitting last replies disclosing submitters' personal information to deleting their tweets.

from a qualitative study," in Proceedings of the 40th Annual Hawaii International Conference on System Sciences, ser. HICSS '07. Washington, DC, USA: IEEE Computer Society, 2007, p. 19. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4076409 [accessed: 2021-06-10]

[8] T. Hirai, "Why does "Enjyo" happen on the Web? : An Examination based on Japanese Web Culture," Journal of Information and Communication Research, vol. 29, no. 4, mar 2012, pp. 61–71. [Online]. Available: http://doi.org/10.11430/jsicr.29.4_61 [accessed: 2021-06-10]

[9] A. Viseu, A. Clement, and J. Aspinall, "Situating privacy online: Complex perception and everyday practices," Information, Communication & Society, 2004, pp. 92–114. [Online]. Available: https://doi.org/10.1080/1369118042000208924 [accessed: 2021-06-10]

[10] S. Livingstone, "Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression." New Media & Society, vol. 10, no. 3, 2008, pp. 393–411. [Online]. Available: https://journals.sagepub.com/doi/10.1177/1461444808089415 [accessed: 2021-06-10]

[11] A. N. Joinson, U.-D. Reips, T. Buchanan, and C. B. P. Schofield, "Privacy, trust, and self-disclosure online." Human-Computer Interaction, vol. 25, no. 1, 2010, pp. 1–24. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/07370020903586662 [accessed: 2021-06-10]

[12] Z. Tufekci, "Can You See Me Now? Audience and Disclosure Regulation in Online Social Network Sites," Bulletin of Science, Technology & Society, vol. 28, no. 1, 2008, pp. 20–36. [Online]. Available: https://journals.sagepub.com/doi/abs/10.1177/0270467607311484 [accessed: 2021-06-10]

[13] Y. Watanabe, H. Onishi, R. Nishimura, and Y. Okada, "Detection of school foundation day tweets that can be used to distinguish senders' schools," in Proceedings of the Eleventh International Conference on Evolving Internet (INTERNET 2019), Jun 2019, pp. 34–39. [Online]. Available: https://www.thinkmind.org/index.php?view=article&articleid=internet_2019_2_30_40026 [accessed: 2021-06-10]

[14] yager. twport. [Online]. Available: https://twport.com/ [accessed: 2021-06-10]

Studying the Dynamics of COVID-19 Misinformation Themes using Topic Streams

Thomas Marcoux, Esther Mead, Nitin Agarwal
College of Engineering and Information Technology
University of Arkansas at Little Rock
Little Rock, AR, USA
{txmarcoux, elmead, nxagarwal}@ualr.edu

***Abstract*—The COVID-19 pandemic has seen the emergence of unique misinformation narratives in various outlets, through social media, blogs, etc. This online misinformation has been proven to spread in a viral manner and has a direct impact on public safety. In an effort to improve public understanding, we curated a corpus of 543 misinformation pieces whittled down to 243 unique misinformation narratives along with independent international organizations debunking these stories. Building upon previous applications of topic modeling to COVID-19 related material, we developed a tool leveraging topic modeling to create a chronological visualization of these stories. From our corpus of misinformation stories, this tool has shown to accurately represent the ground truth. This highlights some of the misinformation narratives unique to the COVID-19 pandemic and provides a quick method to monitor and assess misinformation diffusion, enabling policy makers (such as the Arkansas Office of the Attorney General - Arkansas, USA) to identify themes to focus on for communication campaigns. To further explore the potential of topic streams in understanding online opinion, we experiment with multiple topic models and also apply our methodology to YouTube data. The principal difference between our effort and other similar efforts by Google and social media companies is that we are paying special attention to cases of misinformation and scammers that are affecting our region, while also including global cases.**

***Index Terms*—misinformation, disinformation, topic models, topic streams, COVID-19, misinfodemic, narratives.**

I. Introduction

Social media is characterized as a powerful online interaction and information exchange medium. However, it has given rise to new forms of deviant behaviors such as spreading fake news, misinformation, and disinformation. Due to afforded anonymity and perceived diminished personal risk of connecting and acting online, deviant groups are becoming increasingly common. Online deviant groups have grown in parallel with Online Social Networks (OSNs), whether it is black hat hackers using Twitter to recruit and arm attackers, announce operational details, coordinate cyber-attacks [10], and post instructional or recruitment videos on YouTube targeting certain demographics; or state/non-state actors and extremist groups (such as the Islamic State of Iraq and Syria) savvy use of social communication platforms to conduct phishing operations, such as viral retweeting of messages containing harmful URLs leading to malware [6].

More recently, there is a surge in misinformation and scam cases pertaining to COVID-19. The problem of misinformation is actually worse than the pandemic itself. That is why it is called infodemic or more specifically, misinfodemic. Like the pandemic, misinformation cases are also rising exponentially. These cases are more difficult to track than the epidemic, as they can originate in the dark corners of the Internet. To make matters worse, we cannot enforce lockdown on the Internet to stop the spread of this infodemic. This is in part because, during crises, the Internet is usually the first mode of communication and source of information. Although there are some quarantine efforts, for instance form social media companies such as Facebook, YouTube, and retail companies like Amazon are doing their best to block such content, by suspending bad actors or scammers who are spreading misinformation to further their political agenda or to try to profit off of this adversity. But such cases are simply too many and growing too fast. What makes this problem worse is the fact that the information spreads like a wildfire on the Internet, especially the false or misinformation. Many studies have concluded that misinformation travels faster than its corrective information, and the more questionable the misinformation is the faster it travels. This is simply because on social media people usually have a lot more virtual friends than they do in their real life. So, if they share or retweet some misinformation, wittingly or unwittingly, they expose all their virtual friends to the misinformation.

There are similarities between misinformation about COVID-19 and other misinformation cases that we have studied for NATO, US, EU, Singapore, and Canada, etc. Like in other cases, the motivation for spreading COVID-19 misinformation is monetization or to provoke hysteria. Bad actors or scammers are spreading misinformation to further their political agenda or simply trying to profit off of this adversity. For instance, there exists many cases of scammers selling fake masks, fake cures, using fake websites to ask for private/sensitive information from people by posing as government websites. However, there is a significant difference between COVID-19 and other misinformation campaigns that we have studied before. Being a global and rapidly evolving crisis, the nature of misinformation is also extremely diverse and super-fast. Other misinformation campaigns were specific to an entity, event, region, elections, military exercises. However, misinformation about COVID-19 has both global as well as regional narratives. While fake masks, fake cures,

 ISBN: 978-1-61208-884-6

etc. affect a global audience, the regional narratives include promoting medicines for bovine coronavirus as cure for human coronavirus affecting rural/agriculturalist regions. Moreover, the misinformation about COVID-19 ranges from health to policy to religion to geopolitical affairs, i.e., highly topically diverse. Given the volume, velocity, and variety of COVID-19 related misinformation, research is warranted to study such campaigns and their organization. As resources are stretched too thin, government and other regulatory bodies cannot afford to investigate all the misinformation campaigns and scams. Such research could help prioritize investigation of misinformation campaigns and scams.

Therefore, we propose a study of the themes and chronological dynamics of the spreading of misinformation about COVID-19. Our scope focuses on misinformation geographically relevant to us (Arkansas, USA) as well as some global stories, with our main corpus is a collection of unique misinformation stories manually curated by our team. In collaboration with the Arkansas Attorney General, we have shared our findings with their office and made all reports and misinformation stories publicly available online[8]. In addition, we have collected a variety of YouTube video titles and comments. This allows us to compare a curated corpus to a data set more chaotic and true to life. To highlight and visualize these misinformation themes, we use topic modeling, and introduce a tool to visualize the evolution of these themes chronologically.

The rest of this study is structured as follows. First, we will discuss the work done by other researchers in comparable pieces, then describe our methodology, including data collection, processing, and topic modeling methodology. Then in the results section, we will discuss the subjective findings of our misinformation team as well as the scientific topic streams visualizations that support them. Finally, we will briefly discuss our free online resource where our data can be found before presenting our conclusions.

II. Literature Review

The information community has been tackling the issue of misinformation surrounding the COVID-19 pandemic since early in the outbreak. We base the claims found in this paper on the findings that misinformation spreads in a viral fashion and that consumers of misinformation tend to fail at recognizing it as such [13]. In addition to this, we believe this research is essential as rampant misinformation constitutes a danger to public safety [11]. We also believe this research is helpful in curbing misinformation since researchers have found that simply recognizing the existence of misinformation and improving our understanding of it can enhance the larger public's ability to recognize misinformation as such [13]. In order to better understand the misinformation surrounding the pandemic, we look at previous research that has leveraged topic models to understand online discussions surrounding this crisis. Research has shown the benefits of using this technique to understand fluctuating Twitter narratives [17] over time, and also in understanding the significance of media outlets in health communications [12].

To implement topic modeling, we use the Latent Dirichlet Allocation model (LDA). Within the realm of Natural Language Processing (NLP), topic modeling is a statistical technique designed to categorize a set of documents within a number of abstract "topics"[3]. A "topic" is defined as a set of words outlining a general underlying theme. For each document, which in this case, is an individual item of misinformation in our data set, a probability is assigned that designates its "belongingness" to a certain topic. In this study, we use the popular LDA topic model due to its widespread use and proved performances [4]. One point of debate within the topic modeling community is the elimination of stop-words: i.e., analysts should filter common words from their corpus before training a model. Following recent research claiming that the use of custom stop-words adds few benefits [16], we followed the researchers' recommendation and removed common words **after** the model had been trained.

Our model choice has seen use in previous research using LDA for short texts, specifically for short social media texts such as tweets [1, 7, 19]. Some other social media research using homogeneous social media sources such as tweets or blog posts use associated hashtags to provide further context to topic models [2]. We expand this research on social media corpora by focusing one of the largest information propagator on the web: YouTube.

In this paper, we propose to leverage topic models to understand the main underlying themes of misinformation and their evolution over time using a manually curated corpus of known fake narratives.

As a secondary goal, we observe the performances of different topic models for understanding online discourse. To accomplish this, we repeated our methodology on a secondary data set using a Hierarchical Dirichlet Process (HDP) model [18]. For our purposes, the major difference between the two models is that LDA models require a number of topics prior to training and will actively attempt to fit that number to the corpus, potentially leading to biased results. On the other hand, the HDP model infers the number of topics present in the corpus during training.

III. Methodology

This study uses a two-step methodology to produce relevant topic streams. First, through a manual curating process, we aggregate different misinformation narratives for later processing. We consider misinformation narratives, any narrative pushed through a variety of outlets (social media, radio, physical mail, etc.) that has been or is later believably disproved by a third party. This corpus constitutes our input data. Secondly, we use this corpus to train an LDA topic model and to generate subsequent topic streams for analysis. We describe these two steps in more details in the next sections.

A. Collection of Misinformation Stories

Initially, the misinformation stories in our data set were obtained from a publicly available database created by EU-

 ISBN: 978-1-61208-884-6

vsDisinfo in March of 2020 [9]. EUvsDisinfo's database, however, was primarily focused on "pro-Kremlin disinformation efforts on the novel coronavirus". Most of these items represented false narratives that were communicating political, military, and healthcare conspiracy theories in an attempt to sow confusion, distrust, and public discord. Subsequently, misinformation stories were continually gleaned from publicly available aggregators, such as POLITIFACT, Truth or Fiction, FactCheck.org, POLYGRAPH.info, Snopes, Full Fact, AP Fact Check, Poynter, and Hoax-Slayer. The following data points were collected for each misinformation item: title, summary, debunking date, debunking source, misinformation source(s), theme, and dissemination platform(s). The time period of our data set is from January 22, 2020 to July 22, 2020. The data set is comprised of 543 total stories and 243 unique misinformation narratives. For many of the items, multiple platforms were used to spread the misinformation. For example, oftentimes a misinformation item will be posted on Facebook, Twitter, YouTube, and as an article on a website. For our data set, the top platforms used for spreading misinformation were websites, Facebook, Twitter, YouTube, and Instagram, respectively. All the stories found by our team are made public through our partnership with the Arkansas Attorney General Office and can be found on our website.

B. Collection of YouTube Data

In order to observe results in uncontrolled, relevant social media environments, we also gathered YouTube data. Since there exists many studies that concern themselves with Twitter data, we chose YouTube because it is another principal vector of information and communication between users. Using the official YouTube API, we performed separate searches for the following keywords on April 19th: "Coronavirus, Corona, Virus, COVID19, COVID, Outbreak". The result is a set of the most popular videos at that time, as determined by YouTube's algorithm. From this search, we collected a total of 7,727 videos ranging mostly from January 1st to April 19th. For this particular study, in order to focus on the most relevant videos possible, we selected only videos published between March 1st and March 31st (included). This totals 444 videos, which is comparable to the number of narratives studied. For the purposes of this study, we will only look at the video titles. After selecting this corpus, we used the same API to collect comments posted in these videos and gathered a total of 652,120 comments. In order to comply with YouTube's terms of service, this data cannot be made public.

C. Topic Modeling

In order to derive lexical meaning from this corpus, we built a pipeline executing the following steps. First, we processed each document in our text corpus. All that is needed is a text field identified by a date. Because in most cases of word of mouth or social media it is impossible to pinpoint the exact date the idea first emerged, we use the date of publication of the corresponding third party "debunk piece". We trained our LDA model using the Python tool Gensim, with the methodology and pre-processing best practices as described by its author [14] as well as best stop words practices as described earlier [16]. In this study, we found that generating 20 different topics best matched the ground truth as reported by the researchers curating the misinformation stories.

Still using Gensim, we also trained an alternative topic model using hierarchical Dirichlet process (HDP) [18]. The process is the same except for the number of topics. HDP infers the number of topics in a corpus (with a default threshold of 150). Therefore, we only select the first 20 topics, ordered by α, the weight of each document to topic distribution.

Once the models have been trained, we ordered the documents by date and created a numpy matrix where each document is given a score for each topic produced by the model. This score describes the probability that the given document is categorized as being part of a topic, i.e., if a probability score is high enough (more details below), the document is considered to be part of the topic. Through manual observations we noticed that many documents retain "noise probability", giving them a probability to be in every topic of around 1% to 5%. For this reason, we set the probability threshold to a comfortable 10% and noticed consistent results. This allowed us to leverage the Python Pandas library to plot a chronological graph for each individual topic. We averaged topic distribution per day and used a moving average window size of 20 unless otherwise specified. This helped in highlighting the overarching patterns of the different narratives. Note, however, that this process hides some early and late data in our set as there are less data points around that time.

IV. Results

In this section, we discuss the thoughts of our data collection team and the ground truth as they were observed, and compare these with the results obtained through our topic modeling visualization tool.

A. Prominent Misinformation Themes Over Time

Although a variety of misinformation themes were identified, particular dominant themes stood out, changing over time. These themes were considered as dominant based on a simple sum of their frequency of occurrence in our data set. During the month of March, the prominent misinformation theme was the promotion of remedies and techniques to supposedly prevent, treat, or kill the novel coronavirus. During the month of April, the prominent themes still included the promotion of remedies and techniques, but additional prominent themes began to stand out. For example, several misinformation stories attempted to downplay the seriousness of the novel coronavirus. Others discussed the anti-malaria drug hydroxychloroquine. Others promoted the idea that the virus was a hoax meant to defeat President Donald Trump. Others consisted of various attempts to attribute false claims to high-profile people, such as politicians and representatives of health organizations. Also in April, although first signs of these were seen in March, the idea that 5G caused the novel coronavirus began to become more prevalent. During the

 ISBN: 978-1-61208-884-6

month of May, the prominent themes shifted to predominantly false claims made by high-profile people, followed by attempts to convince citizens that face masks are either more harmful than not wearing one, or are ineffective at preventing COVID-19, and how to avoid rules that required their use. The number and variety of identity theft phishing scams also increased during May. Misinformation items attempting to attribute false claims to high-profile people continued throughout May. Also becoming prominent in May were misinformation items attempting to spread fear about a potential COVID-19 vaccine, and items promoting the use of hydroxychloroquine. During the month of June, the prominent theme shifted significantly to attempts to convince citizens that face masks are either more harmful than not wearing one, and how to avoid rules that required their use. Phishing scams also remained prominent during June. During the month of July, the dominant themes of the misinformation items shifted back to attempts to downplay the deadliness of the novel coronavirus. Another prominent theme in July was the proliferation of attempts to convince the public that COVID-19 testing is inflating the results.

B. Topic Streams

After using the tool described in Section III-C, we generated the graphs and tables described and discussed in this section. Our data for this step contained 243 unique misinformation narratives spanning from January 2020 to June 2020, when we stopped data collection. The data was curated by our research team through the process described in the methodology. Each entry contains, among other fields, a "date" used as a chronological identifier, a "title" describing the general idea the misinformation is attempting to convey, and a "theme" field putting the story in a concisely described category. For example, a story given the title *"US Department of Defense has a secret biological laboratory in Georgia"* is categorized in the following theme: *"Western countries are likely to be purposeful creators of the new virus."* Each topic was represented by an identification number up to 20 and a set of 10 words. We picked the three most relevant words that best represented the general idea of each topic. Notably, obvious words such as *covid* or *coronavirus* were removed from the topic descriptions since they are common for every topic.

In Tables I and II, we described some of the twenty topics found by each of our LDA models. These topics were chosen because they each described a precise narrative and have a low topic distribution (or proportion within the corpus). A low proportion is desirable because this indicates the detection of a unique narrative within the corpus; as opposed to an overarching topic including general words such as "world", "outbreak", or "pandemic". Do note that topic inclusiveness is not exclusive and documents can be part of multiple topics. This becomes apparent in tables I: from our topic model, we found a dominant topic encompassing 68% of narratives. It includes words such as "Trump", "outbreak", "president", etc. Some other narratives also included words such as "flu", "news", or "fake". Because the evolution of these narratives are consistent across the corpus and show little temporal fluctuation, we chose not to report on them further. For these reasons, the narratives we focused on below show a low percentage of distribution (Tables I & II).

TABLE I
MOST FREQUENT DOMINANT TOPICS FROM TITLES.

Topic ID	Word 1	Word 2	Word 3	Proportion
10	china	chinese	spread	2%
12	scam	hydroxy...	health	2%
17	state	donald	trump	2%
18	vaccine	gates	bill	5%

TABLE II
MOST FREQUENT DOMINANT TOPICS FROM THEMES.

Topic ID	Word 1	Word 2	Word 3	Proportion
3	fear	spread	western	2%
9	predicted	pandemic	vaccine	2%
16	phishing	hydroxy...	vaccine	2%

1) Using narrative titles as a corpus: The general narratives described by the topics were thus:

- Topic 10 described the narratives related to the Chinese government and its responsibility in the spread of the virus. These stories represented an estimated 2% of the 243 stories collected.
- Topic 12 described the narratives related to personal health and scams or misinformation such as the benefits of hydroxychloroquine. These stories represented an estimated 2% of the 243 stories collected.
- Topic 17 described the narratives related to the response of Donald Trump and his administration. These stories represented an estimated 2% of the 243 stories collected.
- Topic 18 described the narratives related to the involvement of Bill Gates in various conspiracies, mostly linked to vaccines. These stories represented an estimated 4% of the 243 stories collected.

Related studies have found that finger-pointing narratives usually lead to negative sentiment and toxicity in online communities [1, 5, 7].

Figure 1 shows the evolution of Topic 10, the topic describing China-related narratives. It shows that these narratives were already in full force from the beginning of our corpus and slowly came to a near halt during the month of April. We notice a short spike again towards the end of the corpus during the month of June. This is consistent with the ground truth of online narratives that focused on the provenance of the virus during the early stages.

Figure 2 shows the evolution of Topic 12, the topic describing narratives related to health, home remedies, and general hoaxes and scams stemming from the panic. We can see it was consistent with the rise of cases in the United States and panic increased as with the spread of the virus. It is interesting to

 ISBN: 978-1-61208-884-6

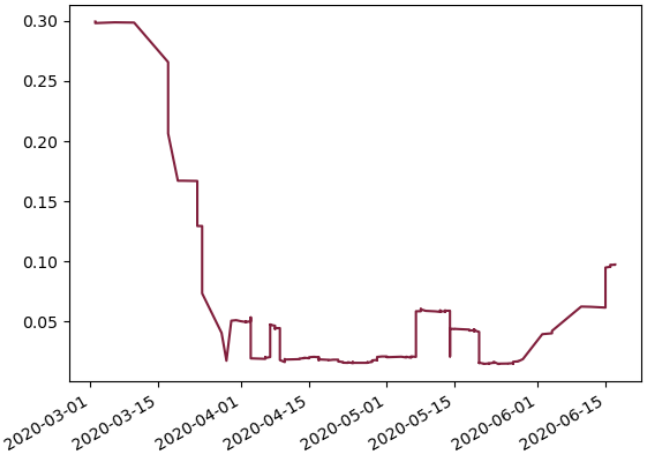


Fig. 1. Topic's probability distribution of titles for topic 10 (keywords: china, chinese, spread)

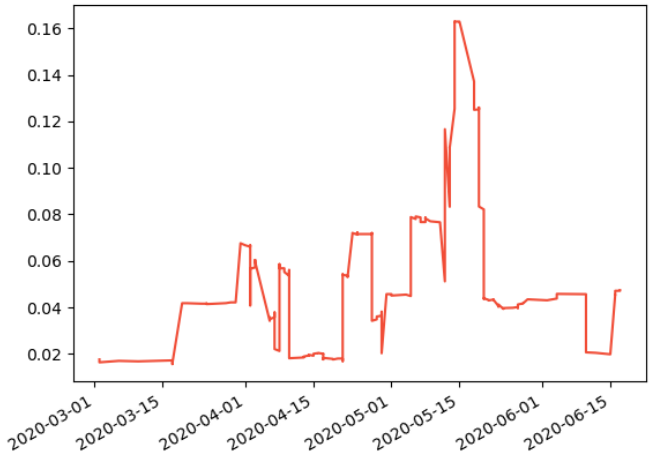


Fig. 3. Topic's probability distribution of titles for topic 17 (keywords: donald, trump, state)

note that this figure roughly coincides with the daily number of confirmed cases for this time period [15].

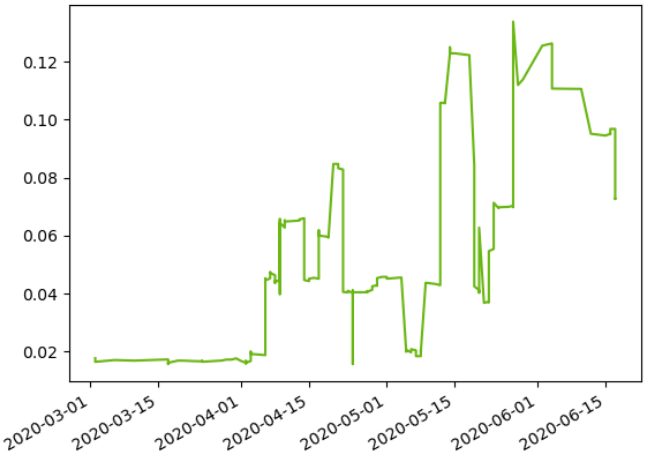


Fig. 2. Topic's probability distribution of titles for topic 12 (keywords: hydroxychloroquine, health, scam)

Figure 3 shows the evolution of Topic 17. This topic described stories related to Donald Trump and his administration. These stories generally referred to claims that the virus was manufactured as a political strategy, or claims that various public figures were speaking out against the response of the Trump administration.

Figure 4 shows the evolution of Topic 18. This topic described stories such as Bill Gates and his perceived involvement with a hypothetical vaccine, and other theories describing the virus' appearance and spread as an orchestrated effort. As with Figure 1, these narratives were especially strong early on (albeit this narrative remained active for a slightly longer time), before coming to a near halt.

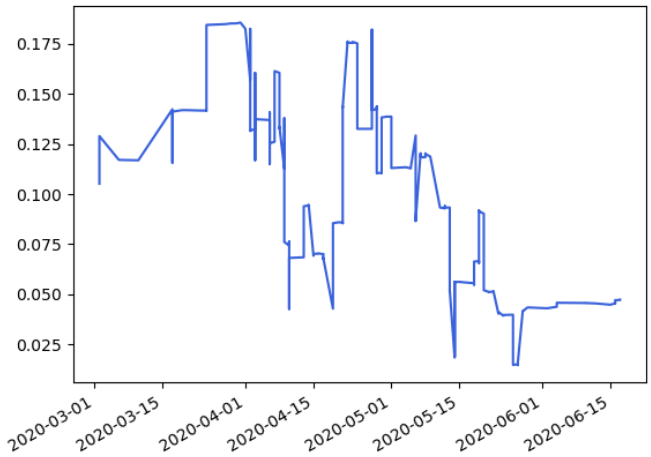


Fig. 4. Topic's probability distribution of titles for topic 18 (keywords: bill, gates, vaccine)

We notice that, as theories about the origins of the virus slowed down, hoaxes and scams increased - as shown on Figure 2. This includes attempts at identity theft, especially toward senior citizens, and attempts to sell miracle cures and miracle personal protection items.

2) Using narrative themes as a corpus: For this section, we inputted narrative themes as the corpus. Note that the topic IDs are independent from the previous set of topics using

 ISBN: 978-1-61208-884-6

titles. Similarly to section IV-B1, we found a dominant topic encompassing 68% of narratives as well. This time including words such as “attempt”, “countries”, and “purposeful”. As for section IV-B1, we chose not to report on that topic as well as other smaller but general topics showing little fluctuation. Therefore, the narratives we focused on below show a low percentage of distribution. The general narratives described by the topics are thus:

- Topic 3 described the narratives related to the speculations on the spread of the virus, especially in an international relations context. These stories represented an estimated 2% of the 243 stories collected.
- Topic 9 described the narratives related to stories claiming the creation and propagation of the virus were either designed or predicted, along with voices claiming a vaccine already exists. These stories represented an estimated 3% of the 243 stories collected.
- Topic 16 described the narratives related to personal health and scams or misinformation such as the benefits of hydroxychloroquine. These stories represented an estimated 2% of the 243 stories collected.

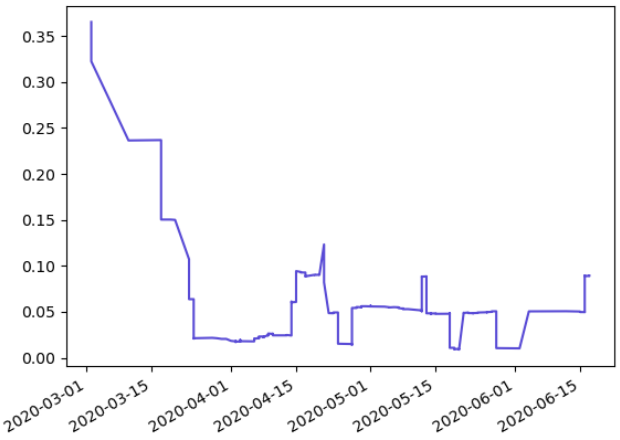


Fig. 5. Topic’s probability distribution of themes for topic 3 (keywords: fear, spread, western)

Figure 5 shows the evolution of Topic 3. It is linked to early fear of the virus and presented narratives as opposing the western block with the East, notably China. It matched closely with Figure 1 and its China-related narratives. In both cases, we see an early dominance of the topic followed by a near halt as the virus touched the United States.

Figure 6 describes the evolution of narratives claiming the virus was predicted or even designed. This figure is consistent with the results shown by Figure 4 which shows claims regarding Bill Gates, early vaccines, etc. They both showed stories of early knowledge of the virus and peaked early, appearing more or less sporadically as time goes on and as cases increased.

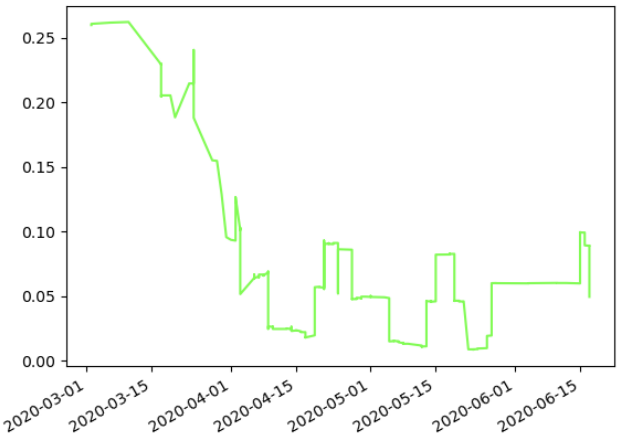


Fig. 6. Topic’s probability distribution of themes for topic 9 (keywords: predicted, pandemic, vaccine)

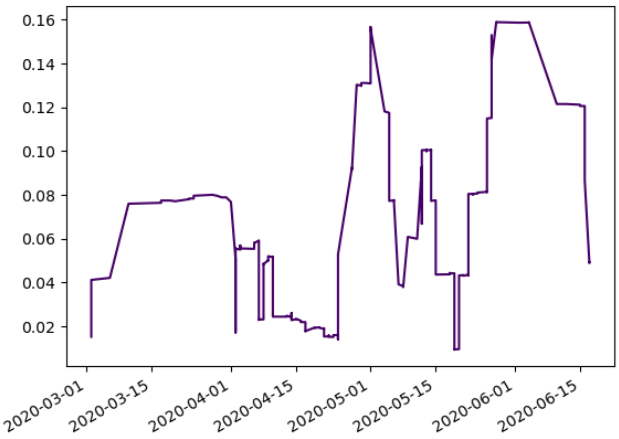


Fig. 7. Topic’s probability distribution of themes for topic 16 (keywords: hydroxychloroquine, vaccine, phishing)

Figure 7 is parallel to Figure 2. Both showed hoax stories promoting scams and health-related misinformation. We noticed an early rise in Figure 7, most likely due to the inclusion of the keyword “vaccines” in the topic, which caused some overlap with Topic 9 as shown in Figure 6.

C. YouTube Data

In this section, we explore how different topic models affect our YouTube data set. We focus on a subset of data published during the month of March to limit the number of comments to process.

1) YouTube Videos: The first observation for this set is that our HDP model did not perform as well as the LDA model. Our HDP model identified one dominant topic present in

87% of videos, with seemingly unrelated identifying keywords (“cases”, “hindi”, “nyc”, “italy”). While the rest of the topics are present in around 1% of the videos. The second most dominant topic (1.8% of documents) also features contradicting words such as “plandemic” and “hospitals”. One would expect language connected to the plandemic narrative in this topic, such as mentions of “Bill Gates” like we saw in the previous sets, but it is missing. There are two possible explanations for this. One is that performance may be due to the size of the set (more in the next section) as there were only 444 video titles processed. The other is that the set features numerous multilingual titles, which may skew results.

TABLE III
RELEVANT TOPICS FROM VIDEO TITLES (LDA MODEL)

Topic ID	Word 1	Word 2	Word 3	Proportion
0	news	update	live	12.4%
17	outbreak	doctor	cases	7.6%
6	plandemic	dempanic	dem	2.7%

Our LDA model, however, behaved as expected and was able to identify major topics, mostly news videos (Topics 0 & 17), as well as what we suspect to be a vehicle of misinformation (Topic 6). As described in Table III and visualized in Figure 8. Figure 8 has been smoothed with a moving average equal to 15% of the total data set size (67) in order to improve legibility and reveal patterns. Due to most of the videos being published late in March, this has removed some granularity towards early March from the plot. However, we notice news topics staying fairly consistent while Topic 6 sees a decline, possibly as the number of covid cases makes maintaining the “fake pandemic” narrative more difficult and other misinformation narratives take over such as various scams and hoaxes as seen in section IV-B1.

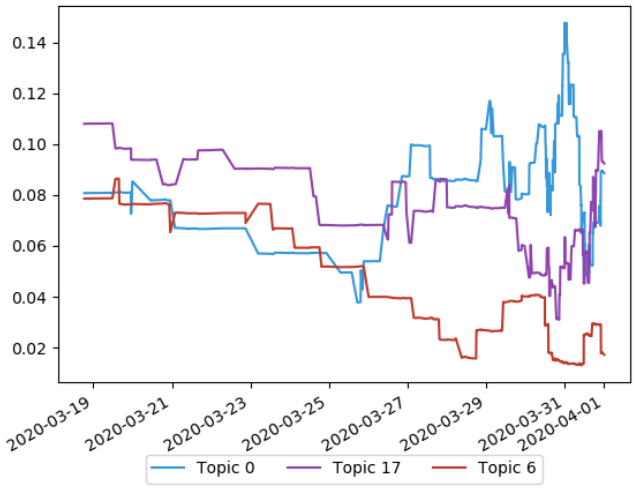


Fig. 8. Topic’s probability distribution of topics 0, 17 & 6 (LDA model)

2) YouTube Comments: Contrary to the previous section, this is a much larger data set of 652,120 comments. This led to better performances, but still inferior to the LDA model. Our HDP model was able to identify non-English comments (11.4% German, 4.5% Spanish, 1.6% French). More importantly, the HDP model identified a topic that could be described as polarizing discourse, some of the most frequent terms including “Trump”, “China’, and “virus”. This topic accounts for 6.6% of the corpus. The evolution of this topic is shown by Figure 9 where we notice that topic is on an upward trend. A moving average equal to 3% of the set size is applied to better identify patterns.

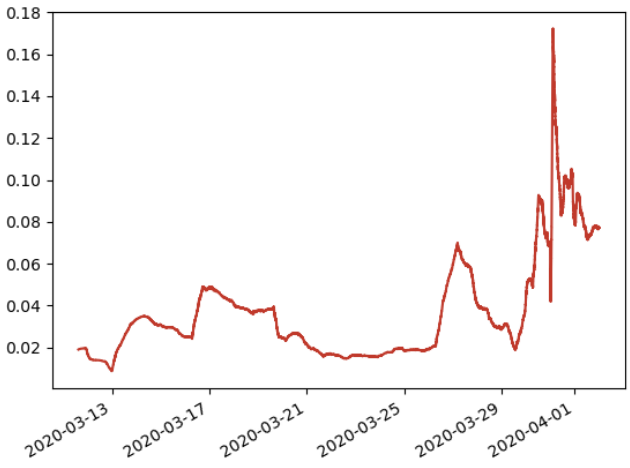


Fig. 9. Topic’s probability distribution of Topic 4 (HDP model)

On this very large set, our HDP model somewhat outperformed LDA for our purposes as it was able to identify a probable topic for misinformation. When applied to our comments set, our LDA model mostly found general terms while also successfully isolating non-English comments. The model did identify a topic with some toxic language and some that could be used in a hostile way or communicate sinophobic sentiments (Topic 7 & 17). See Table IV. While discussion of China has so far been on a downward trend since the start of the pandemic, the mention of the term “virus” along with “china” suggests toxic behavior. See Figure 10.

TABLE IV
RELEVANT TOPICS FROM COMMENTS (LDA MODEL).

Topic ID	Word 1	Word 2	Word 3	Proportion
7	china	virus	made	3.5%
17	trump	dumb	bats	3.3%

D. Public website and citizen science

We have put together a website with known cases of misinformation about COVID-19. As of January 2021, we

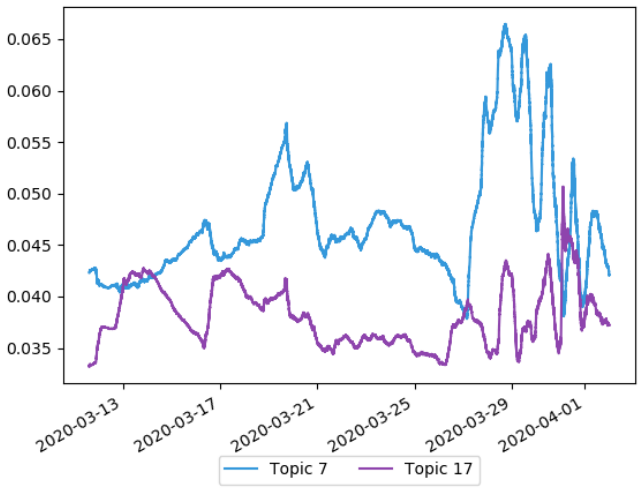


Fig. 10. Topic's probability distribution of Topic 7 & 17 (LDA model)

Fig. 11. COVID-19 Website Front page - Showing the latest misinformation stories

have documented close to 600 cases that we identified from numerous sources (social media - Facebook, YouTube, Twitter, blogs, fake websites, robocalls, text/SMS, WhatsApp, Telegram, and an array of such apps) - see Figure 11 [8]. The principal difference between our effort and other similar efforts by Google and social media companies is that we are paying special attention to cases of misinformation and scammers that are affecting our region, while also including global cases. We update the database periodically with newly detected cases. Moreover, we have put together a list of over 50 tips on the website for people to learn how to spot misinformation. We have also provided a feature for people to report fake websites or scams that are not currently in our database.

Our website uses a three-pronged approach:

- We identify new cases of fake websites, misinformation content, and bad actors. We use social network analysis and cyber forensic methodologies to identify such cases.
- We believe in educating people to be self-reliant because we might not be able to detect all possible cases of misinformation. Therefore, we go through identified cases and prepare a list of common telltale signs to detect whether a piece of information is genuine or not.
- For the cases that are not in our database and people cannot distinguish, we provide a way for people to submit cases of misinformation that we have not captured in our database.

The database of known misinformation cases and scams is publicly available for the research community to use [8]. We envision a tremendous value of this research database to various disciplines. The website is available for regulatory bodies (Arkansas Office of the Attorney General) and any citizen, which serves as an invaluable resource to not only educate people of the misinformation and scams about COVID-19 but also assisting legal authorities in taking action against malicious actors and groups. We are assisting the Arkansas' Attorney General's office by providing reports on cyber forensic evidence about scam/fake websites reported by people. The study presented in this paper will be developed into the system as a real-time campaign tracking feature. We will continue to work with Arkansas' Attorney General's office to assist in their effort to combat COVID-19 misinformation and scams to protect Arkansans.

V. Conclusion

This study has highlighted some of the narratives that surfaced during the COVID-19 pandemic. From January 2020 to July 2020, we collected 243 unique misinformation narratives and proposed a tool to observe their evolution. We have shown the potential of using topic modeling visualization to get a bird's eye view of the fluctuating narratives and an ability to quickly gain a better understanding of the evolution of individual stories. We have seen that the tool is efficient to chronologically represent actual narratives pushed to various outlets, as confirmed by the ground truth observed by our misinformation curating team and independent international organizations. Working with the Arkansas Office of the Attorney General, this study illustrates a relatively quick technique for allowing policy makers to monitor and assess the diffusion of misinformation on online social networks in real-time, which will enable them to take a proactive approach in crafting important theme-based communication campaigns to their respective citizen constituents. We have made most of our findings available online to support this effort.

We have also seen in this study that using carefully curated "themes" - which offer a lexical value close to the abstract topics provided by the LDA model - yields similar results to

 ISBN: 978-1-61208-884-6

using misinformation narratives "title". For this reason, we scaled up our data and repeated our methodology on social media data: YouTube video titles, and their comments. Over concerns of our LDA topic model becoming difficult to scale, we experimented with a HDP (Hierarchical Dirichlet Process) model, which attempts to infer the number of topics.

We found promising but unsurprisingly less precise results. We notice that HDP was able to isolate a probable subset of polarizing comments. One possible way forward would be to use HDP to identify these subsets, filter out irrelevant comments, then apply the LDA model. This may reveal various narratives, some of them spreading misinformation, and further automate the process of identifying online misinformation in uncontrolled spaces.

ACKNOWLEDGEMENT

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920, IIS-1636933, ACI-1429160, and IIS-1110868), U.S. Office of Naval Research (N00014-10-1-0091, N00014-14-1-0489, N00014-15-P-1187, N00014-16-1-2016, N00014-16-1-2412, N00014-17-1-2675, N00014-17-1-2605, N68335-19-C-0359, N00014-19-1-2336, N68335-20-C-0540), U.S. Air Force Research Lab, U.S. Army Research Office (W911NF-20-1-0262, W911NF-16-1-0189), U.S. Defense Advanced Research Projects Agency (W31P4Q-17-C-0059), Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment at the University of Arkansas at Little Rock, and the Australian Department of Defense Strategic Policy Grants Program (SPGP) (award number: 2020-106-094). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

REFERENCES

[1] A. Abd-Alrazaq et al. "Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study". In: vol. 22. 4. 2020, e19016. DOI: 10.2196/19016. URL: http://www.ncbi.nlm.nih.gov/pubmed/32287039.

[2] M. H. Alam, W.-J. Ryu, and S. Lee. "Hashtag-based topic evolution in social media". In: vol. 20. 6. Nov. 2017, pp. 1527–1549. DOI: 10.1007/s11280-017-0451-3. URL: https://doi.org/10.1007/s11280-017-0451-3.

[3] D. M. Blei, J. D. Lafferty, and A. N. Srivastava. *Text Mining: Classification, Clustering, and Applications*. CRC Press, 2009, pp. 71–88.

[4] D. M. Blei, A. Y. Ng, and M. I. Jordan. "Latent dirichlet allocation". In: vol. 3. 2003, pp. 993–1022.

[5] H. Budhwani and R. Sun. "Creating COVID-19 Stigma by Referencing the Novel Coronavirus as the "Chinese virus" on Twitter: Quantitative Analysis of Social Media Data". In: vol. 22. 5. May 2020, e19301. DOI: 10.2196/19301. URL: http://www.ncbi.nlm.nih.gov/pubmed/32343669.

[6] M. Calabresi. *Inside Russia's Social Media War on America.* 2017. URL: https://time.com/4783932/inside-russia-social-media-war-america/. (accessed: 01.19.2021).

[7] R. Chandrasekaran et al. "Topics, Trends, and Sentiments of Tweets About the COVID-19 Pandemic: Temporal Infoveillance Study". In: vol. 22. 10. Oct. 2020, e22624. DOI: 10.2196/22624. URL: http://www.jmir.org/2020/10/e22624/.

[8] COSMOS. *COSMOS - COVID-19 Misinformation Tracker*. 2021. URL: https://cosmos.ualr.edu/covid-19. (accessed: 06.15.2021).

[9] EUvsDisinfo. *EUvsDisinfo. March 16, 2020. The Kremlin and Disinformation About Coronavirus.* 2020. URL: https://euvsdisinfo.eu/the-kremlin-and-disinformation-about-coronavirus/. (accessed: 01.19.2021).

[10] S. Al-khateeb et al. *Exploring Deviant Hacker Networks (DHM) on Social Media Platforms*. Vol. 11. Journal of Digital Forensics, Security and Law, 2016, pp. 7–20. DOI: 10.15394/jdfsl.2016.1375. URL: https://commons.erau.edu/jdfsl/vol11/iss2/1.

[11] R. Kouzy et al. "Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter". eng. In: vol. 12. 3. Publisher: Cureus. Mar. 2020, e7255–e7255. DOI: 10.7759/cureus.7255. URL: https://pubmed.ncbi.nlm.nih.gov/32292669.

[12] Q. Liu et al. "Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach". In: vol. 22. 4. Apr. 2020, e19118. DOI: 10.2196/19118. URL: http://www.jmir.org/2020/4/e19118/.

[13] G. Pennycook et al. "Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention". In: vol. 31. 7. _eprint: https://doi.org/10.1177/0956797620939054. 2020, pp. 770–780. DOI: 10.1177/0956797620939054. URL: https://doi.org/10.1177/0956797620939054.

[14] R. Řehůřek and P. Sojka. "Software Framework for Topic Modelling with Large Corpora". In: May 2010, pp. 45–50. DOI: 10.13140/2.1.2393.1847.

[15] H. Ritchie et al. *United States: Coronavirus Pandemic - Our World in Data*. 2020. URL: https://ourworldindata.org/coronavirus/country/united-states?country=~USA. (accessed: 07.29.2020).

[16] A. Schofield, M. Magnusson, and D. Mimno. "Pulling Out the Stops: Rethinking Stopword Removal for Topic Models". In: *15th Conference of the European Chapter of the Association for Computational Linguistics*. Vol. 2. Association for Computational Linguistics. 2017, pp. 432–436.

[17] H. Sha et al. *Dynamic topic modeling of the COVID-19 Twitter narrative among U.S. governors and cabinet executives*. 2020. arXiv: 2004.11692 [cs.SI].

[18] Y. W. Teh et al. "Hierarchical Dirichlet Processes". In: vol. 101. 476. Publisher: Taylor & Francis _eprint: https://doi.org/10.1198/016214506000000302. 2006, pp. 1566–1581. DOI: 10.1198/016214506000000302. URL: https://doi.org/10.1198/016214506000000302.

[19] Y. Zhang, W. Mao, and J. Lin. "Modeling Topic Evolution in Social Media Short Texts". In: *2017 IEEE International Conference on Big Knowledge (ICBK)*. 2017, pp. 315–319.

 ISBN: 978-1-61208-884-6

Assessing the Impact of Hotel Services on Customer Rating Using Fuzzy String Matching and Belief Networks

Alexandros Bousdekis
Business Informatics Lab, Department of Business Administration, School of Business
Athens University of Economics and Business
Athens, Greece
e-mail: albous@mail.ntua.gr

Dimitris Kardaras
Business Informatics Lab, Department of Business Administration, School of Business
Athens University of Economics and Business
Athens, Greece
e-mail: dkkardaras@yahoo.co.uk

Stavroula G. Barbounaki
Merchant Marine Academy of Aspropyrgos,
Aspropyrgos, Greece
e-mail: sbarbounaki@yahoo.gr

***Abstract*— Online review comments have become a popular and efficient way for sellers to acquire feedback from customers and improve their service quality. These online reviews in the e-tourism era, in the format of both textual reviews (comments) and ratings, generate an electronic Word Of Mouth (eWOM) effect, which influences future customer demand and hotels' financial performance, and thus, have significant business value. This paper proposes an approach for hotel quality evaluation according to online review comments and ratings using Fuzzy String Matching (FSM) for mining customers' opinions and Bayesian Belief Networks (BBN) for evaluating the attributes that contribute to the review rating. The proposed approach was applied to a dataset from TripAdvisor. The results show that the proposed approach is able to model the complex dynamics of online hotel review data, which are derived from both the textual nature of the review comments and the uncertain relationships between these comments and the review rating.**

Keywords-e-tourism; data analytics; machine learning; tourism management; service quality.

I. INTRODUCTION

Online comments have become a popular and efficient way for sellers to acquire feedback from customers and improve their service quality [1]. According to a survey, with the increased popularity of online bookings, 53% of travellers state that they would be unwilling to book a hotel that had no reviews, while a 10% increase in travel review ratings would increase bookings by more than 5% [2]. Customer online reviews of hotels have significant business value in the e-commerce and big data era, while they affect room occupancy [3], revenue, prices [4] and market share [5]. These online reviews in the e-tourism era, in the format of both textual reviews (comments) and ratings, generate an electronic Word Of Mouth (eWOM) effect, which influences future customer demand and hotels' financial performance [6].

Hotel owners want to know the details about hotel guests' experiences, to improve the corresponding product and service attributes, and customers' overall evaluation of the hotel stay experience, to obtain a snapshot of the hotel's operational performance and overall customer satisfaction [7][8]. Although the direct measurement of customer ratings in terms of closed-ended survey questions can show overall customer satisfaction in a direct way [7][8], they suffer from confounding the data of customers' true evaluation because of variations in survey design from different approaches [9].

Recently, many studies have focused on textual reviews [8][10]. In contrast to a pre-designed questionnaire survey, online textual reviews have an open-structured form and can show customer consumption experiences, highlight the product and service attributes customers care about, and provide customers' perceptions in a detailed way through the open-structure form [8]. The provided information is free from obvious bias and is helpful in understanding and assessing hotel performance [11]. In addition, such information is inexpensive and efficient to collect [12]. However, the exploitation of online textual reviews is still largely underexplored [8], while there is a lack of advanced data analytics approaches and algorithms for modeling complex dynamics of online hotel review data.

Hotel quality evaluation from online reviews is an emerging research field; however, the vast majority of existing research works have been performed from a tourism management perspective. Therefore, the applied methods and algorithms are limited to descriptive statistics, e.g., using well-established regression models. However, the increasing amount of online reviews as the core means for customers to express their level of satisfaction about a hotel pose significant challenges to the data analytics and computer science community for the development of advanced data analytics models aiming at providing a higher level of intelligence and thus, increased business value.

In this paper, we propose an approach for hotel quality evaluation from online reviews using Fuzzy String Matching (FPM) and Bayesian Belief Networks (BBN). The objective is to provide a unified algorithm, which both: (i) mines customers' opinions from online hotel reviews (review comments and rating); and, (ii) evaluates the hotel performance by identifying how the various attributes (e.g., location, cleanliness, breakfast, etc.) affect the overall review rating. The rest of the paper is organized as follows: Section II presents the related work on methods and approaches for hotel evaluation based on online review comments. Section III describes the research methodology and the proposed approach for hotel quality evaluation from online reviews using FSM and BBN. Section IV presents the results from

 ISBN: 978-1-61208-884-6

the adoption of the proposed methodology on a dataset from TripAdvisor. Section V concludes the paper and outlines our plans for future work.

II. Related Work

The business value of online consumer reviews has emerged in recent year in the hotel industry aiming at solving the problems confronted by the traditional hotel service quality assessment methods [13]. For example, Kim and Park [14] performed hierarchical multiple regressions in order to examine the effects of traditional customer satisfaction relative magnitude and social media review ratings on hotel performance and found that social media review rating is a more significant predictor. In the traditional hotel quality assessment, domain experts or customers are asked to fill in a questionnaire and score each evaluation index to be used in a service quality assessment model [15]-[17]. On the contrary, online comments are made by a large amount of customers with actual user experience shortly after the consumption is completed. In addition, the increasing amount of reviews-related data pave the way for the use of advanced data analytics and machine learning algorithms that outperform traditional statistical methods based on sampling [2].

Technical attributes of online textual reviews can explain significant variations in customer ratings and can have a significant effect on customer ratings [18][19]. In this direction, Zhao et al. [8] developed an approach for predicting overall customer satisfaction using the technical attributes of online textual reviews and customers' involvement in the review community. They calculated subjectivity and polarity measurements by using naïve Bayes classifier and sentiment analysis. Berezina et al. [10] investigated the underpinnings of satisfied and unsatisfied customers by applying text mining on online reviews.

The literature is rich of methodologies based on descriptive statistics aiming at providing insights on hotel quality performance for various datasets. Xie et al. [20] applied statistical methods in order to assess how several characteristics, such as timeliness of the response, length of the response, number of responses, etc., contributes to the hotel's financial performance. Figini et al. [21] compared the rating dynamics of the same hotels in two online review platforms, which mainly differ in requiring or not requiring proof of prior reservation before posting a review (respectively, a verified vs a non-verified platform). Xie et al. [22] examined the effect of factors of online consumer review, including quality, quantity, consistency, on the offline hotel occupancy (i.e., how popular the hotel is among consumers).

De Pelsmacker et al. [3], the extent to which digital marketing strategies influence hotel room occupancy and revenue per available room and how this mechanism is different for different types of hotels in terms of star rating and independent versus chain hotels was investigated. Li et al. [23] examined the determinants of customer satisfaction in hospitality venues through an analysis of online reviews using text mining and content analysis. Zhao et al. [24] investigated the impacts of online review and source features (usefulness, reviewer expertise, timeliness, volume, valence and comprehensiveness) upon travelers' online hotel booking intentions by applying factor analysis and regression analysis. Zhou et al. [11] compared customer satisfaction by classifying several attributes influencing customer satisfaction in: satisfiers, dissatisfiers, bidirectional forces, and neutrals. Ye and Yu [25] applied qualitative research methods and extracted six main factors influencing the positive or negative emotions of the comments of travelers staying in the hotel.

Radojevic et al. [26] conducted a multilevel analysis of factors affecting customer satisfaction, such as service encounter, visitor, visitor's nationality, hotel, and destination. Nunkoo et al. [27] applied a multi-group analysis and an importance-performance map analysis by means of Partial least squares structural equation modeling (PLS-SEM) in order to differentiate between service quality performance scores and their influences on customer satisfaction across accommodation with a different star grading. Schuckert et al. [28] assessed social media content produced by customers and related review-management strategies of domestic and international hotel chains with the use of multilevel regression.

As mentioned earlier, the increasing amounts of reviews-related data require advanced data analytics and machine learning methods for exploiting the full potential. To this end, Sánchez-Franco et al. [2] assessed whether terms related to guest experience can be used to identify ways to enhance hospitality services. They developed a model based on naïve Bayes classifier in order process vast amount of data and to classify reviews of hotels. Ku et al. [29] developed a framework in order to integrate visual analytics and (deep) machine learning techniques, such as clustering for text classification and Convolutional Neural Networks (CNN), to investigate whether hotel managers respond to positive and negative reviews differently and how to use a deep learning approach to prioritize responses. Reference [1] combined fuzzy comprehensive evaluation and fuzzy cognitive maps aiming at identifying the causal relations among evaluation indexes from online comments. Based on this, their proposed approach recommends more economical solutions for improving the service quality by automatically getting more trustworthy evaluation from a large amount of less trustworthy online comments.

III. Research Methodology

Our research methodology consists of four main steps: (i) Extracting the evaluation criteria from online comments; (ii) Mining customers' opinions using FSM; (iii) Assignment of sentiment scores to a discrete scale; and, (iv) Applying BBN for assessing the impact of hotel services to the customer rating. These steps are described in detail in the following sub-sections.

A. Extracting the Evaluation Criteria from Online Comments

The proposed approach utilizes three fields from the online hotel reviews: (i) *review title*; (ii) *review comments*;

and, (iii) *review rating*. This step of the methodology processes the *review title* and the *review comments* in order to extract the evaluation criteria from the online comments. More specifically, based upon an evaluation index for hotel service quality [1], this step identifies the criteria mentioned in the hotel reviews under examination, e.g., location, price, breakfast, room space, etc. In this way, the criteria are defined dynamically out of the pre-defined list, according to the dataset of the available online comments. The extracted evaluation criteria are further processed with the use of Fuzzy Pattern Match Template (FPMT), as we describe in Section III.B. Moreover, along with the review rating, they derive the parent nodes of the BBN, as we describe in Section III.C and Section III.D.

B. Mining Customers' Opinions Using Fuzzy String Matching

Since online comments are written in natural and informal language, there is the need to mine customers' opinions so that they subsequently feed into the BBN for further processing. FSM, alternatively mentioned as fuzzy string searching or approximate string matching, has been developed in the framework of fuzzy set and possibility theory in order to take into account the imprecision and the uncertainty pervading values, which have to be compared in a matching process [30]. This technique has proved effective for implementing patterns of approximate reasoning in expert system inference engines, and for designing retrieval systems capable of managing incomplete and fuzzy information data bases and vague queries.

In online review comments, different customers may use different words or phrases to express their opinions, while the comments may be vague. For example, poor cleanliness can be expressed as: "The room was too dirty", "Very dirty", etc. Regular expression is an efficient pattern match [31] technology to identify the specific pattern strings from a long text. A simple example of regular expression is "[\s\S]∗?[room|bathroom][\s\S]∗?dirty[\s\S]∗?" that can match "The room was too dirty." However, the regular expression method causes a binary value result: match or not match.

In the proposed approach, we apply FPMT [1] as an effective fuzzy string matching method to deal with the vagueness of the free text online comments. FPMT is a set of pattern strings with membership degrees, denoted as:

$$FPMT = \{(p_1, w_1), (p_2, w_2), \ldots, (p_i, w_i), \ldots, (p_n, w_n)\}$$

where p_i is a pattern string described by regular expression, and w_i is the membership degree that a string falls into the object FPMT when the string matches p_i. When a string matches multiple pattern strings at the same time, the max membership degree of these pattern strings will be selected as the final membership degree. Although this method results in some mismatched cases due to the limitation of pattern strings, this causes little impact on the final result, because there are many redundant comments with similar semantics.

The output of customers' opinions mining is a fuzzy evaluation of the extracted criteria. Specifically, first, the extracted evaluation criteria of hotel quality are assigned to a five-level Likert scale (1 – Very Low, 2 – Low, 3- Neutral, 4 – High, 5 – Very High), which serve as an equivalent to responses of a Likert scale questionnaire. Then, following the approach proposed by [32], this step considers the median of the resulting responses in order to represent the magnitude of causality among the evaluation criteria to be used as FCM concepts in Section III.C.

C. Assignment of Sentiment Scores to a Discrete Scale

In this step, the sentiment scores extracted from the previous step for each criterion are assigned to a discrete scale consisting of ranges of sentiment score values. The number of the scale items should be the same with the respective scale of the review rating so that they are directly comparable. For example, if the review rating takes values between 1 and 5 (which is the most common case), the sentiment scores are classified to a respective discrete scale:

- [-1, -0.6] is assigned to "DISASTER"
- (-0.6, -0.2] is assigned to "MANY THINGS NEED TO BE IMPROVED"
- (-0.2, +0.2] is assigned to "FAIR ENOUGH"
- (+0.2, +0.6] is assigned to "PERFECT"
- (+0.6, +1] is assigned to "ABSOLUTELY PERFECT"

D. Applying Bayesian Belief Networks for Assessing the Impact of Hotel Services to the Customer Rating

In this step, the relationships between the sentiment discrete scale created in the previous step and the review rating of the customer are modelled in a probabilistic model with the use of BBN. A BBN is a powerful tool for knowledge representation and reasoning under conditions of uncertainty and visually presents the probabilistic relationships among a set of variables [32]. A BBN has many advantages, such as combination of different sources of knowledge, explicit treatment of uncertainty and support for decision analysis, and fast responses.

More formally, BBNs are directed acyclic graphs whose nodes represent random variables from the domain of interest, in the Bayesian sense. Therefore, a BN is defined as a pair $B = (G, \Theta)$. $G = (V, E)$ is a Directed Acyclic Graph (DAG) where $V = \{v_1, \ldots, v_n\}$ is a collection of n nodes, $E \subset V \times V$ a collection of edges and a set of parameters Θ containing all the Conditional Probabilities (CP) of the network.

Each node $v \in V$ of the graph represents a random variable X_V with a state space $\boldsymbol{X}_V$ which can be either discrete or continuous. An edge $(v_i, v_j) \in E$ represents the conditional dependence between two nodes $v_i, v_j \in V$ where v_i is the parent of child v_j. If two nodes are not connected by an edge, they are conditional independent. Because a node can have more than one parent, let π_v be the set of parents for a node $v \in V$. Therefore, each random variable is independent of all nodes $V \setminus \pi_v$. For each node, a Conditional Probability Table (CPT) contains the CP distribution with parameters $\theta_{xi|\pi i}$

$:=P(x_i|\pi_i)\ \epsilon\ \Theta$ for each realization x_i of X_i conditioned on π_i. The joint probability distribution over V is visualized by the BN and can be defined as

$$P(X_1, \dots, X_n) = \prod_{i=1}^{n} P(X_i|\pi_i) \quad (1)$$

The outcome indicates the probability of having a specific value of the overall rating given the values of different services (criteria). The user is able to perform queries in order to assess the impact of each criterion on the review rating, but also combinations of criteria.

IV. RESULTS

The proposed methodology was applied to a dataset from TripAdvisor. The nodes of the BBN are shown in Table I. These nodes represent the parent nodes derived from the extracted evaluation criteria from FPMT (C1-C9) along with the review rating (C10), which constitutes the unique child node of the BBN.

TABLE I. THE EXTRACTED EVALUATION CRITERIA

ID	Nodes		ID	Nodes
C1	Location		C6	Quiet
C2	Personnel		C7	Parking
C3	Cleanliness		C8	Interior Design
C4	Room Space		C9	Bed
C5	Breakfast		C10	Review Rating

After the fuzzy evaluation of the aforementioned criteria for each hotel, the BBN is created. The BBN consists of two conceptual layers: the upper layer includes all the evaluation criteria (C1-C9) and the bottom layer includes the review rating provided by the customer (C10). The structure of the BBN is depicted in Figure 1. All the parent nodes are linked to the child node. Based upon this structure, the CPT is calculated for each node. Based upon this structure, the parameters of the BBN are learned.

Upon queries, the Conditional Probability (CP) $P(C10|C_i)$ is calculated. Table II presents the results from some indicative queries. In addition, the queries may deal with specific evaluation criteria in order to assess their impact on the customers' overall review rating. According to the queries, the BBN derives more focused results, e.g., for a specific hotel, group of hotels, location, etc. Finally, the adopted modelling approach may serve as a classifier for predicting the review rating of a customer based upon their review comments. Table III presents the resulting confusion matrix that derives the precision and recall of the classifier as follows:

$$Precision = \frac{TP}{TP + FP} = \frac{41}{41 + 3} = 93.1\% \quad (2)$$

$$Recall = \frac{TP}{TP + FN} = \frac{41}{41 + 9} = 82\% \quad (3)$$

TABLE II. RESULTS FROM INDICATIVE QUERIES

Values of Parent Nodes	Values of Child Node	$P(C10|C_i)$
C1={FAIR ENOUGH}, C2={PERFECT}, C3={MANY THINGS NEED TO BE IMPROVED}, C4={PERFECT}, C5={PERFECT}, C6={FAIR ENOUGH}, C7={MANY THINGS NEED TO BE IMPROVED}, C8={FAIR ENOUGH}, C9={DISASTER}	3 stars	0.332
C1={MANY THINGS NEED TO BE IMPROVED}, C2={PERFECT}, C3={DISASTER}, C4={DISASTER}, C5={PERFECT}, C6={FAIR ENOUGH}, C7={FAIR ENOUGH}, C8={PERFECT}, C9={DISASTER}	2 stars	0.241
C1={PERFECT}, C2={ABSOLUTELY PERFECT}, C3={MANY THINGS NEED TO BE IMPROVED}, C4={PERFECT}, C5={PERFECT}, C6={ABSOLUTELY PERFECT}, C7={FAIR ENOUGH}, C8={FAIR ENOUGH}, C9={PERFECT}	4 stars	0.214
C1={ABSOLUTELY PERFECT}, C2={ABSOLUTELY PERFECT}, C3={FAIR ENOUGH}, C4={ABSOLUTELY PERFECT}, C5={FAIR ENOUGH}, C6={FAIR ENOUGH}, C7={FAIR ENOUGH}, C8={FAIR ENOUGH}, C9={PERFECT}	4 stars	0.183
C1={FAIR ENOUGH}, C2={PERFECT}, C3={MANY THINGS NEED TO BE IMPROVED}, C4={FAIR ENOUGH}, C5={PERFECT}, C6={PERFECT}, C7={MANY THINGS NEED TO BE IMPROVED}, C8={PERFECT}, C9={FAIR ENOUGH}	3 stars	0.144
C1={FAIR ENOUGH}, C2={ABSOLUTELY PERFECT}, C3={FAIR ENOUGH}, C4={PERFECT}, C5={ABSOLUTELY PERFECT}, C6={FAIR ENOUGH}, C7={MANY THINGS NEED TO BE IMPROVED}, C8={FAIR ENOUGH}, C9={FAIR ENOUGH}	3 stars	0.139
C1={MANY THINGS NEED TO BE IMPROVED}, C2={PERFECT}, C3={FAIR ENOUGH}, C4={PERFECT}, C5={ABSOLUTELY PERFECT}, C6={PERFECT}, C7={FAIR ENOUGH}, C8={PERFECT}, C9={FAIR ENOUGH}	4 stars	0.091
C1={PERFECT}, C2={PERFECT}, C3={FAIR ENOUGH}, C4={PERFECT}, C5={ABSOLUTELY PERFECT}, C6={PERFECT}, C7={FAIR ENOUGH}, C8={MANY THINGS NEED TO BE IMPROVED}, C9={FAIR ENOUGH}	5 stars	0.073

The Precision results are quite satisfactory, while the Recall results can be further improved. The BN model sticks to the initially identified relationships, i.e., the ones that have

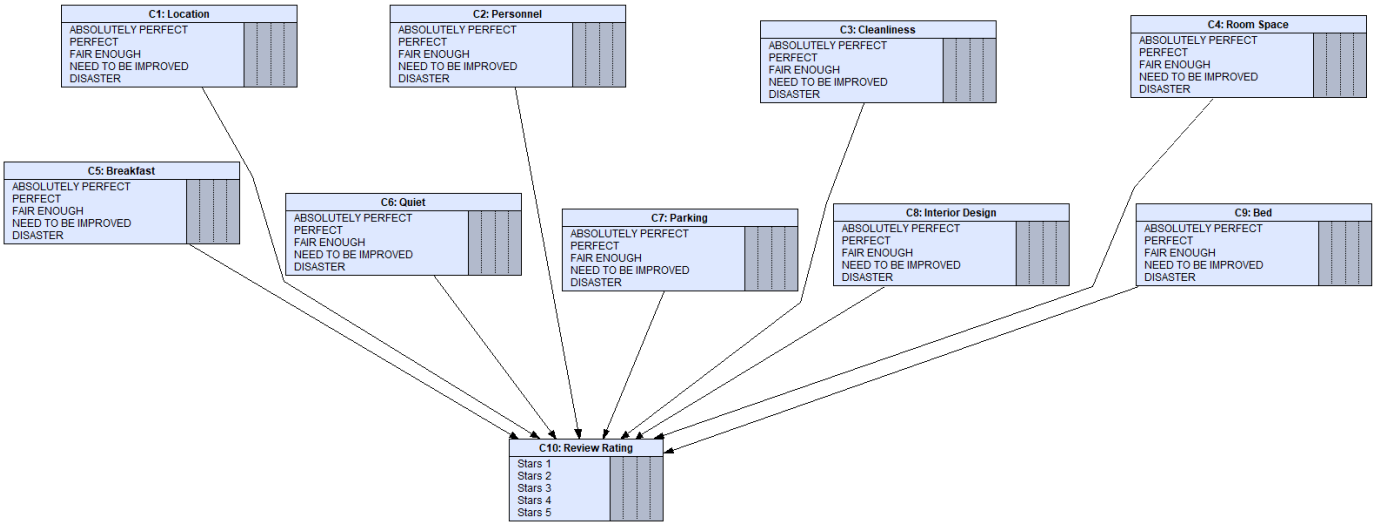


Figure 1. The Bayesian Network structure for assessing the impact of hotel services to the customer rating.

been mined during the model training. Therefore, when new relationships, not previously identified, are added, they may not be classified correctly.

TABLE III. CONFUSION MATRIX

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positive (TP) = 4125	False Negative (FN) = 905
Actual Negative	False Positive (FP) = 307	True Negative (TN) = 3231

V. CONCLUSIONS AND FUTURE WORK

Hotel quality evaluation from online reviews is an emerging research field, while the use of data analytics and machine learning methods are able to exploit its full potential in an e-tourism context. This paper proposed an approach for hotel quality evaluation according to online review comments and ratings using FSM for mining customers' opinions and BBN for evaluating the attributes that contribute to the review rating. The results show that the proposed approach is able to model the complex dynamics of online hotel review data, which are derived from both the textual nature of the review comments and the uncertain relationships between these comments and the review rating.

Regarding our future work, we plan to apply our methodology to further datasets, i.e. from different e-tourism platforms with different data structure and availability, and to investigate the role of user profiling in hotel selection. Moreover, we plan to investigate and develop approaches for detecting the fake reviews in order to increase the accuracy and the reliability of the sentiment analysis methods and algorithms.

REFERENCES

[1] X. Wei, X. Luo, Q. Li, J. Zhang, and Z. Xu, "Online comment-based hotel quality automatic assessment using improved fuzzy comprehensive evaluation and fuzzy cognitive map," IEEE Trans. on Fuzzy Sys., vol. 23, no. 1, pp. 72-84, 2015.

[2] M. J. Sánchez-Franco, A. Navarro-García, and F. J. Rondán-Cataluña, "A naive Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services.," J. of Bus. Res., vol. 101, pp. 499-506, 2019.

[3] P. De Pelsmacker, S. Van Tilburg, and C. Holthof, "Digital marketing strategies, online reviews and hotel performance," Int. J. of Hosp. Man., vol. 72, pp. 47-55, 2018.

[4] Q. Ye, R. Law, B. Gu, and W. Chen, "The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings," Comp. in Hum. Behav., vol. 27, no. 2, pp. 634-639, 2011.

[5] P. Duverger, "Curvilinear effects of user-generated content on hotels' market share: a dynamic panel-data analysis," J. of Trav. Res., vol. 52, no. 4, pp. 465-478, 2013.

[6] K. L. Xie, Z. Zhang, and Z. Zhang, "The business value of online consumer reviews and management response to hotel performance," Int. J. of Hosp. Man., vol. 43, pp. 1-12, 2014.

[7] A. S. Cantallops and F. Salvi, "New consumer behavior: A review of research on eWOM and hotels," Int. J. of Hosp. Man., vol. 36, pp. 41-51, 2014.

[8] Y. Zhao, X. Xu, and M. Wang, "Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews," Int. J. of Hosp. Man., vol. 76, pp. 111-121, 2019.

[9] Z. Xiang, Q. Du, Y. Ma, and W. Fan, "A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism," Tour. Man., vol. 58, pp. 51-65, 2017.

[10] K. Berezina, A. Bilgihan, C. Cobanoglu, and F. Okumus, "Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews," J. of Hosp. Mark. & Man., vol. 25, no. 1, pp. 1-24, 2016.

[11] L. Zhou, S. Ye, P. L. Pearce, and M. Y. Wu, "Refreshing hotel satisfaction studies by reconfiguring customer review data," Int. J. of Hosp. Man., vol. 38, pp. 1-10, 2014.

[12] X. R. Zhao, L. Wang, X. Guo, and R. Law, "The influence of online reviews to online hotel booking intentions," Int. J. of Cont. Hosp. Man., vol. 27, no. 6, pp. 1343-1364, 2015.

[13] E. Boon, M. Bonera, and A. Bigi, "Measuring hotel service quality from online consumer reviews: A proposed method,"

 ISBN: 978-1-61208-884-6

in Information and Communication Technologies in Tourism, pp. 367-379, Springer, Cham, 2014.

[14] W. G. Kim and S. A. Park, "Social media review rating versus traditional customer satisfaction," Int. J. of Cont. Hosp. Man., vol. 29, no. 2, pp. 784-802, 2017.

[15] J. M. Benitez, J. C. Martín, and C. Román, "Using fuzzy number for measuring quality of service in the hotel industry.," Tour. Man., vol. 28, no. 2, pp. 544-555, 2007.

[16] Y. Yu, W. S. Chen, and M. Li, "Fuzzy comprehensive evaluation methods of hotel product quality," Tour. Sc., vol. 5, 2008.

[17] C. C. Yang, Y. T. Jou, and L. Y. Cheng, "Using integrated quality assessment for hotel service quality," Qual. & Quant., vol. 45, no. 2, pp. 349-364, 2011.

[18] M. Geetha, P. Singha, and S. Sinha, "Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis," Tour. Man., vol. 61, pp. 43-54, 2017.

[19] W. He, X. Tian, R. Tao, W. Zhang, G. Yan, and V. Akula, "Application of social media analytics: a case of analyzing online hotel reviews," Onl. Inf. Rev., vol. 41, no. 7, pp. 921-935, 2017.

[20] K. L. Xie, K. K. F. So, and W. Wang, "Joint effects of management responses and online reviews on hotel financial performance: A data-analytics approach," Int. J. of Hosp. Man., vol. 62, pp. 101-110, 2017.

[21] P. Figini, L. Vici, and G. Viglia, "A comparison of hotel ratings between verified and non-verified online review platforms," Int. J. of Cul., Tour. and Hosp. Res., vol. 14, no. 2, pp. 157-171, 2020.

[22] K. L. Xie, C. Chen, and S. Wu, "Online consumer review factors affecting offline hotel popularity: evidence from tripadvisor," J. of Trav. & Tour. Mark., vol. 33, no. 2, pp. 211-223, 2016.

[23] H. Li, Q. Ye, and R. Law, "Determinants of customer satisfaction in the hotel industry: An application of online review analysis," Asia Pac. J. of Tour. Res., vol. 18, no. 7, pp. 784-802, 2013.

[24] X. R. Zhao, L. Wang, X. Guo, and R. Law, "The influence of online reviews to online hotel booking intentions," Int. J. of Cont. Hosp. Man., vol. 27, no. 6, pp. 1343-1364, 2015.

[25] P. Ye and B. Yu, "Customer Satisfaction Attribution Analysis of Hotel Online Reviews Based on Qualitative Research Methods," in Proceedings of the 2nd International Conference on E-Education, E-Business and E-Technology, pp. 93-98, 2018.

[26] T. Radojevic, N. Stanisic, and N. Stanic, "Inside the rating scores: a multilevel analysis of the factors influencing customer satisfaction in the hotel industry," Cornell Hosp. Quart., vol. 58, no. 2, pp. 134-164, 2017.

[27] R. Nunkoo, V. Teeroovengadum, C. M. Ringle, and V. Sunnassee, "Service quality and customer satisfaction: The moderating effects of hotel star rating," Int. J. of Hosp. Man., 102414, in press, 2019.

[28] M. Schuckert, S. Liang, R. Law, and W. Sun, "How do domestic and international high-end hotel brands receive and manage customer feedback?," Int. J. of Hosp. Man., vol. 77, pp. 528-537, 2019.

[29] C. H. Ku, Y. C. Chang, Y. Wang, C. H. Chen, and S. H. Hsiao, "Artificial Intelligence and Visual Analytics: A Deep-Learning Approach to Analyze Hotel Reviews & Responses," in Proceedings of the 52nd Hawaii International Conference on System Sciences, 2019.

[30] D. Dubois, H. Prade, and C. Testemale, "Weighted fuzzy pattern matching," Fuz. Sets and Sys., vol. 28, no. 3, pp. 313-331, 1988.

[31] C. L. Forgy, "Rete: A fast algorithm for the many pattern/many object pattern match problem," in Readings in Artificial Intelligence and Databases, pp. 547-559, Morgan Kaufmann, 1989.

[32] J. Pearl, "Probabilistic reasoning in intelligent systems: networks of plausible inference," Elsevier, 2014.

 ISBN: 978-1-61208-884-6