International Journal on Advances in Internet Technology

2019 vol. 12 nr. 3&4

The *International Journal on Advances in Internet Technology* is published by IARIA.
ISSN: 1942-2652
journals site: http://www.iariajournals.org
contact: petre@iaria.org

International Journal on Advances in Internet Technology
Volume 12, Number 3 & 4, 2019

International Journal on Advances in Internet Technology
Volume 12, Numbers 3 & 4, 2019

CONTENTS

pages: 61 - 75
Blockchain and Image Processing to Reinforce Provenance in the Narrative of a Handwoven Carpet
Amirhossein Roshanzamir, Faculty of Management, Law and Social Sciences, University of Bradford, UK
Vishanth Weerakkhody, Faculty of Management, Law and Social Sciences, University of Bradford, UK
Nripendra Rana, Faculty of Management, Law and Social Sciences, University of Bradford, UK
Mohammad Rahmati, Department of Computer Engineering, Amirkabir University of Technology, Iran
Mehdi Shajari, Department of Computer Engineering, Amirkabir University of Technology, Iran

pages: 76 - 85
Web-based Visualization of Daily Mobility Patterns in R
Antje von Schmidt, German Aerospace Center (DLR) - Institute of Transport Research, Germany
Rita Cyganski, German Aerospace Center (DLR) - Institute of Transport Research, Germany
Matthias Heinrichs, German Aerospace Center (DLR) - Institute of Transport Research, Germany

pages: 86 - 95
How Social Media Factors Influence User's Travel Purchase Intention
Stavros Kaperonis, Panteion University of Social and Political Sciences, Greece

pages: 96 - 108
Digital Adoption Strategy: A Public Sector Ecosystem
Samantha Papavasiliou, University of Adelaide, Australia
Carmen Reaiche, University of Adelaide, Australia

pages: 109 - 121
Smart Shopping Cart Learning Agents
Dilyana Budakova, Technical University of Sofia, Branch Plovdiv, Bulgaria
Lyudmil Dakovski, European Polytechnic University, Bulgaria
Veselka Petrova-Dimitrova, Technical University of Sofia, Branch Plovdiv, Bulgaria

pages: 122 - 132
Performance Catalogs for Just-In-Time Delivery of Web-based Natural Language Processing Services
Soheila Sahami, University of Leipzig, Germany
Thomas Eckart, University of Leipzig, Germany

Blockchain and Image Processing to Reinforce Provenance in the Narrative of a Handwoven Carpet

Amirhossein Roshanzamir[a], Vishanth Weerakkody[a], Nripendra Rana[a], Mohammad Rahmati[b], Mehdi Shajari[b]
[a]Faculty of Management, Law and Social Sciences, University of Bradford
Bradford, UK
[a]{A.H.Roshanzamir, V.Weerakkody, N.P.Rana}@bradford.ac.uk
[b]Department of Computer Engineering, Amirkabir University of Technology
Tehran, Iran
[b]{Rahmati, Mshajari}@aut.ac.ir

***Abstract*—The sharing economy, which is empowered by digital technologies, drives platform business models and creates new forms of growth and value, which disrupts many businesses. The essential challenge for the formation of business interactions between individuals in the context of platforms is mutual trust. Blockchain, with broad applications in different domains, is an emerging technology to build trust and improve transparency and traceability within the sharing economy and platform business models. The purpose of this paper is to explore the role of blockchain in disrupting the luxury industry and leveraging storytelling to create an immutable digital identity. The paper will review the business model of Chamrosh Technology, which transforms a handwoven carpet into the smart carpet. Within the model, blockchain is reinforcing provenance in the narrative of a handwoven carpet, and image processing is generating a digital fingerprint to deal with counterfeits.**

Keywords–Smart Carpet; Startup; Storytelling; Handwoven carpet; Blockchain; IoT; Identity management; Image processing; Carpet fingerprinting

I. INTRODUCTION

This paper is an extension of the work initially presented in the Eighth International Conference on Ambient Computing, Applications, Services and Technologies [1] which discussed the role of digital technologies in transforming handwoven carpet to the smart carpet and co-creating value with carpet fans to influence their behavior.

The rise of the sharing economy, which is empowered by innovation and growth in digital technologies, drives online peer-to-peer platforms where individuals can leverage their existing and under-utilized assets to create value together [1]. The essential challenge for the formation of interactions between individuals is mutual trust [2]-[5]. Many businesses and industries are being transformed by these digitized, open, and participative platforms, which act as a Trusted Third Party (TTP) and commercially connect suppliers and consumers. Through these platforms, companies such as Alibaba, Uber, and Airbnb are disrupting retail, transportation, and accommodation industries, respectively. Disruption is associated in part with “Disruptive Innovation Theory” coined by Clayton Christensen [6] and refers to the fact that even successful companies are at high risk that their product, system, or technology become obsolete. Disruption usually happens when a newcomer displaces the value proposition of the incumbent by offering a cheaper and more accessible solution and later poses a barrier that would prevent the incumbent from quickly imitating a similar model [7].

The luxury goods industry has also experienced disruption by innovation in the sharing economy, which is driven by digital technologies. On the one hand, the sharing economy has displaced the value proposition of luxury brands by providing access to goods and services through online platforms [1]. On the other hand, the entire value chain of luxury goods from procuring raw materials, design, and production up to in-store experience, and even consumption is personalized and transformed by digital technologies [1].

Blockchain with broad applications in different industries can improve transparency and traceability within the supply chain of luxury goods and verify authenticity. A blockchain is a data structure which combines multiple technologies including cryptography and peer-to-peer networks to create a digital and decentralized ledger of data and share it among a network of independent parties [2][3][8]. Decentralized system and cryptography enable each participant to manage the ledger securely without the need for a central authority to enforce the rules [2][8]. Blockchain can, therefore, generate trust for the entities who do not know each by building a secured medium in which data are protected from unauthorized access. That is why blockchain is proliferating in different sectors especially where identity management and verification are required

This study explores the emerging role of blockchain to disrupt the luxury Industry and further discusses how to reinforce the provenance of a handwoven carpet and deal with counterfeiting. The rest of this paper is organized as follows. A brief overview of the smart carpet and background of blockchain is presented in Section II. The four elements of storytelling are presented in Section III where it is also discussed how they can be reinforced by digital technologies. In Section IV, the smart carpet business model is reviewed to show how blockchain can reinforce the provenance and leverage the power of storytelling to prove the origin and verify the authenticity of a handwoven carpet. The collaboration of two startups is discussed in Section V. The challenges of the smart carpet model are identified in Section VI where image processing is reviewed as an innovative solution to generate a digital fingerprint to deal with counterfeit carpets. Discussion is presented in Section VII which is followed by the conclusion and future works in Section VIII. Contribution is summarized in Section IX.

II. BACKGROUND

The concept of smart carpet was first introduced in a research project which aimed at developing a sensor system to detect falls and summon assistance for older adults who cannot use wearable devices due to cognitive deficits [9]. Later, other applications of a smart carpet such as physical therapy emerged where the walking pattern of the owner is detected and analyzed. With the help of the therapist, it is argued that these patterns can predict mobility problems and correct them [10]. Also, it is shown that the smart carpet can identify the presence of an intruder by acting as a kind of alarm system which detects environmental threats, like fires [10].

Roshanzamir et al. [1], for the first time, discuss how the Internet of Things (IoT), blockchain, and platform business can offer a game-changing solution and disrupt handwoven carpet business by bringing transparency and trust to the supply chain, dealing with counterfeiting and increasing customer engagement. They suggested an extension built upon the EIC model [11] in which it is argued that digital technologies can improve operations and contribute to the marketing campaign and create strong brands by generating an immutable digital certificate [1].

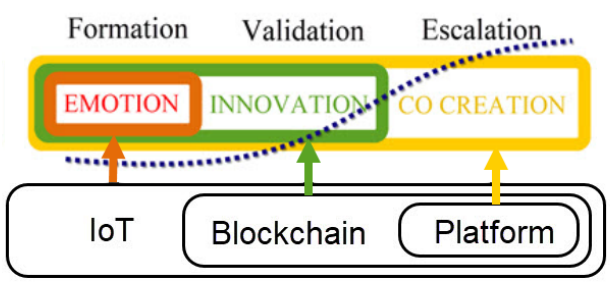


Figure 1. The practical extension of the EIC model [1]

The framework breaks down the narrative of a carpet from the time procuring of knots and weaving starts in a simple connection of cause and effect. The blockchain globally stores, and collaboratively writes a list of all transactions that have ever taken place within a given system [12] and offers the possibility to keep the whole story of a carpet in a decentralized system.

Blockchain technology can, therefore, reinforce trust between individuals because it is a decentralized database, without a central repository and usually without a central authority. Blockchain functions as an immutable digital ledger and generates the list of records that are bundled together in so-called blocks. The ledger is distributed across many participants in a peer-to-peer network as a block that is duplicated thousands of times across a network of computers. Each block has three major components that are: data including facts or transactions plus timestamps, a hash of the previous block, and a hash of the existing block to be added in the next block [13][14].

A hash is created by using a mathematical function or algorithm to convert the input of letters and numbers into an encrypted output of a fixed length. Processing the hash functions to encrypt new blocks or decrypt previous blocks requires substantial computer processing power, which is very costly.

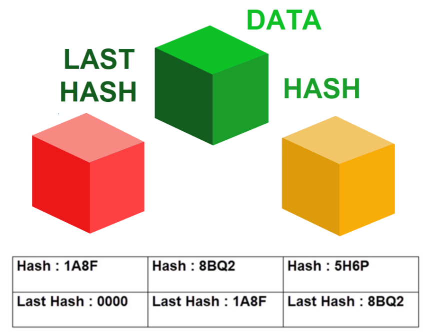


Figure 2. Connecting blocks and checking the hash function

Types of transactions and rules, which can be carried out, must be agreed between participants in advance and stored in the blockchain as a smart contract which was introduced in 1997 by Nick Szabo [15] to ensure that everyone is fulfilling the rules. The main components of the smart contract are a set of executable functions and state variables. Each transaction has some input parameters which are required by a function in the contract. During the execution of a function, the status of the state variables is changed depending on the logic implementation. The new blocks are constantly added and updated to the chain, and data in any given block can only be altered by modifying all previous blocks, which would require the control over a majority of computing power in the network. The benefits and innovation of blockchain, described in below key characteristics from a combination of four technologies: a distributed data storage; peer-to-peer networks; cryptographic algorithms; and decentralized consensus mechanism [16]-[18]:

- Distributed ledger: Blocks are recorded, duplicated, shared and synchronized on independent computers (referred to as nodes) with no centralized data storage. The indistinguishable copies available on different nodes allow participants to validate information without a central authority and make it failure resistant, even if few nodes fail.
- Immutable: Blocks are time-stamped, and each block is connected to the one before and after it within the same transaction in a growing chain structure. This creates an irreversible, immutable chain in which the trail and history of blocks and transactions cannot be modified.
- Secure: Blocks are permanently recorded and encrypted in a growing chain, which makes them impossible to delete, edit, or even copy already created block and add them in the network. This ensures a high level of robustness and trust.
- Consensus mechanism: A transaction on the blockchain can be executed once the parties on the network unanimously approve it. This stop illegal and bad transactions and blocks.

Blockchain can, therefore, generate trust for the entities who do not know each other by making data and the relevant digital identity tough to change due to decentralized features and very hard to read due to cryptography (hash function). That is why blockchain has been successfully expanding its application across different domains including investment, logistics, global organizations, and even governments where identity management and verification are required. More importantly, blockchain can provide emerging opportunities in the platform business model since it facilitates real-time granular visibility, ensures trust, and enforces security using a chronological order of smart contracts. These contracts can automatically implement a binding agreement between two or more parties, where every entity must fulfill their obligations according to the agreement. In the next section, we will explore the role of storytelling with an implication in the luxury industry and how these stories can be verified through the blockchain.

III. AUTHENTICATED STORYTELLING IN THE LUXURY INDUSTRY

Storytelling is defined as ‘‘sharing of knowledge and experiences through narrative and anecdotes to communicate lessons, complex ideas, concepts, and causal connections’’ [19]. It is one of the oldest and most powerful modes of communication and has been an integral part of cultures throughout history, both as entertainment and as a means of passing on knowledge, values, and desired behaviors from generation to generation [20]. The fact that people naturally organize their experiences through the construction of stories has an essential implication in marketing and branding. In fact, stories are the pillars of Word of mouth (WOM) communication, and a significant dimension of brands and advertising needs to support them [21]. McKee [22] argues that stories are effective at persuasion because they emotionally involve audiences as many people are interested to know more than just the product or service by observing behind the scenes stories of organizations according to market research [23]. That is why the appropriate way to convince someone to purchase something is by telling a compelling brand-oriented story rather than providing rational arguments, statistics or facts. For example, Woodside et al. [19] consider storytelling through blogs as a more effective way of driving purchase intentions than traditional websites. These stories can provide a strong point of differentiation that cannot be copied [23] [24], and that fulfills the purpose of a brand to differentiate goods or services of one seller from a competitor, according to the American Marketing Association’s definition. In fact, a brand is what people feel about us and our products and services. Therefore, it is partly rational but mostly emotional. Nonetheless, our brand and story, however visionary and brilliant, carry no weight, unless we can win the hearts and minds of customers and deeply motivate them on a personal level to act not because they need to, but because they want to. We must be mindful that at the heart of a strong brand are great products and services, but every brand contact matters [24] and storytelling is the most important one [25]. Therefore, organizations need to build a strong brand story to communicate with their potential and existing customers and regularly reveal the compelling benefits of using our ideas, products, or services.

Therefore, we can argue that stories are considered to be powerful tools to structure reality, facts, conclusions, and recommendations about a brand in a persuasive manner that play to the emotions and rely upon empathy. Literature review and different studies have identified four elements of good brand stories [26], and we will discuss here how two elements including authenticity and humor, can be verified and reinforced by integrating technologies including the Internet of Things (IoT), blockchain, and platform business. The goal should be to cultivate an authentic, trustworthy, and compelling narrative to inform, engage, persuade, and build trust with customers.

A. Authenticity

Authenticity is a sense which readers or audiences obtain that makes them believe and associate the story with reality. In fact, the authenticity of narratives, which should tell real and incredible stories, is a critical issue since how they reflect the truth within an organization can be questioned [27]. The corporate reputation is defined as a stakeholder’s perception of the organization. Brown et al. [28], and Dowling [29] suggest that if the story causes stakeholders to perceive the organization as more authentic, distinctive, expert, sincere, powerful, and likable, then it is likely that the overall corporate reputation will be enhanced. Recent studies show that digital technologies, including blockchain, can provide promising possibilities for authentication and verification of corporate stories. For example, Maxwell et al. [12] argue that digital technology has transformed the ways we create and consume narratives, from moving images and immersive story worlds to digital long-form and multi-branched story experiences.

B. Conciseness

A story should deliver complete thoughts in a few words, while still covering essential points. This is sometimes referred to as plot, which is a sequence of events starting from anticipation and then progresses towards crisis, getting help, and finally achieving a goal [24]. The story must also be presented as something inherently valuable where the value comes from the story itself and from the role it is playing for a broader significance that is promoting a brand. In this manner, we try to give the events a meaning that our potential and existing customers may not otherwise grasp.

C. Reversal

A climax is a turning point when the emotion takes a surprising twist. Corporate stories bring an emotional dimension to an organization and enable people to make a personal connection to learn, understand, and share information and ideas [23]-[25]. Here participating in our customer’s transformation can give new life and meaning to our business suggested by Miller [30]. He argues that we need to think about who we want our customers to become and how we can improve the way they see themselves [30].

D. Humour

A statement, picture or other things which call for action or give rise to a kind of emotions. Peter McGraw [31] posited the “Benign Violation Theory,” which states that humor occurs when a situation is both benign and a violation of moral or social norms, or some other expectation. That is how humor can challenge expectations, make connections between conflicting ideas or emotions to surprise the audiences.

Also, it should be noted that we have moved from mere consumers of content in web 1.0 towards increasing our participation and documenting our activities in social media, which is referred to Web 2.0 as proposed by Tim O’Reilly [32][33]. Unlike Web 2.0, which refers to the participation of users, Web 3.0 is based on users’ cooperation [34]-[36], which integrates the generated data to create new meaning. In fact, the concept of Web 3.0 is often associated with the idea of the Semantic Web which was first coined in 1999 by Tim Berners-Lee, the creator of the World Wide Web, who foresaw the possibility of enabling machines to ‘talk to one another’ and to understand and create meaning from semantic data [37] [38]. The blockchain is a technology that can reinforce semantic data by building up trust between the parties and ensuring data integrity. That is why blockchain empowers Web 3.0 to have a user-centric transparent and secured internet, which is based on a decentralized network. Maxwell et al. [12] conduct three experiments that draw on some of the fundamental principles of blockchain including the ledger, the blocks and the mining process and argue that blockchain opens up new possibilities to explore how storytelling might adapt as distributed ledger technologies become part of how we read, write and share stories. They further argue that blockchain could significantly transform the distribution, promotion, and propagation of stories, especially with regards to data archival, originated from time-stamping principles. Managing collaboration, contribution, and attribution, as well as ethical and moral issues, are other implications of generating narrative through blockchain as they discuss.

In Section IV, we are going to review the case study of a new venture who integrates blockchain technology to offer a solution in one of the most traditional businesses i.e., handwoven carpet.

IV. SMART CARPET BUSINESS MODEL

Founded in early 2018, Chamrosh Technology [39] is a startup in the Middle East, who aims to transform a handwoven carpet to the smart carpets through digital technologies. It is an Intelligent System of identifying, recording, and verifying the provenance of a handwoven carpet by generating the fingerprint and issuing and storing a digital certificate on the blockchain with data-backed evidence. Chamrosh integrates digital technologies including the Internet of Things (IoT), image processing, steganography, blockchain, asymmetric encryption and mobile technology to issue an irrevocable and immutable digital certificate for a handwoven carpet and store it on the blockchain.

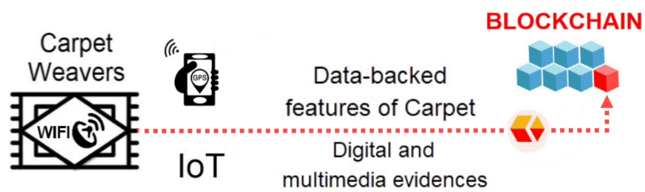


Figure 3. The narrative of carpet is stored on the blockchain [1]

The narrative of a handwoven carpet begins by collecting data-backed features as well as multimedia evidence from the weaving process. Also, a smart chip is embedded inside the carpet from the beginning with a low energy wireless IP network via the Internet of Things (IoT) or Bluetooth to integrate with Global Positioning System (GPS). This is achieved through a mobile app that pings out to the GPS every month to fetch the location of the carpet weaving and communicate it through the smartphone of the weaver(s) with the system as illustrated in Figure 3. Data-backed features of carpet including original map, designer’s

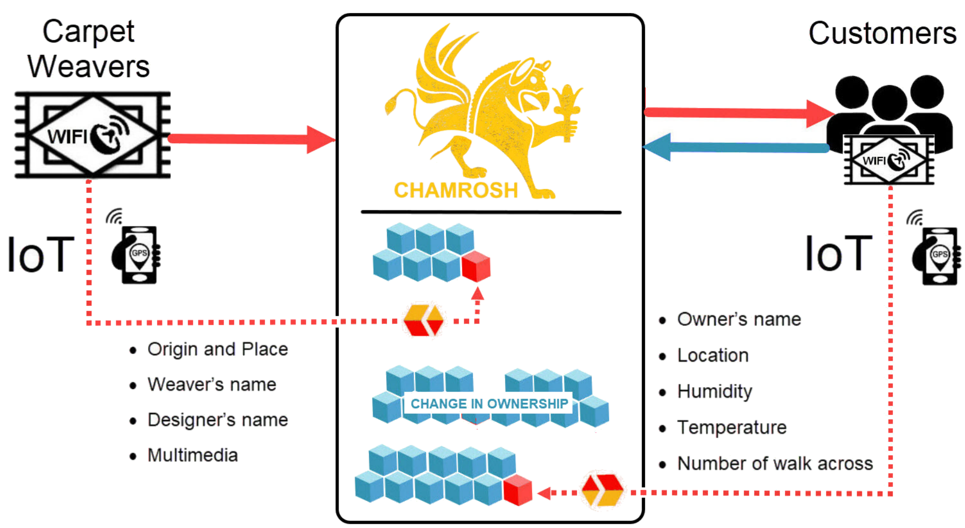


Figure 4. Chamrosh business model can reinforce the provenance by immutable digital certificate [1]

name, size, number of knots and colors, etc. together with audio, video, and pictures from the weaving process are stored in story blocks as successive stages. Therefore, the potential customers feel what the weaver feels and see what the weaver sees [1]. Most importantly, the ownership history and other environmental data, including temperature, moisture, location of use, and even the number of times carpet was walked over, can be recorded as new blocks and be verified later, as shown in Figure 4. Here, we see a paradigm shift in story-giving where the customers can actively contribute to composing the story of the brand before selling the carpet [1] which makes the product part of a bigger story and more meaningful to the customers' lives [40]. Customers can both buy and sell their smart carpets within the platform. This is highlighted by blue and red arrows within Figure 4. Also, the owners of a smart carpet can use the Chamrosh platform to sell his or her carpet, and the new customer can verify the ownership, and that is how the legal norms and ownership rights are enforced without the need for a third-party authority. This resonates with recent arguments that the hybrid institution of property is a distributed ledger that can hold information about the intellectual property of right holders instead of a centralized government database [41].

Kapferer [42] argues that a luxury strategy places a high priority on localized production to support the brand story and increase intrinsic value. That is why the country of origin symbolizes expertise and cache. For example, Swiss watches and German cars are endowed with unique local know-how and the magic spirit of place of production. The same argument holds true for Persian handwoven carpets or Persian rugs which were first referenced around 400 BC, by the Greek author Xenophon in his book "Anabasis". He described them as precious, and worthy of being used as diplomatic gifts which put them into a context of luxury and prestige[1]. "Carpet" and "Rug" are terms that are used interchangeably. In the US, a carpet is a floor covering that is laid wall to wall, and rugs cover a small specific area. But in business "carpet" and "rug" are used to denote a covering over the floor.

In fact, Persian rugs are the singular and invaluable symbol from one of the most ancient civilizations in the world. Persian rugs carries more than 2500 years of heritage, culture, and art that covers stories of failures, victories, peace, passion, prosperity, and love and that is how two pieces of rug can be made of the same material within the same size, and even have a similar design, yet, Persian origin can reel the price by 5 to 10 times [1]. That is why the Persian rugs have always been subjected to counterfeiting due to their supremacy in all aspects over handwoven carpets from other countries.

Adored for their intricate designs, elegant colors, and perfect craftsmanship as well as intrinsic value, Persian rugs have found their way into households, catwalks, and concert stages, as well as in scores of artworks [43]. Persian rugs also become important ingredients of the luxury ambiance in many premium locations such as China Room on the Ground Floor of the White House. Despite advancements in technology used to produce machine-made carpets, the handwoven carpets or rugs have retained their value as an attractive and masterful product.

Persian handwoven carpets are exported to about 80 countries, and the exported value is estimated to be $500B per year, according to the Persian Carpet Research Center [44]-[46]. The tallies, however, do not cover the potential value of the millions of carpets woven and remaining in Iran as well as sold before throughout the world, since, it is nearly impossible to track or quantify them [44]. Most of the designs are named after the village or area that is expert at a unique pattern or style of weaving such as Tabriz, Kashan, Naeen, Isfahan, Qom, and Kerman. In addition to data-backed evidence of carpet including origin, place, pattern, map, designer name and other physical features like size, the number of colors, weaving style. Every handwoven carpet can carry provenance and much trustworthy information about the owner and environmental parameters that form the ingredients of a successive story and potentially increase the intrinsic value of a carpet. For example, those carpets which are placed at the holy sites of Islam, Christianity, and other religions are traded at premium prices because the potential buyers assumed them to be blessed and cherished by the location of use [1]. The smart chip which pings out to the GPS offers an innovative solution to trade these used carpets by verifying the location. This positively influences the purchase intentions and willingness of customers to pay a premium [1].

V. CHAMROSH AND BLOCKTAC

BlockTac [47] is a Spanish startup who offers solutions in the Identity Management and Authentication field for various industries through in house technology, which is built on the integration of blockchain, cloud, and mobile app. Founded in 2018, BlockTac has developed a cost-effective technology in the educational sector which empowers Universities, Business Schools and other educational institutes to register the certificate of their students on the blockchain. BlockTac further provides a mobile platform to enable the third party to verify the issued certificate. The technology is implemented in five simple steps as illustrated in Figure 5.

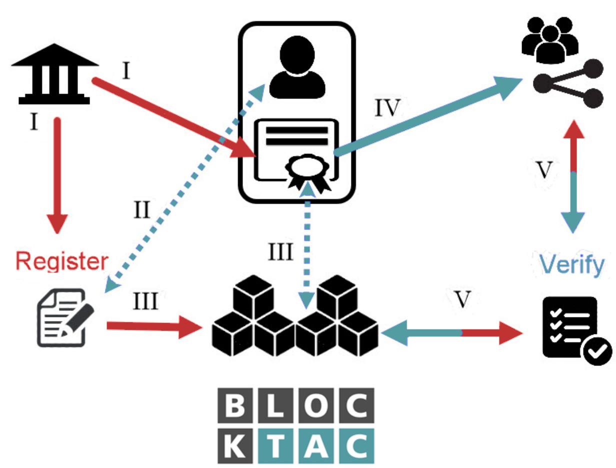


Figure 5. BlockTac methodology to verify certificates

First, the University or Educational Institute who has the record of registered students will issue the certificate for those who fulfilled the requirements and then communicates the details to BlockTac through a private and secure channel. Second, BlockTac notifies the student and facilitates an Android or iOS application to view the certificate and get the student's confirmation on personal and other details. Third, BlockTac registers the certificate on the Blockchain through in-house technology and sends the blockchain certificate with a relevant link to the student. Forth, the student can share the certificate with whomever he or she wishes. Fifth, any person or third party can verify the authenticity of the certificate without having to contact the University or Institute.

Relying on Blockchain technology to provide the properties of inviolability, immutability and open verification for all digital certificates, BlockTac is now diversifying the technology and integrating it with the value chain of industries such as luxury, healthcare, pharmaceutical, insurance, supply chain, food, and logistics as well as government administration and financial sectors. For example, BlockTac has developed a digital seal for food products, including oil, wine, and saffron as well as medicines and perfumes, to battle counterfeiters. In the food supply case, for instance, all information is digitally connected to a particular food including farm-origination, processing protocol, factory information, storage, and transportation details, most importantly expiration dates are collected, agreed and stored in a digital certificate which becomes a permanent record that cannot be changed. The records are approved by consensus among participants and then verified through the mobile app by consumers.

The solution works as follows: in the first step, the seal of a bottle or package will be opened, and the mobile scans the Quick Response (QR) trackable code. The digital certificate in the blockchain is just valid for one-time verification. Upon scanning, the system verifies the credentials and confirms that the package is original and the seal has never been opened. In the second step, if we rescan the code or try to use it on another product or package, the system disproves the credentials and indicates that the package was opened before.

Figure 6. Digital seal for food products by BlockTac [47]

In most industries and sectors, including investment, logistics, and even governments, the issue of identity management, traceability and authentication are critical. BlockTac technology would contribute to reinforcing identity via blockchain and enabling verification through a mobile application without Trusted Third Party (TTP).

BlockTac and Chamrosh are now collaborating on a well thought strategic partnership to forge a technology to transform handwoven carpets to smart carpets. Synergizing each other's core competencies, both startups are expecting to seize evolving opportunities by leveraging mutual resources and expanding their geographical reach in the digital era where most companies become increasingly interdependent. The technology is developed by using a smart chip and seal containing QR code, embedded in a handwoven carpet, to store the data-backed features of carpet on the blockchain system as successive story blocks. The process is shown in Figure 7 when data to be collected through a protocol every month and recorded in an individual block basis where the last block will contain the whole story. The immutable history of ownership will also be recorded, and once the real owner verifies a purchase, as an update, the system automatically forms a new block of ownership with all other features for the next owner.

The sharing economy can also provide an emerging opportunity to rent or lease luxury handwoven carpet on a daily or weekly basis for special occasions including weddings, and anniversaries. Blockchain and IoT can help to trace this luxury item during the rental period.

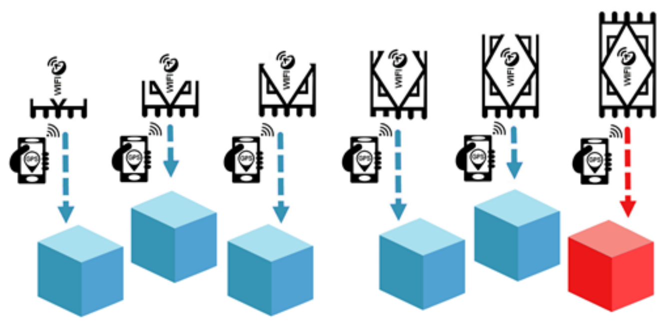


Figure 7. A narrative of weaving to be stored in story blocks

Powered by blockchain, the smart carpet business model can offer significant benefits for carpet weavers, dealers and customers, by improving efficiency, transparency, and traceability of supply-chain. The platform gathers, shares and verifies key information and evidences about a woven carpet by integrating secure, trustworthy and accessible technologies. We can enumerate the benefits as indicated in Figure 8.

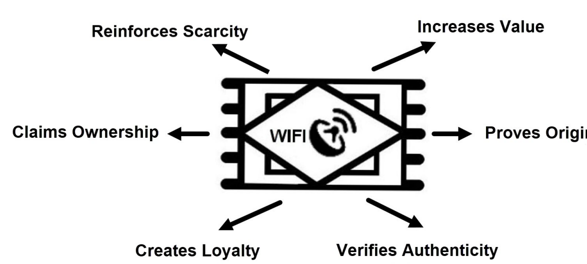


Figure 8. Smart carpet can offer six significant benefits [39]

First, through smart chip, which identifies the place of production, the system determines genuine carpets woven within Persian territory and essentially differentiates them from fake carpets that copy the design and are produced in other places. This proves the origin and craftsmanship quality. Second, the system creates an immutable and live digital identity with a mobile application to verify authenticity. This deals with counterfeiting and fake items and gives peace of mind to the customers. Third, the system generates a platform to directly interact with consumers and make them part of a weaving story. This wins the hearts and minds of Handwoven carpet's fans in a new and innovative way and creates loyalty. Forth, the system records and tracks the ownership, and this makes the carpet transferable and updatable once the status changes. Fifth, the system reinforces scarcity and exclusivity by making the item unique and irreplaceable. Sixth the system increases the value of a smart carpet since its aesthetic, historical, and sole significance and all former benefits contribute to the value.

The aforementioned business model aims to bring back confidence to handwoven carpet fans and empowers them to interact with their luxury items in the most direct, secure, and transparent way that cutting-edge technologies allow. The possibility to track ownership and update the changes once the real owner of carpet requests, provides a unique opportunity to earn money for transferring ownership to the next owner. This can justify the commercialization of technology and business model.

The prototype of the solution is shown in Figure 9 in which the credentials of a carpet is framed into a digital certificate and then stored on blockchain with Quick Response (QR) trackable code. A QR code, which is a two-dimensional code in square shape image, mostly represented by black and white pixels in a binary format and is used in consumer advertising such as web pages and posters.

Figure 9. Digital certificate of carpet stored on the blockchain [48]

QR code has fast readability, and larger storage capacity compared to barcodes and used to hold and record the credentials of carpet on the blockchain by distributing a full copy of the database to multiple computers or nodes. All participants must approve the changes, so it is incredibly challenging for individuals to tamper with the carpet's digital certificate or commit fraud.

The QR tag is attached at the back of the carpet and once an app scans it, the relevant certificate paired with the carpet credentials will appear on the mobile. The user can scroll down to verify the certificate as shown in Figure 10 by clicking on the verification button, and within a few seconds, the features of the carpet will be authenticated through blockchain.

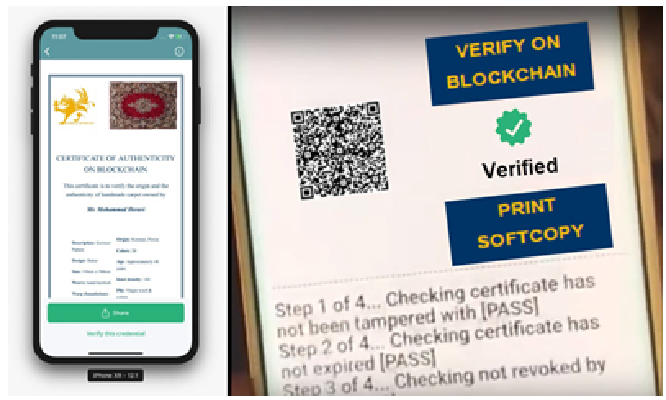


Figure 10. Mobile App verifies the credential of carpet [48]

VI. CHALLENGES AND IMPLICATIONS

One of the emerging challenges in the smart carpet business model is the fact that smart tag can be removed and then attached to a replica carpet. In these circumstances, the owner has an option to push another carpet of similar kind and lower value into the system. Obviously the owner cannot own two carpets with the same smart tag, yet, he or she can still sell the first carpet in the market and leverage the value of the second replica because of smart tag. In this section, we offer a solution to address this challenge by digital image processing and steganography.

With the advancement of Artificial Intelligence abbreviated as AI technology and the power of computers, digital image processing is now among growing technologies with various applications in business, science, and everyday life. Traditional image processing technologies mainly focus on image acquisition, transformation, enhancement, restoration, compression coding, segmentation, and edge extraction [49]. Yet, image feature analysis, image registration, image fusion, image classification, image recognition, content-based image retrieval, and image digital watermarking have made significant progress in different fields [49][50], which reflect on human intelligence activities [49] and are empowered by Artificial Intelligence. The recent and major applications of image processing include but not limited to aerospace, land mapping, urban planning, medical research and treatment, product anti-counterfeiting, surface damage identification, real-time monitoring, iris recognition, and military, cultural, artistic and communications aspects of human life and work [50].

Image processing includes two key steps or two-dimensional functions, namely sampling and quantization

[49][50], which enable a computer to recognize images. Sampling and quantization are sometimes referred to as acquisition/scanning and preprocessing in image processing literature, respectively. Ponti et al. [51] argue that the image-based visual recognition pipeline includes a step that converts colored images into images with a single channel, obtaining a color-quantized image that can be processed by feature extraction methods.

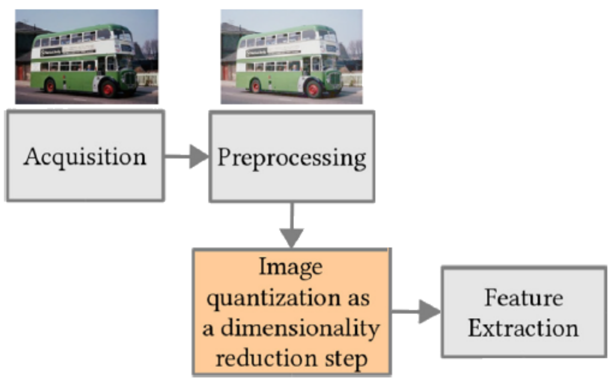


Figure 11. The pipeline of image processing [51]

Digital images are generally obtained by optical scanners or cameras. The allocation of pixels to locations in the original image is called sampling, which means each pixel represents a sample of the image from a particular place [50]. Quantization assigns numbers from 0 to 255 in color images for each pixel from 3 channels (Red, Green, Blue). Each channel has 8 bits, which contains 256 grades starting from 0 which is the darkest color to lightest that is 255. Therefore, (2^8) x (2^8) x (2^8) = 16,777,216 choices of color are available for each pixel. The next step in image processing would improve the quality of the image, which can be done via Contrast & Brightness, Filters, Threshold, Edge Detection and Contour, Sharpening and Blurring, and Noise Reduction.

A. Carpet fingerprinting by detecting human errors

It is often said that all carpets are perfectly imperfect which means that slight inconsistencies and imperfections prove that a carpet was woven by hand [48]. The fact is reflected in David Benioff and DB Weiss interview with Japan's Star Channel addressing Game of Thrones' controversial eighth season in which a coffee cup left in a scene in front of Daenerys Targaryen. David Benioff called it their "Persian rug" to justify the issue by saying, "In Persian rugs, it's a tradition that you make a little mistake when making the rug because only God can do anything perfectly." [52]

A carpet is usually woven by a reader who speaks out the color of each knot pixel by pixel from the original carpet map. The weavers, who listen to the reader, select the colors and weave them in the same sequence, which they hear. Yet, they are likely to commit a few random mistakes by weaving a wrong color specially when they are tired or lose their concentration. Other errors such as knot loss, and carpet ragging i.e., changing the width of the carpet due to the aggregation of the pulling force of the carpet weft are all exclusive and distinct features of any handwoven carpet. Therefore, we can argue that every carpet has unique and distinct features that cannot be replicated whatsoever, and that is how every carpet is perfectly imperfect. This is exactly like human fingerprints, which are detailed and unique i.e., no two people, not even identical twins, have the same fingerprint. That is why we use fingerprint term for handwoven carpets since no two handwoven carpets can have exactly the same errors.

For example, it is very much possible that in a handwoven luxury carpet of 5 Million knots, 5,000 knots are woven with the wrong color which indicates 0.1% errors. These errors are the unique characteristic of a carpet that randomly took place and are unlikely to happen in the same knot or pixel in another carpet, even if the original map, reader, and weavers are the same. These errors are considered to be the fingerprint of a carpet and are stored on the blockchain as a part of the immutable digital certificate. We can then use this innovative idea to identify a handwoven carpet by using image processing technology. The first step is to utilize a high-resolution scanner for scanning the back of the first carpet to record the woven layout in the computer. The scanned layout then must be compared to the original map through image processing algorithms to reduce noise and contrast and sharpen the brightness. For example, in Figure 12, we can observe that in pixel 6, 13, and 15, there are slight brightness errors, and the real colors are not properly sampled and quantized. That is why our image processing algorithms should enable us to correct the color according to the color of the original map.

Moreover, the weavers of the first carpet have committed few mistakes on the color of pixels 4, 18, and 30. These are originally red, green, and yellow respectively in the carpet's original map. Yet, we can see that they are wrongly woven as blue instead of red (pixel 4), yellow instead of green (pixel 18), blue instead of yellow (pixel 30).

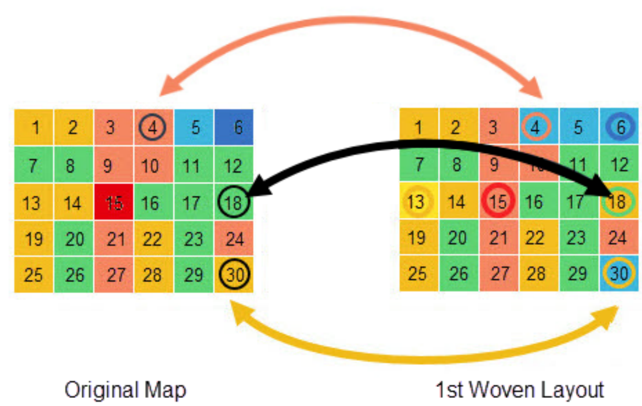


Figure 12. Comparison between original and 1st woven layout

Therefore, we identify pixels 4, 18, 30, which are mistakenly woven in different colors, as the unique feature of the first carpet and add them to the next block in the blockchain as an element of carpet's digital certificate.

TABLE I. THE ERRORS IN WEAVING PROCESS

Pixel	Original layout	1st Woven Pattern
4	Red	Blue
18	Green	Yellow
30	Yellow	Blue

Now, let us say that in a very optimistic scenario, there is a new weaving team who have access to the original map and try to weave a replica. They cannot succeed to copy exactly the same carpet simply because the errors will randomly take place in different pixels during the weaving process of the second carpet. Therefore, it is absolutely unlikely that the errors on the second carpet happen in the same pixel as the first carpet.

For example, we assume that the random errors of the second carpet happen in pixel 6, 21, 26 as indicated in Figure 13. These are originally dark blue, red and green respectively in the carpet hardcopy layout. Yet, we can see that they are wrongly woven as red instead of dark blue (pixel 6), blue instead of red (pixel 21), red instead of green (pixel 26).

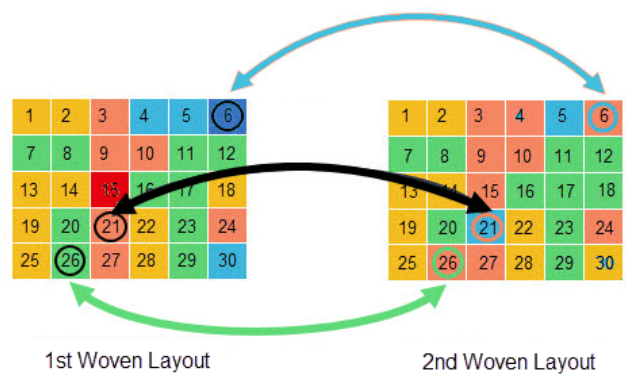


Figure 13. Comparison between 1st and 2nd woven patterns

Therefore, we can identify more dissimilarities (here six) by comparing the layout of the 1st and the 2nd carpet, which are highlighted in Table II.

TABLE II. COMPARISON OF ORIGINAL AND WOVEN PATTERNS

Pixel	Original	1st Woven	2nd Woven
4	Red	Blue	Red
18	Green	Yellow	Green
30	Yellow	Blue	Yellow
6	Dark Blue	Dark Blue	Red
21	Red	Red	Blue
26	Green	Green	Red

The theoretical assumptions discussed with examples are based on human errors during the weaving process. These errors provide an excellent identifiable characteristic for any handwoven carpet to be integrated within the digital certificate and then stored on the blockchain. These errors are as unique as human fingerprints which are detailed, very difficult to alter, and durable over the life of an individual

In addition to colors, there is a good possibility to identify symmetric and asymmetric issues within the woven layout. The original map of a handwoven carpet is based on either symmetric or asymmetric pattern. Yet, due to human errors, the woven layout does not fully match with the original map in terms of the size and dimensions of features. Therefore, we can reutilize image processing to identify the errors which took place in the pattern of the woven carpet. These errors are always unique for each and every handwoven carpet because they happen randomly.

B. Carpet fingerprinting by a hidden message

In addition to random errors of a handwoven carpet, the possibility of storing and communicating secret and/or private information can provide an invaluable solution to identify and verify the original carpet and act as the second fingerprint. This can be achieved by committing a few intended errors during the weaving process.

Information security systems are separated into two major categories i.e., encryption and information hiding [53]. Whereas cryptography keeps the data secret by converting it into an unreadable form, information hiding which is divided into watermarking and steganography does not change the format of data or message and keeps the presence of its actual data [54]. Watermarking is a process in which the information or message which verifies the owner is embedded into carriers such as images, audios, or videos.

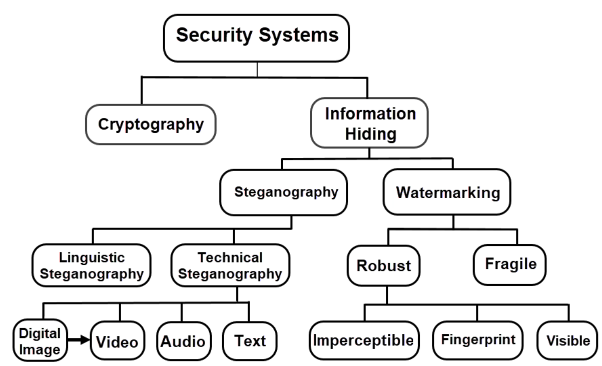


Figure 14. Classification tree of information hiding [53]

In 1993, Tirkel et al. [55] first introduced the digital watermarking system, which embeds hidden copyright marks or other verification messages into digital multimedia. When a dispute of copyright occurs, the owner can prove his ownership by revealing the copyright mark from the embedded watermark. Watermarking is of two types; visible and invisible. The embedded watermark must be robust to accidental processes, such as JPEG compression, or malicious attacks like cropping, noise adding, and rotation. Otherwise, the embedded watermark may be distorted or totally lost.

The goal of watermarking is one to many communication in which the intruder cannot remove or replace the message. Figure 15 shows a watermark in EURO banknote which is a recognizable image that appears when viewed by reflected light. Here, it is challenging to extract the watermark because the image will be distorted under some capturing process, such as scanning.

Figure 15. Watermark in Euro banknote

Steganography is another technique using digital images for secret communication where a hidden message is embedded in the image that only the sender and intended recipient can reveal. The goal of steganography is one to one communication in which the intruder cannot detect a message. Therefore, steganography can be used to reinforce the identity of a woven carpet by embedding a message such as the designer's name and time and date of design during the weaving process. Carpet's layout offers an excellent platform to embed messages because of two major reasons. First, layout can be altered to a certain extent without losing their beauty and quality. Second, humans are unable to distinguish minor changes in image color and a carpet pattern once the layout is woven. For example, we can turn 1000th to 1020th pixels or knots in the layout to colors that correspond with a letter of the alphabet or a number. While the carpet appearance would not appear exceedingly distorted or perhaps slightly corrupted at worst, we can decrypt colors and find the message.

Figure 16. Steganography embeds a secret message inside the patttern

Steganography, as discussed, can embed few intended errors during the weaving process that corresponds to a secret message like date or place which is hidden in the pattern. This method ultimately produces another unique feature like a fingerprint that can be stored on the blockchain. As such, we can spot a counterfeit carpet that does not carry the embedded message within the pattern.

To sum up, we can integrate both solutions, i.e., human errors and a hidden message to generate an integrated digital fingerprint for every handwoven carpet and store it on the blockchain.

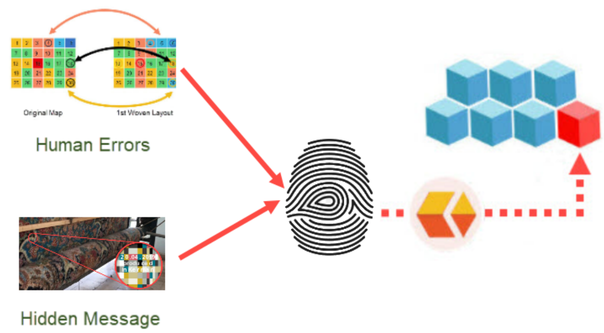


Figure 17. Carpet fingerprinting by integrating errors and hidden message

This digital fingerprint builds a robust and immutable connection between physical carpet and digital carpet. Even if the QR code is lost from the back of carpet, the digital fingerprint (random human errors and secret message) will maintain this connection and we just need to take a few photos from the back of carpet and utilize the fingerprint stored on the blockchain to compare and contrast the so-called layout with the original map.

Figure 18. Digital fingerprint connects physical carpet to digital certificate

The fingerprint verifies the authenticity of the original carpet by matching the errors with the original map. If errors do not match, the system crackdowns and reports counterfeit.

VII. DISCUSSION

The platform, as discussed by Parker et al. [56], is an emerging business model that leverages digital technology to connect people, organizations, and resources in an interactive ecosystem in which tremendous amounts of value can be created and exchanged. The most valuable companies in the world Microsoft, Apple, Amazon, and Google — have harnessed platform power to achieve rapid growth and market dominance as their products and services permeate our daily life [57]. These platforms usually consist of common technological building blocks that the owner and ecosystem partners can share to create new complementary products and services, such as smartphone apps or digital content from Apple iTunes or Netflix [58]. Uber and Airbnb are two other famous companies whose business model is to create transaction platforms which connect suppliers to the consumers. These platforms are largely intermediaries or online marketplaces that make it possible for people and organizations to share information or to buy, sell, or access a variety of goods and services [58]. Platforms generate network effects and positive feedback loops where the value each user experiences, increases as more users adopt the platform [56].

The idea of generating a fingerprint and issuing and storing digital certificate of handwoven carpet on the blockchain through the platform is novel and still under thorough research in order to study the feasibility and value proposition for stakeholders who are brought together to produce network effects. Yet, Chamrosh Technology is already demonstrating the “start-up soul” argued by Gulati [59] in which he observed that startups have an essential and intangible notion such as energy and soul which inspire people to contribute their talent, money, and enthusiasm and fosters a sense of deep connection and mutual purpose. Chamrosh is intended to leverage digital technologies to co-create the legacy of luxury handwoven carpet to globally protect and promote Persian carpet, which struggles to keep the centuries-old industry alive. This is an ambition to fill needs that handwoven customers had not realized before, and it is as if Chamrosh is articulating new problems that cannot be solved by existing solutions. The concept of an utterly new category was highlighted in “ How Unicorns Grow” [60] as a successful strategy formulated by Facebook, LinkedIn, and Tableau who were carving out entirely new niches. Ramadan et al. [61] believe that the most successful start-ups become Category Kings and bury the competition and therefore, they argue that a common phrase in technology ‘There’s room in this market for several players’ seems to be utterly false. The data shows that there’s room in any well-defined category for one super successful player and a bunch of companies that wind up as category serfs and those that just die [60].

The business intent of Chamrosh is, therefore, the first element in Gulati’s [59] dimension of the soul which can run deeper and spark a different, more intense kind of commitment and performance as argued in his framework. Moreover, the research-oriented approach resonates with improvisation concept suggested by Gulati and DeSantola [62] in which they argue that young ventures need to try new things and react to dynamic markets, but with an eye toward larger objectives and sustaining the business. The brand is created with this strategy in mind since Chamrosh is a bird in Persian mythology with the body of a lion and head and wings of an eagle. It is said to live on the summit of Mount Alborz and is sent by an angel to snatch invaders and drop them from mountaintops in order to protect the Persian Land.

Chamrosh vision is to be the global platform to cherish the odyssey of Persian art and culture.

Figure 19. Chamrosh Brand [39]

The second dimension is customer connection where Chamrosh business model tries to inspire carpet fans to contribute and be part of the narrative of a handwoven carpet and co-create a brand for example through the style and location of the carpet which are all collected and stored while observing the confidentiality and privacy issues.

The third dimension is to allow people and employees to have a voice and choices. Here, Chamrosh must work very hard to give employees ‘freedom within a framework”—the liberty, i.e., to operate within well-delineated boundaries—as well as opportunities to influence key decisions, such as which strategies or products to pursue as suggested by Gulati [59]. This dimension is critically highlighted in other studies. For example, Henrich [63] argues the internal startups, filled with a small number of passionate believers dedicated to one project at a time, can experiment rapidly and scale their impact. Gulati and DeSantola [62] consider culture a big part of what draws people to join start-ups—and what keeps them going. As employees battle the odds to turn a fledging business into a viable company, working late nights and weekends to make it happen, they’re motivated by camaraderie and a sense of belonging to something important. Also, the founders must develop and practice some crucial skills and traits including building a great team, leadership, team management, selling, marketing, product design, strategy formulation recommended on a survey of 141 founders of enterprises by Eisenmann et al. [64].

VIII. CONCLUSION AND FUTURE WORKS

Chamrosh like other successful startups can possibly preserve and revive its soul and remain innovative and grow as long as it keeps those three elements suggested by Gulati [59] at the forefront of its strategy and daily operations. This will take time, effort, and iteration to improve the business model and enhance technology. But the critical question is whether the handwoven carpet market can be cultivated into an active ecosystem that appreciates Chamrosh solution. This question resonates with the “scalable startup” definition [65] which argues that startup must aim not only to prove its business model but to do so quickly, in a way that will have a significant impact on the current market. There are three key ingredients in this definition :

• First, the goal is to seek, explore, examine, and validate an unmet need within the handwoven carpet market by providing a vision of a smart carpet with a set of features.

• Second, assumptions and hypotheses about carpet platform model are made to iterate until they are proved.

• Third, Chamrosh must quickly validate the model and check if customers behave as the model predicts. The feedback and input of stakeholders, including weavers, dealers, and carpet fans, are instrumental in answering this.

The above three ingredients are in line with three innovation challenges raised by Eppinger [66]. Here Chamrosh suggestes digital certificate solution to build trust with handwoven carpet fans (People). Then it leverages blockchain and image processing to develop the solution (Technical). Finally, it formulates a platform model to encourage stakeholders to contribute.

Figure 20. Three innovation callenges [66]

Building on platform model, Chamrosh enjoys two significant benefits of platforms highlighted by Parker el. al [56]. First, it leverages marginal economics of production and distribution, since Chamrosh can expand with minimum marginal cost just by integrating digital technologies on selected new carpets when the listing fee of carpets is very small. Second, the possibility to scale quickly is enhanced by network effects of carpet fans.

The other important question after successfully passing former stages is that how quickly Chamrosh can become a Category King, dominate the entire technology industry, bury the competition and reach a pinnacle that even founders could not have dreamed. This demands high levels of trust, cooperation and innovation which can be achieved by infinite mind set in business as suggested by Sinek [67]. He argues that the game of business fits the very definition of an infinite game and infinite mind ensures that an organization’s employees, customers and shareholders remain inspired to continue contributing with their effort, their wallets and their investments [67]. Chamrosh can embrace surprises and adapt with and even transform by changes if it is structured to last forever and this is the nature of a company built for resilience i.e., as highlighted by Sinek [67]. Time will reveal if Chamrosh can do this as “ most overnight successes took a long time” Steve Jobs.

Future research is suggested on evaluating the feedback and comments of stakeholders, including handwoven carpet fans and the owners of luxury carpets. Also, the image processing solution must be put into practice to assess the feasibility and technical implications

Chamrosh has developed Minimum Viable Product (MVP) and is planning to introduce the first Blockchain-Empowered “Carpet” or “Rug” focused on verifying the origin and authenticity in the first quarter of 2020. The steering committee of the startup has identified a masterpiece that crystalizes the craftsmanship quality of Persian rug.

Figure 21. The first Blockchain-Empowered rug in the world [68]

The creator of this unique 1.5m x 2.46m item, Ali Lahiji, spent 8.5 years designing the original map and weaving the layout. The rug was completed in early 2019 and has 1000 different artistic elements, including flowers, birds, and animals in 110 compound colors of pure silk. These elements have a significant meaning which reflects their close affinity with nature. The rug is woven with 6 million knots and will be the first in the world that has equipped with electronic fingerprint and certificate of authenticity stored on the blockchain. The aim is to create a unique user experience through a mobile application where porential customers can authenticate and verify the history of this masterpiece.

IX. CONTRIBUTION

The sharing economy, platforms, and digital technologies are disrupting the luxury sector by offering a new catalyst for growth and enabling individuals to share goods and services and even contribute to their design, production, and consumption.

The first and significant contribution of this study is to bring insights into the emerging and disrupting role of blockchain in the sharing economy and luxury industry for reinforcing provenance, building trust, and verifying the credentials and ownership.

The real application of blockchain technology in one of the most traditional businesses, and enumerating the benefits is the second contribution. Here, it is highlighted how a handwoven carpet can be transformed into the smart carpet as a blockchain-empowered product which enables carpet fans to verify the authenticity, claim ownership, and enjoy scarcity and exclusivity.

Third, the study points to the vital role of a strategic partnership between new ventures to explore and seize evolving opportunities for maximizing value for both parties especially in the digital era where most companies become increasingly interdependent.

Forth, the study suggests an innovative revenue model for strategic partners to earn money from second-hand trade onwards on the blockchain.

Last but certainly not least, this study offers state-of-the-art solutions by utilizing image processing to generate an identifiable feature for any handwoven carpet. This feature can be as unique as human fingerprint.

ACKNOWLEDGMENT

The first author would like to express gratitude to several academics and industry experts who have supported him: Kevin Lane Keller, the E. B. Osborn Professor of Marketing at the Tuck School of Business at Dartmouth College, for his guidance, and invaluable moral support. David Rogers, Faculty Director-Digital Business Strategy & Leadership at Columbia Business School, whose knowledge was the source of inspiration. Francisco Guillén and Enrique Lizaso, the co-founders of BlockTac whose dedication and expertise were instrumental in blending this research with practice. Ali Roshanzamir, Shahin Emai Alaei, and Narges Roshanzamir for their invaluable support. Ahmad Rahnema, Professor of Financial Management at IESE Business School for his immeasurable direction and dedication. Ahmad Kardan, Emeritus Professor of Information Technology and eLearning at Amirkabir University of Technology for his wisdom and encouragement. Ahmad Roosta, Professor in Marketing and faculty member of Shahid Beheshti University for his priceless support. Hamid Reza Baradaran Shoraka, Visiting Faculty at the University of Florida for his support and guidance. Zahir Irani, Pro-Vice-Chancellor (Academic, Innovation and Quality) at the University of Bradford for his support and encouragement. Andrew Smith, Director of DBA at the University of Bradford for his continuous support. Ahmad Ghoneim, Senior Lecturer and Director of Accreditations at the Faculty of Management and Law, University of Bradford and Martin Sedgley, Head of Effective Learning at the University of Bradford for their encouragement and support. Ali Najafi, Seyed Behrouz Safawi and Mohsen Ghaniyari for continuous guidance and direction. Arezoo Kheiltash for her immeasurable patience and support. Masoud Meshkat for his creativity and craftsmanship. Jalal Mirdavoodi, CEO of Rayan Group for his feedback and sharing industry experience. Gita Beyk for her invaluable support and reviewing the paper. The authors are also grateful to the Industry Experts, Farrokh Bahmanyar, Mehdi Fakhkharian, Mohammad Khakbazan, Ali Lahiji, Mohmmad Heravi, Hossein Bidabadi, Kambiz Norouzi, Mohammad Taheri, Ali Farzad Mohajeri, Hamid Ismaili, Alireza Aminifar, Mohsen Roshanzamir, Mojtaba Roshanzamir, Alireza Afraz, Farzin Raeeszadeh, Shahram Pourshams, Payam Saltanatkhah, Forouzandeh Farrokhi and Allahyar Mansubi for their idea, valuable support and contribution.

REFERENCES

[1] A.H. Roshanzamir, V. Weerakkody and V. Yoganathan, "A Rhetoric of Smart Carpet in the Age of Disruption," The Eighth International Conference on Ambient Computing, Applications, Services, and Technologies Athens, Greece, 1-10, 2018.

[2] T. Laurence, "Blockchain For Dummies," Wiley. Kindle Edition, 2017.

[3] W. Mougayar and V. Buterin, "The business blockchain: promise, practice, and application of the next Internet technology," Hoboken, New Jersey, John Wiley & Sons, Inc., 2016.

[4] B. Blöbaum, "Trust and Communication in a Digitized World: Models and Concepts of Trust," Research by Springer; 1st E., 2016.

[5] T. Hardjono, D. Shrier and A. Pentland, "Trust Data: A New Framework for Identity and Data Sharing," Visionary Future LLC, 2016.

[6] C. M. Christensen, "The innovator's dilemma: When new technologies cause great firms to fail," Harvard Business School Press. 2016.

[7] D. L. Rogers, "The Digital Transformation Playbook: Rethink Your Business for the Digital Age," Columbia Business School Publishing, Columbia University Press. Kindle Edition, 2016.

[8] O. Malekan, "The Story of the Blockchain: A Beginner's Guide to the Technology Nobody Understands," Triple Smoke Stack. Kindle Edition, 2018.

[9] M. A Aud et al. "SMART CARPET Developing a Sensor System to Detect Falls and Summon Assistance," Journal of gerontological nursing 36, 8-12, 2010.

[10] D. Zax, "A Smart Carpet to Detect Intruders," https://www.technologyreview.com/s/429130/a-smart-carpet-to-detect-intruders/ 2019.09.01

[11] A.H. Roshanzamir, "Dilemmas of Branding for Start-ups in Digital Era, "The Twelfth International Conference on Digital Society and eGovernments Rome, Italy, 63-72, 2018.

[12] D. Maxwell, C. Speed and L. Pschetz, "Story Blocks: Reimagining narrative through the blockchain Convergence," The International Journal of Research into New Media Technologies, 23(1), 79-97, 2017.

[13] A. Morrison, "Blockchain and smart contract automation: How smart contracts automate digital business.", 2015. http://www.pwc.com/us/en/technology-forecast/blockchain/digital-business.html / 2019.09.01,

[14] C.E. Cherry and N.F. Pidgeon, "Is sharing the solution? Exploring public acceptability of the sharing economy," Journal of Cleaner Production, 939-948, 2018.

[15] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[16] N. Popper, "Digital Gold: Bitcoin and the inside story of the misfits and millionaires trying to reinvent money," New York City: HarperCollins Publishers. 2015.

[17] N. Szabo, "Smart Contracts: Building Blocks for Digital Markets" 1996.http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart_contracts_2.html/ 2019.09.01

[18] A. Wright and P. De Filippi, "Decentralized Blockchain Technology and the Rise of Lex Cryptographia," SSRN Electronic Journal. 2015.

[19] A.G. Woodside, S. Sood and K.E. Miller, "When consumers and brands talk: storytelling theory and research in psychology and marketing" Psychology and Marketing 25(2), 97-145, 2008.

[20] D. Sole and D. G. Wilson, "Storytelling in Organizations: The power and traps of using stories to share knowledge in organizations," Training and Development, 53(3), 1-12, 1999.

[21] C. Fill, " Marketing Communications," Pearson Education Limited 6th edition, 2013.

[22] R. McKee, "Storytelling That Moves People: A Conversation with Screenwriting Coach Robert McKee," Harvard business review, 81(6), 51-55, 2003.

[23] S. Roper and C. Fill, "Corporate Reputation Brand and Communication,' Pearson Education Limited, First Edition, 2012.

[24] K. L. Keller, "Strategic Brand Management," 4th Edition Pearson HE, Inc. Kindle Edition, 2013

[25] H.-C. Chiu, Y.-C. Hsieh and Y.-C. Kuo "How to align your brand stories with your products," Journal of Retailing 88(2), 262-275, 2012.

[26] C. Fill and S. Turnbull, "Marketing Communications," Pearson Education Limited 7th E., 2016.

[27] S. Spear and S. Roper, "Using corporate stories to build the corporate brand: an impression management perspective," Journal of Product and Brand Management, 22(7), 491-501, 2013.

[28] T.J. Brown, P.A. Dacin, M.G. Pratt and D.A. Whetten, "Identity, intended image, construed image, and reputation: An interdisciplinary framework and suggested terminology," Journal of the Academy of Marketing Science, 34(2), 99-106, 2006.

[29] G. Dowling, "Communicating Corporate Reputation through Stories." California Management Review, 49(1), 82-100, 2006.

[30] D. Miller, "Building a Story Brand: Clarify Your Message So Customers Will Listen," HarperCollins Leadership. Kindle Edition. 2017.

[31] https://www.youtube.com/watch?v=ysSgG5V-R3U / 2019.09.01

[32] V. Barassi, and E. Treré, "'Does Web 3.0 come after Web 2.0? Deconstructing theoretical assumptions through practice," New Media & Society, 14(8), 1269-1285, 2012.

[33] T. O'Reilley, "What is Web 2.0: design patterns and business models for the next generation of software," MPRA, available at http://mpra.ub.uni-muenchen.de/4578/ 2007. /2019.09.01

[34] C. Fuchs et al., "Theoretical foundations of the web: cognition, communication, and co-operation. Towards an understanding of Web 1.0, 2.0, 3.0", Future Internet, 2(1), 41-59,2010.

[35] D. Harris, "Web 2. 0 Evolution into the Intelligent Web 3.0: 100 Most Asked Questions on Transformation, Ubiquitous Connectivity, Network Computing, Open Technologies, Open Identity, Distributed Databases and Intelligent Applications," Emereo Publishing, Newstead. 2008.

[36] M. Tasner, "Marketing in the Moment: The Practical Guide to Using Web 3.0 to Reach your Customers First," FT Press, Upper Saddle River, NJ. 2010.

[37] M. Watson, 'Scripting Intelligence: Web 3.0 Information, Gathering and Processing' (Apress Series: Expert's Voice in Open Source), Springer Publishing, New York, NY. 2009.

[38] T. Berners-Lee, "Web 2.0 Summit 09 discussion: A Conversation with Tim Berners-Lee interview by Tim O'Reilly [video]. San Francisco: O'Reilly Media. Retrieved from https://youtu.be/KY5skobffk0, 2009 /2019.09.01

[39] www.chamrosh.com / 2019.09.01

[40] K. Smith and M. Wintrob, "Brand Storytelling: A Framework for Activation," Design Management Review, 24(1), 36-41, 2013.

[41] A. Ha, " Blockai Uses the Blockchain to Help Artists Protect Their Intellectual Property," 2016; Retrieved from https://techcrunch.com/2016/03/14/blockai-launch/ 2019.09.01

[42] J. N. Kapferer, "Kapferer on Luxury How Luxury Brands Can Grow Yet Remain Rare," Kogan Page. 2015.

[43] http://www.bbc.com/culture/story/20171206-the-timeless-appeal-of-the-persian-rug /2019.09.01

[44] B. Murphy, "The Root of Wild Madder: Chasing the History, Mystery, and Lore of the Persian Carpet," Simon & Schuster. Kindle Edition. 2005.

[45] Anonymous,"Getting acquainted with samples of Persian Carpets,' Iran Carpet Co., Vol. 2, 1984.

[46] https://financialtribune.com/articles/economy-business-and-markets/80480/Hand-woven-carpet-exports-to-hit-400m /2019.09.01

[47] www.blocktac.com /2019.09.01

[48] A.H. Roshanzamir, A. Roosta, and M. Shajari "CHAMROSH Technology: Transforming handwoven carpet to smart carpet," World of Export, 10, Fourth year/Summer and Autumn, 9–11, 2019.

[49] X. Zhang and W. Dahu, "Application of artificial intelligence algorithms in image processing," Journal of Visual Communication and Image Representation, Vol.61, 42–49, 2019.

[50] W. Burger and M.J. Burge, "Digital Image Processing: An Algorithmic Introduction Using Java, Springer" 2nd Edition. 2016.

[51] M. Ponti, T. Nazare and G. Thume, "Image quantization as a dimensionality reduction procedure in color and texture feature extraction," Neurocomputing Vol. 73, 385–396, 2016.

[52] BBC News " Game of Thrones makers on that coffee cup blunder and season eight ", September 2019. https://www.bbc.com/news/newsbeat-49542412?ns_mchannel=social&ns_campaign=bbcnews&ocid=socialflow_twitter&ns_source=twitter /2019.09.10

[53] A. Cheddad, J. Condell, K. Curran and P.M. Kevitt, "Digital image steganography: survey and analysis of current methods," Signal Processing, 90, 727–752, 2010.

[54] I.J. Kadhim, P. Premaratne, P.J. Vial and B. Halloran , "Comprehensive survey of image steganography: Techniques, Evaluations, and trends in future research," Neurocomputing, 335(28), 299-326, 2019.

[55] AZ. Tirkel et al. 'Electronic water mark.Proceedings of Digital Image Computing: Techniques and Applications," Macquarie University, 666–673, 1993.

[56] G.G. Parker, M.W. Van Alstyne and S.P. Choudary, "Platform Revolution: How Networked Markets Are Transforming the Economy--and How to Make Them Work for You," W. W. Norton & Company. Kindle Edition.

[57] M. A. Cusumano, A. Gawer and D.B. Yoffie, "How Digital Platforms Have Become Double-Edged Swords," MIT Sloan Review, 2019.

[58] M.A. Cusumano, "The Business of Platforms," HarperBusiness. Kindle Edition, 2019.

[59] R. Gulati, " The Soul of a Start-Up", Harvard Business Review July–August 2019, 85-91, 2019.

[60] Anonymous "How Unicorns grow' January–February 2016 Harvard Business Review, 28-30, 2017.

[61] Al Ramadan, C. Lochhead, D. Peterson and K. Maney, “Time to Market Cap: The New Metric That Matters,” Play Bigger Advisors, LLC, 1-23, 2014.

[62] R. Gulati and A. DeSantola “Start-ups that last,” Harvard Business Review March 2016, 54-61, 2016.

[63] S. M. L. Henrich, “Teaching a giant to think like a startup,” fortune.com /Nov.01.17, 94-102, 2017.

[64] T.R. Eisenmann, R. Howe and B. Altringer "What Does an Aspiring Founder Need to Know?," Harvard Business Review Digital Articles. 6/21/2017, pp.1-6, 2017.

[65] E. Ries,"The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses,” The Crown Publishing Group. Kindle Edition. 2011.

[66] S.D. Eppinger, “MIT’s Approach to Design Thinking” Online Course by Emeritus Institute of Management in collaboration with MIT Sloan School of Management, 2018.

[67] S. Sinek, “The Infinite Game,” Penguin Publishing Group. Kindle Edition. 2019.

[68] https://www.instagram.com/p/BssaWndgKZp/

Web-based Visualization of Daily Mobility Patterns in R

Antje von Schmidt, Rita Cyganski, Matthias Heinrichs
Institute of Transport Research
German Aerospace Center (DLR)
Germany, Berlin
e-mail: antje.vonschmidt@dlr.de, rita.cyganski@dlr.de, matthias.heinrichs@dlr.de

***Abstract*-Human mobility reflects how, when, where and why people move from one location to another. Transport demand models can provide answers to these questions. Such models usually require a large amount of data as input and provide detailed information about the trips made by each individual during a day. Exploring this data can become very complex. Usually, several types of aggregation and disaggregation are performed on a spatial, temporal or demographic level. Often a variety of tools is used for analyzing, communicating and validating the data. This paper introduces an interactive and scalable web application for analyzing, communicating and validating the data of a transport model. The presented approach could also be used within a different research domain. Therefore, recommendations and implementation notes are given on how such an application can be realized in R. A special focus is set on the representation of daily mobility patterns within this application.**

Keywords-transport demand; daily mobility patterns; visualization; R; Shiny.

I. INTRODUCTION

This paper is an extended version of previous work presented at the IARIA Seventh International Conference on Data Analytics [1].

Human mobility reflects how, when, where and why people move from one location to another. Biking, going by car, using public transport, and walking are often considered modes to accomplish activities such as education, working, shopping, leisure or other things. Transport models can be used to provide a realistic picture of the current traffic situation, to predict future developments of transport demand or to undertake scenario-based analysis of various possible development paths, such as an aging population, changed prices or new mobility trends. To estimate the demand that is analyzed and visualized in this paper, the microscopic transport demand model TAPAS [2][3] is used. Such models usually require a large amount of data as input. Within a simulation run, TAPAS calculates the activities and trips performed during a day for each person in the research area. Thereby, it provides individual trip chains with specific spatial and temporal information as well as a detailed description of each person and the associated household as simulation output. The sum of these trips results in an overall picture of the transport demand within a specified study area and timescale.

Not only the validation of input and output data, but also the analysis and representation of simulation results are not easy tasks, especially if the target audience is heterogeneous and several output media have to be covered. There are many visualization types available, and which one to choose strongly depends on the research question and the domain of interest. In the field of transport research, several types of aggregation and disaggregation are performed, may they be on a spatial, temporal, or demographic level. A wide range of visualization tools are available [4]. Some of them can be used out of the box, and others require programming skills or are only intended for a certain type of visualization, spatial or non-spatial. Consequently, often a combination of these tools is used, also because commercial software solutions or eye catching animations are usually too expensive for being applied within research projects. In general, a flexible and extensible approach is preferable, which allows adaptation to the respective needs.

The aim of this paper is to introduce the application “Transport Visualizer” (TraVis). It is an interactive, user-friendly and scalable web-based solution for analyzing, communicating and validating the simulation data of a transport model, such as TAPAS. A special focus is set on the representation of daily mobility pattern within this application. TraVis is based on the programming language R, which is known as a flexible and powerful open source language. R is widely used in the scientific field for statistical computing, data analysis and visualization [5]. This application architecture was chosen because of the possibility of combining the benefits of R with the interactivity of modern user interfaces by using the Shiny web-framework [6].

The paper is organized as follows: Related work is discussed in Section II, followed by an introduction of the simulation data used with TAPAS in Section III. Section IV describes the implementation requirements and Section V gives recommendations, on how to set up a web-based application in R. The usage of spatial data and the realization of the TraVis application are outlined in Section VI respectively in Section VII. The visualizations of the daily mobility patterns within this application will be presented in Section VIII, using a baseline scenario for the city of Berlin in 2010. Finally, Section IX summarizes this paper and gives an outlook on future work.

II. RELATED WORK

For a long time, the role of visualization in transport modelling was largely limited to the static presentation of aggregated final results, primarily through statistical tables and simple graphs or the provision of key indicators describing the development of transport demand (see for instance [7][8]). In recent years, map-based representations were progressively used within the transport domain. In particular, the level of detail is increasing: from a general overview of the study area, to a certain spatial resolution of it or temporal selections, down to the visualization of individual travel behavior [9]-[12]. Modern web technologies [13]-[15] as well as the growing availability of data and computing power have contributed to a significant increase in the usage of dynamic and interactive forms of visualization. In transport as in other research fields, it is necessary to analyze the visualization requirements and consider the type of data to be visualized, the audience to be addressed, the purpose of the visualization, the appropriate level of detail, the aspects of the data that should be transmitted, and the target medium for which the visualization is generated. A general framework for the visualization of transport data and models is defined in [16].

Another important factor is the type of visualization. Charts are often used to represent information on a high or intermediate level of detail. Which type of illustration should be used depends on the data type and on what is to be shown. In [17], a chart selector guide and four basic methods of data analysis are defined that can help to choose the right chart type for comparison, composition, distribution, and relation. The comparison of single values, such as totals or means, is best shown with regular bar or column charts, while line charts are more suitable for identifying distributions of continuous values or the development of a measure over time. Stacked charts can be used to represent the absolute or relative changes within the composition of categorical variables. Scatter and bubble charts are suitable for representing correlations between two respective three numerical variables, whereas parallel coordinates can be used to point out relationships between multivariate data. A chord diagram is a common way to illustrate interrelation between data in a matrix, whereas data with a spatial context are typically represented by suitable maps. For example, a choropleth map can help to identify regional varieties, while flow maps show the quantity of movements between geographical units. Spatial changes between scenarios are usually achieved by difference maps, whereas animated maps are suitable to represent differences over time. There are many more ways for visualizing data, and not all of them can be listed here. Hence, only a brief insight into the diversity is given. A good overview can be found at [18].

Although many visualization concepts and tools are available, the challenge remaining is to integrate these different approaches and to enable the users to perform their analyses without mastering programming languages or using a variety of tools.

III. Simulation Data

Microscopic transport demand models usually require a variety of different input data, such as population, locations, network, transport offers, timetable and land use. All this data is very heterogeneous in terms of format, spatial resolution, and time frame. TAPAS, for instance, requires all possible locations within a specified study area where activities such as working or shopping can be realized, but also the respective capacities indicating the number of people that can be there at the same time. Furthermore, a main input is a highly differentiated synthetic population, which is generated by the SYNTHESIZER [19]. Within such transport models, individual and household information play an important role in the choice of destination and transport means. Therefore, a synthetic population contains information both at the personal and household level. In the case of TAPAS, age, sex and an employment status variable are used on the individual level. In addition, information on available mobility options is required, such as a driving license, a public transport ticket, a bike as well as the budget a person can spend on mobility. Household information comprises the number of persons, the total household income, the number and type of vehicles that belong to the household as well as the spatial reference of the address. The simulation results, on the other hand, provide detailed information for each individual trip, including, among other, start time and duration, transport mode chosen, the distance as well as information about the trip purpose, and the location of origin respectively destination. Table I provides an excerpt of the simulation output. The identifier of each person and the corresponding household can be used to merge further information. Overall, data associated with an average simulation run is quite vast, reaching roughly 11.5 million trips using a baseline scenario for the city of Berlin with a population of 3.3 million inhabitants in 2010.

IV. Implementation Requirements

Exploring the simulation data of a transport model can become very complex. The tabular representation of the simulation output, shown in Table I, is not easily to interpret by humans, especially when handling large amounts of data. Based on related work, the following aspects were selected as requirements for the implementation.

TABLE I. Generated Individual Activity Based Daily Mobility Plan For A Sample Person

P-ID	**HH-ID**	**Trip**				**Purpose**			**Location**	
		Start time	*Duration*	*Mode*	*Distance*	*Activity*	*Start time*	*Duration*	*Start*	*End*
1	5	7:15	10 min	bike	2 km	shopping	7:25	5 min	at home	at a shop
1	5	7:30	30 min	public transport	10 km	working	8:00	8:30 h	at a shop	at work
1	5	16:30	15 min	bike	3 km	leisure	16:45	1:00 h	at work	at a gym
1	5	17:45	45 min	bike	9 km	leisure	18:30	12:00 h	at a gym	at home

A. Data type

As mentioned before, the simulation data includes a variety of heterogeneous data types, which all have to be taken into account for exploring the data and the resulting daily mobility pattern in a disaggregated or aggregated manner. Besides the data with a spatial context such as activity location, traffic flow, population density or origin-destination matrices, there are also descriptive variables. These can be further subdivided into continuous, discrete and categorical data, such as the trip length, person age, or the activity performed.

B. Target audience and purpose

The visualizer shall help modelers to review the model values and disseminate the results to a wider audience, including both the scientific community and the public. Therefore, the usage of different media, including scientific publications, static presentations or interactive visualizations is intended.

C. Level of detail

The target application should contain all levels of detail of spatial, temporal, and demographic dimensions. It should be possible to aggregate simulated data at different viewing modes – for the complete study area or parts of it. This can be used, e.g., to validate the applied synthetic population, including the vehicle fleet and mobility options. The simulation result should be used for computing common key indicators of the transport demand (e.g., modal split) by aggregation. At the most detailed level, the individual travel behavior should be extractable from the simulation output and visualized. In addition, the usage of public space should be included in the analysis. It might be interesting to know how much space is used for parking vehicles. This is particularly interesting due to the space limitations and competition in cities. The space requirements for parking may change, for example through an increasing use of mobility services (e.g., bike-, car- or scooter-sharing) or the introduction of autonomous vehicles. Therefore, this information can provide useful information for future urban planning.

D. Output medium and interactivity

According to the different audience and purpose, both static as well as dynamic media have to be addressed. To understand changes, for example between different scenario parameters or time frames, it is important to provide the possibility to compare the corresponding simulation results. To take a closer look at certain aspects of the simulation output, the use of filters should be supported, including the following filter types: specific groups of persons or households, mode choice, location, trip purpose, time of traveling, distance, trip type and specific part of the study area. Detailed filter options are shown in Table II. These filters should be used to limit the simulation result accordingly.

TABLE II. FILTER OPTIONS

Filter	**Elements**
Region type	agglomeration, urban, sub urban, rural, zone identifier
Household size	1, 2, 3, 4, 5+
Number of cars	0, 1, 2
Person group	kids (< 6), pupil, student, employed, unemployed, pensioner
Transport mode	walk, bicycle, car, car (co-driver), public transport, other
Activities	education, leisure, private matters, shopping, working, other
Locations	education, leisure, private matters, shopping, work, other
Trip type	local people, commuters, origin, destination
Time of traveling	early (0 - 6), morning (7- 12), afternoon (13 - 18), evening (19 - 24)
Distance	1000m, 2000m, 3000m, 4000m, 5000m, 5000m+

Another main aspect is the opportunity to adjust the spatial scale, whereby different resolution of zones or the European wide standardization for geographical grids [20] may be applied. To address the various target media, it is also required to export appropriate figures for print media and to show the results on screen.

E. Visualization type and customization

With regards to the different types of data and the various target media, the application should contain suitable visualizations, such as different charts, maps and animations. This is necessary for presenting the simulation data according to the desired level of detail. Furthermore, the application should be scalable so that new types of reporting can be integrated immediately. Besides the choice of the right visualization type, it is also very important to have the opportunity to adopt its appearance. Therefore, it is required to customize attributes related to the layout, such as the color of the bars, the chart background, and the position of labels or of the legend as well as the appropriate font type and size. With the focus on internationalization, it is also necessary to supply different languages in order to automatically generate the required labels for the visualization.

V. RECOMMENDATIONS ON HOW TO SET UP AN WEB-BASED APPLICATION IN R

The development of a web-based application is not always straightforward. If a comprehensive application has to be created, the implementation can quickly become complex and time-consuming. It is therefore recommended to create readable and reusable code. In the following section, some recommendations are given, on how to set up a web-based application in R.

A. Use a development environment

RStudio, for instance, has established its position as a development environment in the R community. This tool should be used in conjunction with an up to date R version. Both are platform independent and a freely accessible version is available. When creating a new application, select "Shiny Web Application" as the project type within RStudio. This automatically integrates the Shiny web-framework. After the creation, a first example application is available and can be adapted accordingly.

B. Apply a usable application structure

A basic Shiny application is usually structured either in one script (*app.R)* or split into two: a user-interface script (*ui.R)* and a computational script (*server.R)*. The *ui.R* script includes the layout and all the user-interface elements of the application, whereas the logic is specified in the *server.R* script. Here, for example, is defined what will occur if the user changes the value of a select box. Both scripts interact with each other through input and output objects. Splitting the programming code in only two scripts may be sufficient for smaller applications, but the programming code can quickly become unwieldy in larger ones. Therefore, it is recommended to split these main scripts further to make the code easier to maintain for developers. This can be additionally supported by a modular application design.

C. Create modules to use elements several times

Recently, the capability of using modules was added to Shiny as a new feature [21]. Shiny modules can be used to capture functionality and avoid name collisions by using namespaces for the input and output elements. This is especially important as element identifiers must be unique within a Shiny application. A module can be considered as a function that can be called several times within an application and is also structured either in one script *(module.R*) or split into two: a user-interface script (*ui.R*) and a computational script (*server.R*). To use modules across applications, it is recommended to pack them into a separate R package. This provides the full benefits of packaging, such as an update of the module package will automatically update all affected Shiny applications accordingly. Before creating modules, one should consider what recurring features are needed in the application.

D. Define the data store used

Data is often stored in a database. PostgreSQL, for example, can be used as a database system. A free version is also available. With the PostGIS extension, spatial data can also be managed within such a database. But, data can also be stored in text files instead of in a database. Within R it is possible to access all common data types within various data stores.

E. Reduce the data volume

Within the R environment, all data is stored in the working memory. When analyzing large amounts of data, this can slow down the performance. Therefore it is recommended to reduce the data volume. This can be done, for example, by aggregating the data accordingly.

F. Create a concept for managing spatial data

Spatial related evaluations within the field of transport research have to be often carried out for different spatial scales. Furthermore, the boundaries of a spatial zone may change or new zonal levels are added. Therefore, the data have often been spatially adapted accordingly. The spatial mapping of data on the fly can be very time consuming. For this reason, it is recommended to create a concept to avoid data redundancies and an unnecessarily inflation of the data volume.

G. Use HTML widgets for interactive visualization types

A Shiny application can easily be extended by including HTML widgets. They provide an interface to a specific JavaScript library and generate interactive visualizations. The usage within R usually does not require further web development skills. There are already several widgets for interactive visualization types available, such as for charts, maps, or data tables.

H. Provide multi-language support

Results of research projects often have to be presented on national and international conferences or in journals. Therefore, it is necessary to provide the labels of the figures in the respective language. The labels should therefore not be integrated statically into the application, but instead be dynamically adaptable. This can be achieved, for example, by providing and using resource files for each language.

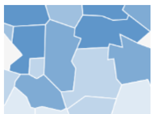
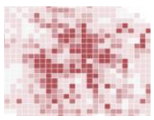
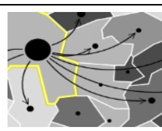
I. Decide, how to share the application with others

A Shiny application can be started locally within RStudio. To make this application available, the entire application folder with the associated files can be passed to other users. To start this application, it is necessary that the user has RStudio and R installed on the computer. This type of application sharing may be sufficient within a workgroup. However, it is also possible to run the application on an own web server. Therefore, the Shiny Server can be installed on a Linux operating system. This server is freely available as open source. In addition, such an application can also be hosted for free or commercially via the RStudio service.

VI. USING SPATIAL DATA IN R

Spatial data plays an important role within the simulation data. Not only for the preparation of the input data, but also in the later analysis and visualization of the simulations. Table III includes possible geometry types to represent transport related issues. Whereby, relevant attributes of the types can be used to highlight content related issues. The color of a point for example can be used to present different activity locations, whereas the size can be used to visualize the location capacity and the line thickness can help to show a different traffic volume.

TABLE III. GEOMETRY TYPES TO REPRESENT TRANSPORT RELATED ISSUES

Geometry type		**Intended purpose (e.g.)**
Point		activity location, location capacity
Line		origin and destination of a trip
Polygon		population density
Grid		car density
Combination		aggregated traffic flow

This kind of data may be available in various formats, e.g., as shapefile, text file or in a database. R offers several packages for handling spatial data. The package *rgdal* [22] for instance includes functions for reading and writing shapefiles. Data available in the geoJSON format can be integrated using the *geojsonio* package [23]. The *rpostgis* package [24] provides an interface to PostGIS enabled PostgreSQL databases. Currently, spatial objects in R are often based on spatial classes which are specified in the *sp* package [25]. The recently created *sf* package [26] supports simple features, a standardized way to encode spatial vector data and could replace the *sp* package in the future. The package links to a lot of other R packages, include many operations to manipulate spatial data, and can be used for reading and writing the mentioned formats. Furthermore, the *sf* package interacts very nicely with the *tidyverse* package [27] and makes it quite easy to create maps in conjunction with other appropriate packages.

Depending on the intended use, maps may be needed either in a static or interactive way. For instance, the packages *cartography [28]*, *ggplot2 [29]* or *tmap [30]* can be used to plot static maps, whereas interactive maps can be created with packages such as *leaflet [31]*, *highcharter [32]* or as well with *tmap*.

VII. REALIZATION OF THE TRAVIS APPLIACTION

Based on the implementation requirements, the following parts of the application were realized as described below.

A. Concept for managing spatial data within a database

All simulation data is stored in a PostGIS enabled PostgreSQL database. The database structure of TAPAS is designed to avoid redundant data and unnecessary increase in data volume. Therefore, each spatial scale of a study area is kept in its own table, including an identifier, the coordinates of the boundaries as polygon and further attributes like the zone name. In addition, all household and location addresses as well as other places without direct address, such as playgrounds or parks, are stored as point coordinates in the *geocodes* table. All geometry columns of these tables have been provided with a spatial index. The use of spatial indexes can reduce the time required for spatial queries compared to tables without such an index. Every point coordinate within the *geocode table* is mapped to the desired spatial scales. Finally, the corresponding zone identifier has been stored within the *geocodes* table, each scale in its own column. These identifiers can be used to join the data from the belonging *spatial scale* table. Each point coordinate has also a unique identifier which is joined to the *households* and *locations* table. At the same time, the household and location identifier is included in the *trips* table. With this approach, the spatial mapping of the point coordinates to the polygon of the respective zone only has to be performed once. It therefore offers a great advantage for the later spatial analysis and visualization of the data, as the aggregation to a certain zone is much faster using identifiers than a spatial reallocation. Furthermore, new spatial scales can be integrated easily. An outline of the database structure is shown in Figure 1.

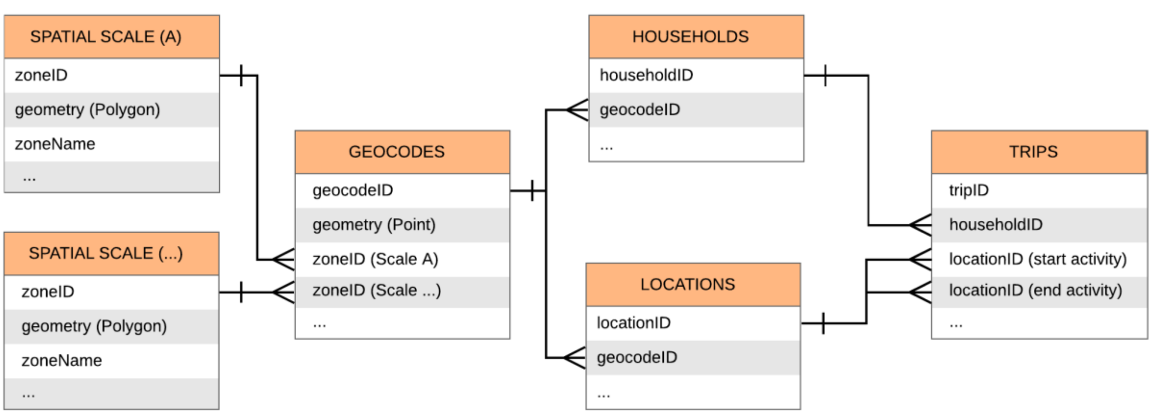


Figure 1. Outline of the TAPAS database structure

B. Approch to reduce the data volume

Loading a full sample, with roughly 11.5 million trips for the city of Berlin 2010, is not really manageable within R. Especially the representation within a web browser does not work well, because the loading of the data is rather slow. To reduce the amount of data, the disaggregated simulation data is not loaded directly into R. Instead, several SQL functions for data preparation were implemented within the database e.g., for data filtering, or data aggregation. This also makes parsing the data with other programs possible. However, the goal is to enable users to perform their analyses automatically and without the use of additional tools.

The simulation results generated by TAPAS are not overwritten when simulation parameters are adjusted; new result data is generated instead. Once generated, the data is static and can be stored in corresponding R objects. This has the advantage that the data does not have to be retrieved from the database again for a new simulation evaluation.

C. Application structure

The two main scripts of a basic Shiny application are split into several parts to make the code easier to maintain. Global objects for example, are defined once in the *global.R* file and used multiple times within other files. Application settings such as the used packages, property files for each provided language, or database configurations are defined within the settings folder. Each part of the application frontend (e.g., header, body, or sidebar) is placed in its own file and referenced within the *ui.R* script. Computational parts are outsourced as functions and used within the *server.R* script. Each created module is stored in its own folder, which contains the corresponding *ui.R* and *server.R* files of the module. Once generated, data (e.g., evaluated data) is stored within the data folder. The developed and used file structure within TraVis is shown in Table IV. An overview of the application architecture is given in Figure 2.

TABLE IV. USED FILE STRUCTURE

Folder/File	Description
/data	stored spatial and evaluated data (*.rdata)
/functions	outsourced functions including: database connection, text formatter
/modules	includes subfolder for each module, which contains the corresponding ui.R and server.R files
/server	computational snippets
/settings	application settings, e.g, used packages, properties (de/en), database config
/ui	user-interface snippets
/www	Stylesheets, JavaScript functions
global.R	global objects, with reference to /settings and /scripts
server.R	main computational file, with reference to /server and /scripts/modules
ui.R	main user-interface file, with reference to /ui and /scripts/modules

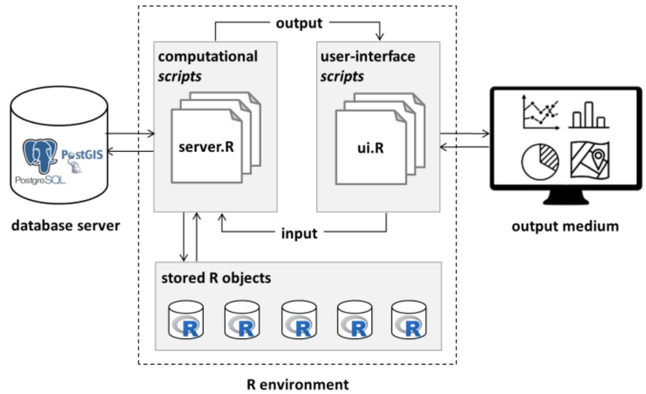


Figure 2. Overview of the TraVis application architecture

D. Modularizing

The following recurring features were defined for the TraVis application so far:

- Render a map view, enabling support for panning, zooming and switching layer on/off, as well as for exporting the map as a printable image. Supply a settings panel to adjust the map layout individually. This includes customization options, such as the color scheme and legend position.
- Supply a map data panel with elements to switch between multiple scenarios and possible categorical values.
- Render a chart view and supply a settings panel to adjust the chart layout individually. This includes customization options, such as the type of the chart, an adoptable color scheme for matching the respective project’s corporate design, text alignment and legend position. Furthermore, provide the possibility to export the generated chart as a printable image.
- Render a panel with value boxes to highlight key indicators.
- Supply a filter panel with elements to limit the simulation result according to the filter options given in Table II.
- Enable an interactive link between the above mentioned elements.

For flexibility reasons and further usage within other applications, these defined features have been split into separate Shiny modules: *map, chart, box* and *filter* module. In the following, each of the sub-modules is described in more detail.

The user-interface of the *map* module contains a data panel with selection fields for the choice of the respective simulation and the associated categorical data. Furthermore, there is a map container and a corresponding settings panel for adapting the map layout. The server part of the *map* module receives several parameters as input, such as spatial aggregated data and the associated geometry. Additionally, it provides a function to view the data according to the selected scenario as well as the chosen category (e.g., the mode walking). For rendering the map, the *leaflet* HTML widget [31], related to the JavaScript *Leaflet* library is integrated within this module. This widget is perfect for displaying maps within a web browser, as all the interactive features, such as panning, zooming and switching layer on/off, come into use. The module structure, including the definition and use of the namespace as well as the linkage of the input and output objects between the user- and server part of the *map* module is outlined in Figure 3.

The *chart* module includes a chart container and a corresponding settings panel for adapting the chart layout within the user-interface. The server part receives several parameters as input, such as aggregated data and the colors to use as default. For rendering the chart, the *highcharter* HTML widget [32], related to the JavaScript *Highcharts* library is used within this module. *Highcharts* contains a large number of interactive chart types and supports the export into different output formats (e.g., png, svg or jpeg). This module includes additionally a function to render the chart according to the selected spatial object within the map.

The *box* module contains a panel with several boxes in the user-interface. The server part, on the other hand, receives the values as parameter and contains a function for rendering these boxes. For this purpose, the value boxes from the *shinydashboardPlus* [33] are used. Additionally, this module contains a function to render the boxes according to the selected spatial object within the map.

The user-interface of the *filter* module contains a panel with select boxes for each defined filter option in Table II. The server part contains a function to render this panel and to filter the data accordingly. Finally, this module is used to return the filtered data to a parent module.

```
#definition of the map module user-interface
mapUI <- function(id){

 #namespace definition
 ns <- NS(id)
 …
 #use of the namespace for the data panel elements
 uiOutput(ns("simulation"))  #list  of simulations
 uiOutput(ns("category"))    #list  of modes, activities, …
 …

 #map container
 leafletOutput(ns("map_leaflet"), …)
 …
}

#definition of the map module server part
map <- function(input, output, session, data, geometry, named.list, …){

  ns <- session$ns
  …
  #render ui-elements for the data panel
  output$simulation <- renderUI({… selectInput(ns("simulation")) …})
  output$category    <- renderUI({… selectInput(ns("category")) …})
  …
  #render map
  output$map_leaflet<- renderLeaflet({…})
}
```

Figure 3. Outline of the *map* module

The *map*, *chart*, *box* and *filter* module are wrapped within a higher-level *simOut* module. This module includes a grid layout for the arrangement of the four sub-modules and a reference to the respective user-interface elements of each module. The server part of the *simOut* module receives several parameters such as spatial aggregated data, the associated geometry and a named list. Some of these inputs are provided as parameters for the sub-modules, others are used within function. The named list, for example, is used for transforming categorical values into adequate factors. Furthermore, it includes a function to aggregate data as well as the reference to each sub-module. The *simOut* module has been created for the evaluation of the simulation output. An outline of the *simOut* module is shown in Figure 4. It illustrates the link between the four sub-modules and the *simOut* module. Furthermore, a *simIn* module has been created for the evaluation of the simulation input data. Both main modules are used several times within the application.

```
#module user-interface
simOutUI <- function(id){

    #namespace definition
    ns <- NS(id)
     …
     #insert the user-interface parts of the corresponding modules
     filterUI(ns("filter"))
     mapUI(ns("map"))
     boxUI(ns("box"))
     chartUI(ns("chart"))
     …
 }

#module server
simOut <- function(input, output, session, data, geometry, named.list, …)
    …
    #call the server parts of the corresponding modules
    callModule(filter, "filter", …)
    callModule(map, "map", …)
    callModule(box, "box", …)
    callModule(chart, "chart", …)
}
```

Figure 4. Outline of the parent module *simOut*

E. Application frontend

The application frontend of TraVis is built with an extended dashboard version for Shiny applications, *shinydashboardPlus* [33]. It includes four parts: header, body, left and right sidebar. The header is currently used for displaying the application name and to toggle both sidebars. Further elements can be integrated. The left sidebar contains the main navigation of the application. Each menu item refers to an evaluation topic. Currently, parts of the population, vehicle fleet and stationary traffic from the simulation input can be visualized as well as daily mobility patterns which are obtained from the simulation output. The right sidebar is used as the main settings panel with several tabs. The first one enables the user to get access to the data. Here, it is possible to select one or more simulation runs belonging to a chosen research project. The spatial scale can be defined in the second tab. The third tab allows for defining the target language to be used throughout the entire session. Following the approach of multi-language support within a Shiny application [34], switching between English and German is currently implemented. The body includes the main panel. Depending on the chosen evaluation topic, the corresponding module *simIn* respective *simOut* is used.

VIII. Daily Mobility Patterns For Berlin

Daily mobility patterns can be described by three key indicators that are related to the selected mode of out-of-home travelers [35]: the number of trips, the distance and the travel time. Table V contains common daily mobility key indicators using a baseline scenario for the city of Berlin in 2010. Approximately 3.3 million inhabitants lived in 1.9 million households that year.

The result of a microscopic transport model allows working with even more detailed patterns of travel behavior. For instance, the daily mobility can be influenced by various socio-demographic factors such as age, gender and income. Especially, the interaction between the trip purpose and the used mode can be analyzed. But, also the spatiotemporal information is very important because these values can be used to point out spatial differences in the traffic volume during a day and to visualize traffic flows. The following figures illustrate various forms of representing daily mobility patterns within TraVis. The focus is on the visualization of the computed traffic volume grouped by transport mode for the city of Berlin in 2010. Within the application it is possible to visualize the spatial distribution for different spatial scales and for each transport mode as well as the activity made.

TABLE V. Key Indicators Of Daily Mobility Using A Baseline Scenario For The City Of Berlin In 2010

Mode	Number of Trips	Distance (km)	Time (min)
	Sum Values	*Average Values*	
Bike	1.469.023 *(12.5%)*	5.1	29
Car	3.625.876 *(30.9%)*	7.6	25
Public transport	3.210.767 *(27.4%)*	7.8	50
Walk	3.422.100 *(29.2%)*	1.8	26
Total	11.727.766	5.9	31

Figure 5 (A) shows the spatial distribution of all walking trips within the city. The total share of walking trips for the entire city is shown in Figure 5 (B). The modal split of two selected districts is represented in Figure 5 (C). These districts, one within the city center and the second further out, serve to compare the traffic volume during a day. While, the traffic volume in the morning is shown at the top and in the afternoon at the bottom for each district, see Figure 6 (A-B). In both districts the morning pick takes place at 6 and 7 o'clock. Throughout the day there are more trips within the inner city district which are made by walking respectively by using public transport. In the further out district, however, going by car dominates. In addition, it is evident that in the afternoon many more trips are undertaken within the inner city district.

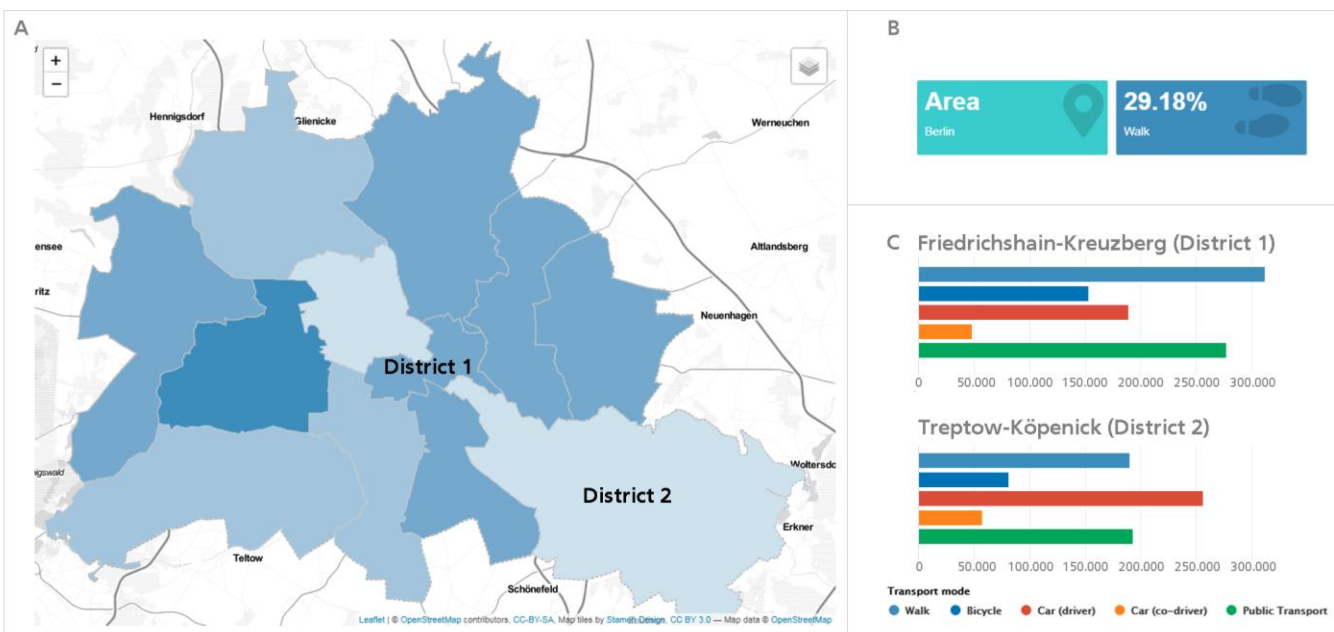


Figure 5. Visualization of the computed traffic volume for a selected mode of transport (walk). The spatial distribution within the city is shown in the map on the left side (A), while (B) represents the total share of walking trips for the entire city of Berlin. The modal split of two selected districts, one within the city center (Friedrichshain-Kreuzberg) and one further out (Treptow-Köpenick), is shown on the right-hand side (C).

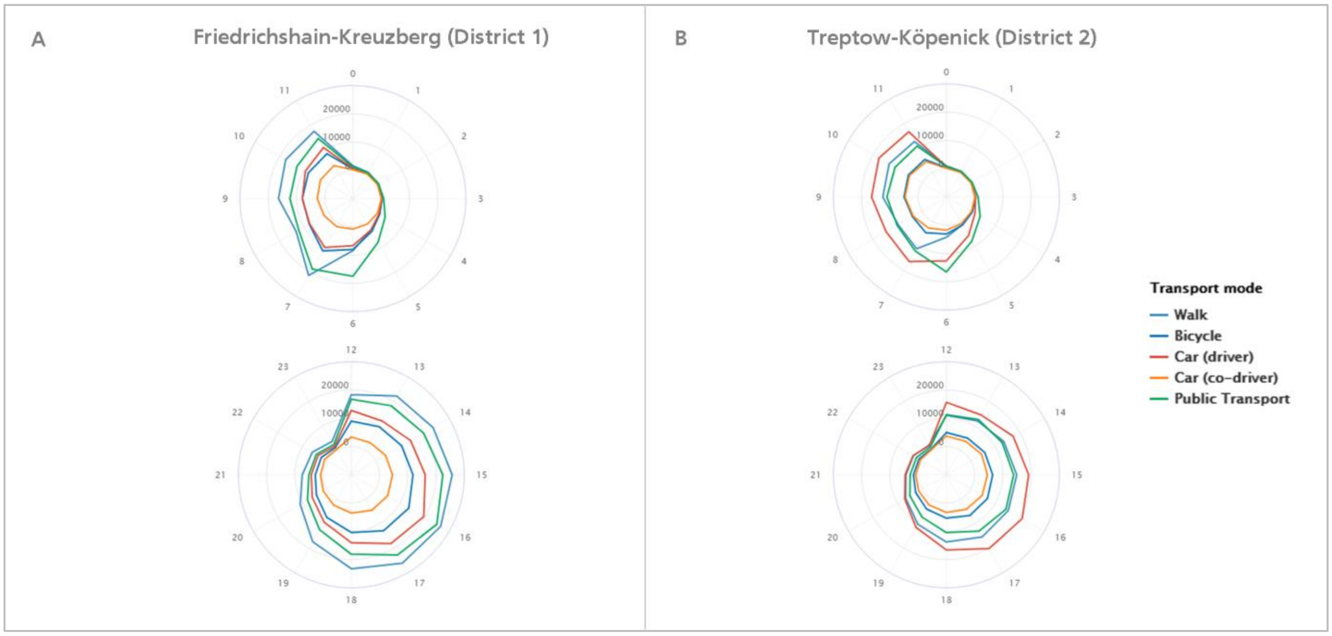


Figure 6. Visualization of the computed traffic volume during a day. Two districts within the city of Berlin are compared, one within the city center (A) and the second further out (B). The traffic volume in the morning is shown at the top and in the afternoon at the bottom.

IX. Summary and Furure Work

Transport demand models can provide answers to the questions of how, when, where and why people move from one location to another. Exploring the data of such a model can become very complex. This paper introduces the web-based application TraVis, which integrates several types of aggregation and disaggregation on a spatial, temporal, or demographic level. The application is implemented in the programming language R, which is widely used in the scientific field for statistical computing, data analysis and visualization. By using the Shiny web-framework it is possible to convert R analyzes into an interactive web application. Therefore, the application architecture was chosen. TraVis is currently used to analyze and communicate the simulation data of the transport model TAPAS, but is also suitable for evaluating real-time mobility observation data. In this case, legal issues such as data privacy related aspects have to be considered.

The presented approach could also be used to analyze research results within a different domain. Therefore, several recommendations and implementation notes are provided on how such application can be realized in R. This includes the use of a development environment, a usable application structure, a modular application design, the use of a reduced data volume, a concept for managing spatial data, to integrate HTML widgets for interactive visualization types, the support of multi-language, and advice on how to share the application with other users. Finally, the visualization of daily mobility patterns within the TraVis application is presented for a baseline scenario for the city of Berlin in 2010.

Upcoming work will focus on the visualization of spatiotemporal related data such as the individual travel trajectory and on implementing the ability to toggle between single and multi-window view. Furthermore, it is planned to make the application freely available as open source. TraVis is currently intended for internal use only, but it will be accessible from the following repository in the near future: *https://github.com/DLR-VF/TraVis*.

References

[1] A. von Schmidt, R. Cyganski, and M. Heinrichs, “Shine on Transport Model Simulation Data: Web-based Visualization in R using Shiny”, in DATA ANALYTICS 2018, The Seventh International Conference on Data Analytics, pp. 67-72, 2018, ISBN 978-1-61208-681-1, ISSN 2308-4464.

[2] M. Heinrichs, D. Krajzewicz, R Cyganski, and A. von Schmidt, “Disaggregated car fleets in microscopic travel demand modelling”, 7th International Conference on Ambient Systems, Networks and Technologies, pp. 155-162, 2016, doi: 10.1016/j.procs.2016.04.111.

[3] M. Heinrichs, D. Krajzewicz, R Cyganski, and A. von Schmidt, “Introduction of car sharing into existing car fleets in microscopic travel demand modelling”, Personal and Ubiquitous Computing, Springer, pp. 1055-1065, 2017 doi: 10.1007/s00779-017-1031-3.

[4] N. Yau, “Visualize this: the flowingdata guide to design, visualization, and statistics”, Wiley Publishing, 2011, ISBN: 978-0-470-94488-2.

[5] R Core Team, “R: A language and environment for statistical computing”, R Foundation for Statistical Computing, Vienna, Austria, 2018. http://www.R-project.org, accessed: 2019.11.26

[6] W. Chang, J. Cheng, J. J. Allaire, Y. Xie, and J. McPherson, “Shiny: Web Application Framework for R”, 2017. https://CRAN.R-project.org/package=shiny, accessed: 2019.11.26

[7] J. L. Bowman, M. A. Bradley, and J. Gibb, “The Sacramento Activity-based Travel Demand Model: Estimation and Validation Results”, presented at the European Transport Conference, 2006.

[8] R. M. Pendyala, R. Kitamura, A. Kikuchi, T. Yamamoto, and S. Fujii, “Florida Activity Mobility Simulator: Overview and Preliminary Validation Results”, Transportation Research Record (1921), pp. 123-130, 2005.

[9] X. Liu, W. Y. Yan, and J. Y. Chow, “Time-geographic relationships between vector fields of activity patterns and transport systems”, in Journal of Transport Geography, 42, pp. 22-33, 2015.

[10] O. Klein, “Visualizing Daily Mobility: Towards Other Modes of Representation”, A. Banos and T. Thevenien (Eds.), Geographical Information and Urban Transport Systems, Wiley Online Library, pp. 167-220, 2013.

[11] H. Guo, Z. Wang, B. Yu, H. Zhao, and X. Yuan, “Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection”, Visualization Symposium (PacificVis), IEEE Pacific, pp. 163-170, 2011, ISBN: 978-1-61284-935-5.

[12] R. Cyganski, A. von Schmidt, and D. Teske, “Applying Geovisualisation to Validate and Communicate Simulation Results of an Activity-based Travel Demand Model”, GI_Forum, in Journal for Geographic Information Science. pp. 575-578, 2015.

[13] M. Bostock, V. Ogievetsky, and J. Heer, “D3: Data-Driven Documents”, IEEE Transactions on Visualization and Computer Graphics, IEEE Press, 17 (12), pp. 2301–2309, 2011, doi: 10.1109/TVCG.2011.185.

[14] V. Agafonkin, “Leaflet: a JavaScript library for interactive maps”, 2011. Available from: https://leafletjs.com, accessed: 2019.11.26

[15] Highsoft, “Highcharts: interactive JavaScript charts for your webpage”, 2009. https://www.highcharts.com, accessed: 2019.11.26

[16] M. Loidl et al., “GIS and Transport Modeling-Strengthening the Spatial Perspective”, ISPRS International Journal of Geo-Information, 5(6), 2016, doi: 10.3390/ijgi5060084.

[17] A. Abela, “Advanced Presentations by Design: Creating Communication That Drives Action”, Pfeiffer, 2nd edition, 2013, ISBN: 978-1-118-34791-1.

[18] Ferdio, “Data Viz Project”. http://datavizproject.com, accessed: 2019.11.26

[19] A. von Schmidt, R. Cyganski, and D. Krajzewicz, “Generation of synthetic populations for transport demand models, a comparison of methods taking Berlin as an example”, “Generierung synthetischer Bevölkerungen für Verkehrsnachfragemodelle, ein Methodenvergleich am Beispiel von Berlin” (original title), In HEUREKA'17 - Optimierung in Verkehr und Transport, FGSV-Verlag, pp. 193-210, 2017.

[20] INSPIRE, http://inspire.ec.europa.eu, accessed: 2019.11.26

[21] J. Cheng, “Modularizing Shiny app code”, 2017. https://shiny.rstudio.com/articles/modules.html, accessed: 2019.11.26

[22] R. Bivand, T. Keitt, and B. Rowlingson, “rgdal: Bindings for the 'Geospatial' Data Abstraction Library”, R package version 1.4-4, 2019. https://CRAN.R-project.org/package=rgdal, accessed: 2019.11.26

[23] S. Chamberlain and A. Teucher, “geojsonio: Convert Data from and to 'GeoJSON' or 'TopoJSON'”, R package version 0.7.0, 2019, https://CRAN.R-project.org/package=geojsonio, accessed: 2019.11.26

[24] D. Bucklin and M. Basille, “rpostgis: linking R with a PostGIS spatial database”, in The R Journal, 2018, 10(1), pp. 251-268. https://journal.r-project.org/archive/2018/RJ-2018-025/index.html, accessed: 2019.11.26

[25] E. Pebesma and R. Bivand, “Classes and methods for spatial data in R”, in R News, 5 (2), pp. 9 - 13, 2005.

[26] E. Pebesma, “Simple Features for R: Standardized Support for Spatial Vector Data”, in The R Journal, 10 (1), pp. 439-446, 2018, doi: 10.32614/RJ-2018-009.

[27] H. Wickham, “tidyverse: Easily Install and Load the 'Tidyverse'”, R package version 1.2.1, 2017. https://CRAN.R-project.org/package=tidyverse, accessed: 2019.11.26

[28] T. Giraud and N. Lambert, “cartography: Create and Integrate Maps in your R Workflow”, in JOSS, 1(4), 2016, doi: 10.21105/joss.00054.

[29] H. Wickham, “ggplot2: Elegant Graphics for Data Analysis", Springer-Verlag New York, 2016.

[30] M. Tennekes, “tmap: Thematic Maps in R”, in Journal of Statistical Software, 84(6), pp. 1–39, 2018, doi: 10.18637/jss.v084.i06.

[31] J. Cheng, B. Karambelkar, and Y. Xie, “leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library”, R package version 2.0.2, 2018. https://CRAN.R-project.org/package=leaflet, accessed: 2019.11.26

[32] J. Kunst, “highcharter: A Wrapper for the 'Highcharts' Library”, 2019. https://github.com/jbkunst/highcharter, accessed: 2019.11.26

[33] D. Granjon, “shinydashboardPlus: Add More 'AdminLTE2' Components to 'shinydashboard'”, R package, 2019. https://CRAN.R-project.org/package=shinydashboardPlus, accessed: 2019.11.26

[34] C. Ladroue, “Multilingual Shiny App”, 2014. https://github.com/chrislad/multilingualShinyApp, accessed: 2019.11.26

[35] S. Schönefelder and K. W. Axhausen, “Urban Rhythms and Travel Behaviour: Spatial and Temporal Phenomena of Daily Travel”, Transport and Society, Ashgate, 2010, ISBN: 978-0-7546-7515-0.

How Social Media Factors Influence User's Travel Purchase Intention

Stavros D. Kaperonis
Communication, Media and Culture
Panteion University
Athens, Greece
email: skap@panteion.gr

Abstract—**This research aims to investigate the impact of specific factors on user's travel purchase intention and attitude in Social Media. Data collected from Social Media users to measure the relationship between the above factors of Social Media and user attitude on travel purchase intention with the Structural Equation Model. The primary purpose of the research is to explore whether the following Social Media factors are positively related to user attitude and travel purchase intention. The factors are source credibility, information reliability, user pleasure, and perceived value in travel services information. The findings of the research showed that there is a positive relationship between pleasure, attitude, and travel purchase intention, a positive relationship between reliability and customer attitude, and a positive relationship between perceived value and travel purchase intention. We also found that there is no relationship among credibility, user attitude, and travel purchase intention and an additional factor concerning atmospherics, positively affecting the perceived value and user pleasure. Finally, we found that customer attitude plays a vital role in mediating the relationship between credibility, pleasure, reliability, value, and travelers' purchase intention.**

Keywords-Social media; customer attitude; purchase intention; source credibility; pleasure; information reliability; perceived value; web atmospherics; SEM.

I. INTRODUCTION

The impact of social media on user's travel purchase intention through advertising from tour operators, social friends, and influencers, is more than ever timely [1][2]. Social Media (SM) users outran 2.70 billion in 2019, globally. Facebook was the most popular social network, with 2.38 billion active users per month in the first quarter of 2019. In 2018, daily Internet users spent 136 minutes on Social Media. In 2019, 96 percent of the active Facebook users per month, accessed via mobile devices [3].

Currently, Information Technology has reinforced Social Media, and the words "connecting" and "exchanging" have been replaced by the words “searching” and “selling” [4]–[6]. Tourist industry and hospitality has also become an essential tool for accessing different sources of tourism [5][6]. Users trust the Internet to obtain information. Studies have shown that usefulness and loyalty play a decisive role in user behavior, as far as understanding the information is concerned. Also, the usability of travel services has a positive influence on the consumer's loyalty and integrity [9]. The Internet has conquered the travel industry. Younger generations, especially Gen Y, are much more active in planning trips; they send and receive information via a variety of sources, including mobile devices (e.g., videos, Social Media). They make online reservations and plan potential destinations to visit. Users seek to be part of a wide range of travel experiences, and they are more responsive to online advertising. Social Media and mobile devices support these new ways of expression.

The purpose of the research is to study the interaction of the characteristics of source reliability, information reliability, perceived value, and user pleasure to purchase travel services. We are investigating the effect of source credibility features, information credibility, perceived value, and user pleasure, while browsing on social media to understand the attitude of the user when searching for travel destinations.

The paper is structured as follows: Section II presents the research background. Section III presents the research methodology and hypotheses. Section IV contains the conclusion of our research study.

II. RESEARCH BACKGROUND

Today SM networks, like Facebook and Instagram, allow people from different locations to interact and develop relationships or share travel experiences (e.g., posting photos and videos, sharing context) [10]. This information can be beneficial to potential travelers and can be personalized [11]. Most users install SM apps on their smart devices, and they use it as a tool for finding more travel information, with search engines providing direct access [12]. Researchers, [13] found that purchase intention is one dimension of customer behavior. Behavior is assessed through purchase intention, and consumers' behavioral patterns are examined [14]. Behavior is correlated to purchase intention [11][12], and this relationship has been empirically tested on the tourist industry [13][14], showing that customers' information reliability and satisfaction constitute an essential factor of e-behavioral intentions. Website design and information quality are essential for user satisfaction. There are many theories about value, such as consumption value, service value, consumer value, and perceived value [18]–[21]. When we talk about the perceived value [18][22], we refer to consumers' perception, price, and quality of a product, evaluating cost and benefit factors. To proceed with the conceptual research model presented in Figure 1, via Structural Equation Model (SEM), we measured the following four variables: source credibility, pleasure,

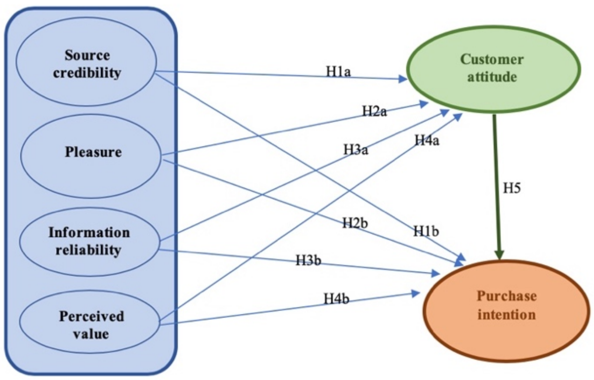


Figure 1. Theoretical framework and hypotheses

information reliability, and perceived value, in interaction with users' traveling attitude and purchase intention.

In order to confirm our Research Model, we used SEM, which incorporates the confirmatory approach, needed to justify our hypotheses. SEM uses confirmatory analysis rather than exploratory analysis for data. Confirmatory Factor Analysis (CFA) assess the measurement model validity, which compares the theoretical measurement with an actual model. SEM provides precise estimates of the errors in our parameters. As a first step, we are going to find which elements of each factor play an important role in user attitude and travel intention by using exploratory factor analysis (EFA), reliability analysis, and linear regression. However, the linear regression ignores mistakes and could lead us to severe inaccuracies, especially if mistakes are significant. Such methods are avoided by using SEM. SEM can incorporate both measurable and obscure variables. SEM method is preferred because it estimates the multiple and interrelated dependency in a single analysis.

The "state of the art" in this research is that investigates the role of credibility, pleasure, reliability, and value on customer attitude when browsing on social media searching for travel information, covering a gap in the field of tourism industry. This research wants to be a useful guide for tour operators, web developers, and advertising companies to design better platforms and creating more attractive travel products or services.

III. RESEARCH METHODOLOGY AND HYPOTHESES

Our study examines the factors that affect travelers' purchase intention and attitude in SM networks, [23]–[25]. In mental accounting theory, travel information mainly examines the context of SM use, rather than users' technology knowledge [15][16]. Based on the perceived value and the usage of SM in the tourism sector, we developed our research model, Figure 1. In this context, perceptual value is defined in terms of quality and price-performance or cost-benefit. Variables are categorized as source credibility, information reliability, pleasure, and perceived value. Also, perceived value leads to the usage of SM to search for travel information as a consequence of value perception.

Travelers' purchase intention defined in terms of credibility, pleasure, information reliability, and perceived value as a trade-off between costs and benefits [15][16][23]. The confirmation of our model, Figure 1, will be a useful guide for tour operators knowing the online factors, which are crucial to the user's travel decision. In this way, we are expanding previous researches on the use of SM by travelers in order to enhance our understanding of how travelers choose travel destinations through SM and in which way these four essential variables affect traveler purchase intention and interact with customer attitude [27].

A. Source credibility

Source credibility is defined as the factor upon which information is perceived as believable and trustworthy by users [28]. Source credibility constitutes an essential factor in decision-making procedures and involves high levels of risk [29]. Thus, source credibility is relevant to the context and the information acquired following computer and user interaction [28]. Researchers have found that users used SM before traveling [30]. Users are looking for information and ideas in SM for their travel, accommodation, and activities [30][31]. Therefore, SM provides travelers with all the information they need to make their travel dreams come true, thus reducing the risks it may pose [32]. Findings have shown that travelers use SM to obtain information from other users [33]. In word of mouth communication (WOM), reliability, and trust are essential factors for the consumer to accept the receiving message [34][35]. Recent researches have shown that the credibility of a source plays a vital role in making a piece of information reliable [36], [37]. The research of [38] showed a weak or insignificant relationship between source trust and intention to use user-generated travel content. Social Media requires users to create profiles and therefore provide personal information. Besides, most SM and especially those related to tourism (TripAdvisor), provide rating systems so that each user can submit his / her positive or negative reviews [39]. Moreover, TripAdvisor provides a distinct evaluation system for each service, thus enabling the user to get more detailed and specific information about the tourist service [40].

Researches have shown that source credibility has a significant impact on user attitude as an external factor because it affects human judgment [41]. Source credibility influences persuasion when evidence is fuzzy. In this case, hands-on processing can partially become cognitive processing [42]. For example, celebrities are one type of exogenous factor, which may enhance source credibility by influencing users' judgment [41].

H1a. Source credibility has a positive influence on customer attitude

H1b. Source credibility has a positive influence on purchase intention

B. Pleasure

Pleasure has a significant effect on technology admittance that enforces the meaning of usefulness [43] and internal motivation that enhances the feeling of using a computer because it is enjoyable. [43] referred to enjoyment as the "extent to which the activity of using the computer is perceived to be enjoyable in its own right." When people use technology and feel pleased or joyful, they perceive technology as a contributory value, and they are willing to use it again and again. [44][26] stated that the meaning of perceived value incorporates two different values (utilitarian and hedonic). The hedonic value explained the entertainment and emotional value of shopping. Researchers have shown that pleasure positively affects perceived value [15][17] and the intention of using hedonic information [45]. The pleasure as a feeling derives from emotions such as love, joy, and enjoyment [46]. Users receive satisfaction from experiences and feelings and seek pleasure from the process experienced during browsing [47][48].

Web atmospherics are playing a significant role in user satisfaction and pleasure. Elements like design, simplicity, layout, colors, video, are playing a substantial role in making the user feel pleasant and affect him in his purchase decision [49]. According to [49], designers must use beautiful and nice-looking pictures, and animations and colors, which are distinctive, visually appealing, and thus, design features should be separated from information contents [50]. Emotional is linked to web atmospherics but also is related to the decision process of the consumer on tour operators' websites [51]. According to [49], graphic design contributes to pleasing and arousing customers when they are engaged in searching for information and purchasing products or services provided by online stores.

The feelings of pleasure created by using SM apps encourage travelers not only to look up information on travel destinations but also to interact with other users. Travelers interact with each other by sharing photos or videos [8]. We, therefore, assume the following:

H2a. Pleasure has a positive influence on customer attitude

H2b. Pleasure has a positive influence on purchase intention

C. Information reliability

Studies on marketing have shown that consumer preferences are driven by value. Consumers are people of essence who seek to maximize usage [26]. The Internet provides travelers with various choices of many possible destinations to visit and make users pursue information reliability through a strenuous information search [52]. Direct access to alternative sources of information through SM builds trust between electronic word-of-mouth (eWOM) users and expert travel agents. The combination of convenient search and information reliability helps travelers to search, and evaluate a destination, and study new experiences related to a trip. The reliability of information is considered to be a significant factor for the traveler to perceive value when using SM [8].

Travelers prefer searching for reliable and credible information provided by the interaction between users of SM rather than obtaining the information through travel websites [8]. Travelers use SM networks like Facebook or Instagram, which are connected to User Created Contents (UCC) travel destinations, to share their experiences (e.g., photos, videos). By doing so, some travelers evaluate this reliable and credible information about a trip, thus reflecting their desire to engage online — users study reviews to decide on their travel destinations. We should not forget that the choice of a travel destination is complicated, as it involves factors such as accommodation, transport, restaurants, attractions, and many more. The diversity of each traveler and their different needs have not been included in these factors. Therefore, SM should be able to provide information relevant to the different needs of each user. When it comes to information reliability, we mean it is free of advertisements and hotel promotion as it is based on real user experience [39]. Therefore, if travelers find that the information is reliable and useful, they trust SM more because they are confident, they come from real customers rather than from biased sources of information.

The structure of information reliability is akin to the source information concept of information quality [53], which constitutes the output characteristics of the accuracy, timeliness, and completeness offered by the source of information. Quality of information of a travel destination has become a driving force on user decision making [54] assuming that the reliability of information influences purchase intention, hypothesizing the following:

H3a. Information reliability has a positive influence on customer attitude

H3b. Information reliability has a positive influence on purchase intention

D. Perceived value

The perceived value theory has also been adopted in travel destinations and shows high levels of influence on the intention of travelers to discover new or similar destinations, [55][56] showing that, in cruise travel services, emotional factors are essential in the perceived value. Cruise vacationers' behavior is influenced by the factors of hedonics or pleasure within the perceived value. The perceptual value in the tourism industry has been the subject of interest among many researchers. Some researchers have studied the perceived value of the received benefits and the sacrifices made by consumers [57][58]. Tourism and consumer decisions are associated with desires, fantasies, and emotions. The estimated value is subjective and varies among different tourists and cultures, also taking the emotional reactions of consumers into account [59]–[62]. Another finding on the perceptual value of [56] is about the pleasure and emotional reactions that cruise travelers get out of cruising.

Travelers evaluate the travel information in SM based on their perceptions of what they are willing to achieve and what to sacrifice. Perceived value involves a balance between costs and benefits and interaction between customer and service [63][43]. Analyzing the cost-benefit theory, the discrimination of perceived ease of use and perceived usefulness is similar between product performance and the effort of using the product. In high levels of perceived value in Social Media, travelers are likely to use a travel information search whereas, in the low levels of the perceived value, travelers show higher resistance toward travel information searches in SM [54]. When travelers search for information, they are more likely to select or to reject it based on the perceived benefits and the associated sacrifices of use, according to [43].

H4a. Perceived value in travel services information from Social Media has a positive influence on customer attitude

H4b. Perceived value in travel services information from Social Media has a positive influence on purchase intention

E. Customer Attitude

Consumer behavior is an integral part of marketing a successful business [64][65]. It has shown a positive relationship between consumer behavior and purchase intention [66][67]. The positive attitude of the consumer towards a product or service can positively contribute to purchase intention, especially on social networks [68]. According to all the above, we have the following hypothesis:

H5. Customer attitude has a positive influence on purchase intention

F. Data collection

The data for this study has been collected through a convenient sample e-survey from SM users. The e-survey was sent via a Facebook link to over 500 user profiles. The number of responders was 404 users. The age range of the responders was between 18 and over 54. The sample was primarily composed of individuals aged 18-44 years (98.1% of the sample) from European countries, mainly from Greece. Research [69] showed that over 70% in Greece of this age group (18-44) navigate on social media; therefore, it cannot be considered as a limitation (Table I).

The e-survey was divided into three sections. The first section outlines users' habits on SM and aims to discover the most important criteria for purchasing products or services for them. The second section outlines the factors that influence user travel purchase decisions. In order to complete the survey, the third section collected data for classification and statistical processing.

TABLE I. SOCIO-DEMOGRAPHIC CHARACTERISTICS OF THE RESPONDENTS

Dimension	Items	Percentage
Gender	F	23.8
	M	76.2
Age	18-24	84.7
	25-34	10.9
	35-44	2.5
	45-54	1.5
	54>	0.5
Income	0-500	61.9
	501-1000	18.8
	1001-2000	2.5
	2000>	0.5
	No answer	16.3

G. Data analysis

This research study adopted Structural Equation Modeling (SEM) to test the hypotheses. By using SEM, we want to evaluate our proposed model, analyze and explain the collected data [70]. All variables can be directly observed and thus qualify as manifest variables, called path analysis. In SEM terms, y, as in (1), (2), (3), (customer attitude and purchase intention), encloses the endogenous variables and χ (credibility, pleasure, reliability, and value) encloses the exogenous variables [71]. Variables that are influenced by other variables in a model are called endogenous variables. Variables that are not influenced by other variables in a model are called exogenous variables. Covariances, such as the one between χ_1, χ_2, χ_3, and χ_4, as in (1), (2), are represented by two-way arrows, Figure 2. One-way arrows represent paths acting as a cause. Each effect of source credibility, pleasure, information reliability, and the perceived value can be separated and are said to be related to customer attitude and purchase intention (Table II). The structural equations for this model are:

$$y_1=\gamma_{11}\chi_1+\gamma_{12}\chi_2+\gamma_{13}\chi_3+\gamma_{14}\chi_4+e_1 \quad (1)$$

$$y_2=\gamma_{21}\chi_1+\gamma_{22}\chi_2+\gamma_{23}\chi_3+\gamma_{24}\chi_4+e_2 \quad (2)$$

$$y_2=\psi_{21}y_1+e_3 \quad (3)$$

In our proposed research, we can see a model with two y variables and four χ variables. The multiple dependent variables, covariances, and variances of the exogenous factors x's are given and estimated by the values of the sample. As a result, it is complicated for the model to be falsified. Freedom degrees of our model counts the elements in the Φ matrix containing four values of γ and one of ψ.

Thus, there are precisely as many free parameters as there are data points. The parameters create transformations of the data. In our SEM analysis, we include all individual items that load their relationships, variances, disturbance, or errors onto the observation variables.

In our research, the metrics were designed after taking into consideration studies for our four manifest variables, source credibility, information reliability, pleasure, and perceived value. For the manifest variable of credibility, we used ten measurement scales [72][73][74]. For the manifest variable of pleasure, we used sixteen (including items for the

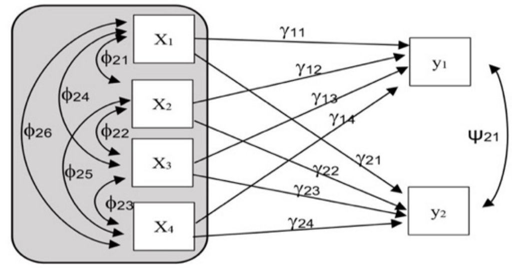


Figure 2. SEM research model

web atmospherics of the Social Media platforms) measurement scales [75][76][77]. For the manifest variable of information reliability, we used twelve measurement scales [33][78], and finally, for the manifest variable of the perceived value, we used nine measurement scales [15][16]. For every single item, we use multi-measurement items to overcome the limitations.

Because every single item has a high rate of measurement errors, we usually aim to capture all the attributes of a structure. All of these 47 items were measured on 5-point Likert scales, ranging from strongly disagree (1) to strongly agree (5).

The factor analysis drives us to erase some of the 47 items because either they load in many factors or their loading values are very low. Next, we noticed that the items of the factor pleasure split into two different factors. As a result of this, we named the new variable web atmospherics. The first important step in factor analysis is the normality of the data, which has been tested with the measurement of Kaiser-Meyer-Olkin, providing a value of 0.903, indicating a great fit of our model [79]. In factor analysis, the pattern matrix loads five factors, credibility, atmospherics, pleasure, reliability, and value. The variable credibility loads seven items, and the variable pleasure loads four items, the variable reliability loads four items, the variable value loads three items, and the variable web atmospherics loads eight items. Also, in the results, the variable pleasure was divided into two factors: pleasure and the new factor named web atmospherics (Table III). The first five factors together account for 72.66% of the total variance.

In linear regression, the model summary showed that R^2 is explained by 56.8% of the population when the role of dependent value played by purchase intention and the predictors are the items of credibility, pleasure, reliability, value, and atmospherics. For the dependent value customer attitude with the same predictors as the above, the R^2 explained by 49.3% of the population.

In reliability analysis, we found that Cronbach's alpha, for each one of the five factors, is very high, supporting the importance of each factor for our model (Table IV). In correlation analysis, findings give us α potential rendition of

TABLE II. MODEL VARIABLES DESCRIPTION

Variables	Description
x_1	Credibility
x_2	Pleasure
x_3	Reliability
x_4	Value
y_1	Intention to Purchase
y_2	Customer Purchase attitude

TABLE III. THE LOADING FACTORS

Pattern Matrix					
	Factors				
	1	2	3	4	5
Credibility					
The buying process must be clear	0.945				
Payment/debit methods must be clear	0.936				
All the steps to the purchase should be simple	0.920				
Product or service instructions must be clear before purchase	0.865				
The perceived service is significant	0.849				
The buying process should be quick	0.830				
The process should inspire confidence	0.789				
Web atmospherics					
Fonts in social media are significant		0.916			
Special effects in social media are significant		0.861			
The unique design of social media platforms is significant.		0.858			
The layout of social media is significant.		0.850			
The colors in social media are significant		0.805			
The pleasant design of social media is significant.		0.794			
Videos in social media are significant.		0.742			
The simplicity of designing social media platforms is essential		0.697			
Pleasure					
Searching for information on social media is enjoyable.			0.916		
Searching for travel information on social networks is fun.			0.903		
Searching for travel information through social media is fascinating.			0.819		
Searching for travel information through social media is a pleasant experience.			0.792		
Reliability					
The authors of travel information in social media is usually trustworthy				0.939	
The authors of travel information in social media is usually reliable.				0.914	
Professionals usually write travel information on social media for travel destinations.				0.775	
Travel information in social media is usually written by someone who knows well the destinations				0.739	
Value					
The perceived value of a product or service influences the purchase decision.					0.805
The price of the products or services influences the purchase decision					0.694
The quality/variety of the products or services influences the purchase decision					0.651

the impact of some social factors on user attitude while searching for travel experiences (Table V).

In this research, we test and estimate our model through SEM. Two different approaches used for testing our research hypotheses. The first approach is exploratory factor analysis (EFA), with linear regression and reliability analysis. After the validation of the factors, we use the SEM to test the validity of the proposed model and hypotheses (Table VI). For the validity of our model, we tested the goodness-of-fit, [80], assisted by the goodness-of-fit index (GFI) [81], adjusted goodness-of-fit index (AGFI) [80], comparative fit index (CFI) [82], and root mean square error of approximation RMSEA [83].

After running many different models and taking into account all those parameters, improving our model, and eliminating all the possible errors, we found the following results, Figure 3. The GFI, AGFI, and CFI have values between 0.893 and 0.951, indicating an absolute fitting model. Also, the RMSEA with a value 0.073 is supporting the excellent fit of our model (Table VII) [45][46].

TABLE VI. THE LITERATURE SUPPORT FOR THE RESPECTIVE FITNESS INDEX

Index	Index full name	Literature
Chi-Square	Discrepancy Chi-Square	[86]
RMSEA	Root Mean Square of Error Approximation	[87]
GFI	Goodness of Fit Index	[88]
AGFI	Adjusted Goodness of Fit	[89]
CFI	Comparative Fit Index	[90]
TLI	Tucker-Lewis Index	[82]
NFI	Normed Fit Index	[91]
Chisq/df	Chi-Square/Degrees of Freedom	[92]

TABLE V. CORRELATIONS

		Customer Attitude
Social media information helps me in my travel decision	Pearson Correlation	**.847**
	Sig. (2-tailed)	.000
	N	404
Social media could change my first decision	Pearson Correlation	**.782**
	Sig. (2-tailed)	.000
	N	404
Information and advertising help me to decide to book a trip	Pearson Correlation	**.660**
	Sig. (2-tailed)	.000
	N	404
Product or service usability, time and energy I need to spend for my travel decision is important	Pearson Correlation	**.641**
	Sig. (2-tailed)	.000
	N	404
My social friends travel suggestions affect my buying decision	Pearson Correlation	**.629**
	Sig. (2-tailed)	.000
	N	404
An influencer can affect my attitude about a travel service or product	Pearson Correlation	**.609**
	Sig. (2-tailed)	.000
	N	404
The provided information through social media affect me on about a travel service or product	Pearson Correlation	**.599**
	Sig. (2-tailed)	.000
	N	404
Social media comments affect me in my decision for a travel service or product	Pearson Correlation	**.590**
	Sig. (2-tailed)	.000
	N	404

The overall model fit was measured using the chi-square, yielding a value of x^2/df=3.156 [93], the chi-squared value was 533.392, and all factor loadings were statistically significant (p=0.000).

TABLE IV. THE RELIABILITY STATISTICS FOR EACH FACTOR

	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
Credibility	.957	.958	7
Pleasure	.927	.928	4
Reliability	.910	.910	4
Value	.821	.824	3
Atmospherics	.943	.943	8

In the confirmed research model (Figure 3), we observe that the factor of credibility does not interact with the intention and attitude variables, the factor of pleasure interacts with the travel purchase intention and customer attitude variables, the factor of reliability interacts only with the variable attitude, and the factor of value does not interact with the travel purchase intention and customer attitude variables. We also find that the factor of web atmospherics plays a significant role in customer attitude, interacting with the pleasure and value variables. The credibility variable does not interact with the main variables. Travel purchase Intention and customer attitude play an essential role in the way the other variables behave.

IV. CONCLUSIONS

This article aimed to create a model by which we could interpret the effect that specific SM factors have on user purchase intention, as far as travel services are concerned. Studying the theoretical background, we noticed a research gap that led us to our hypotheses, which we elaborate on in the article. We analyzed the collected data from the e-survey, and we tested our research model validity by detecting the factors that influence the user purchase intention when traveling, after using Social Media. Linear regression, factor analysis, Correlation, and reliability analysis used to test the validity of our variables (Tables III, IV, V).

The results gave us a good fit for our Conceptual model, allowed us to go one step further and confirmed our model through SEM. SEM enabled us to check our assumptions and confirm our model. Using the SEM data analysis, we

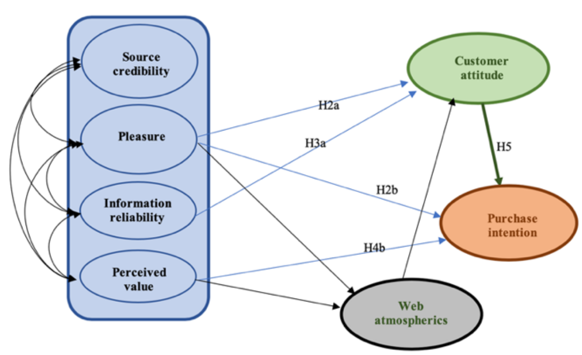


Figure 3. The confirmed research model

determined the critical factors concerning the customers' attitude towards travel purchase intention.

Furthermore, the SEM method showed us that the variable web atmospherics influences purchase intention in SM platforms, playing the role of a moderator. Also, findings confirmed that the following factors are related to the travel service purchase intention and customer attitude. Source credibility has no positive relationship with travel purchase intention and customer attitude (H1a, H1b, not confirmed). The factor of pleasure has a positive relationship with travel purchase intention and customer attitude (H2a, H2b confirmed). The reliability variable has a positive relationship only with customer attitude (H3a confirmed, H3b not confirmed), whereas the factor of value has a positive relationship with travel purchase intention (H4a not confirmed, H4b confirmed). Finally, the travel purchase intention variable, interacts with the customer attitude variable positively (H5 confirmed), Table VIII.

Thus, we can state that the pleasure, value, and reliability variables, play a critical role in customer attitude and purchase intention, when choosing a travel destination. Also, the credibility factor plays a significant role in positively interacting with the factors pleasure, reliability, and value, means that all the steps in the booking procedure must be simple, clear, and make you feel confident when you want to book a trip. The new variable, that is web aesthetics, has emerged, as it plays an essential role in traveling users' decision making, interacting with the pleasure and value variables, meaning that it plays the role of moderator in customer attitude.

TABLE VII. MODEL FIT AND THE LEVEL OF ACCEPTANCE

	Index	Level of acceptance
Absolute fit	Chi-Square	P-value<0.001
	RMSEA	0.073
	GFI	0.893
	AGFI	0.854
	CFI	0.951
	TLI	0.939
	NFI	0.930
	Chisq/df	3.156

TABLE VIII. FINAL RESULTS OF OUR HYPOTHESES

	Hypotheses	Confirmed or Not
Source credibility	**H1a.** Source credibility has a positive influence on customer attitude	No
	H1b. Source credibility has a positive influence on purchase intention	No
Pleasure	**H2a.** Pleasure has a positive influence on customer attitude	Yes
	H2b. Pleasure has a positive influence on purchase intention	Yes
Reliability	**H3a.** Information reliability has a positive influence on customer attitude	Yes
	H3b. Information reliability has a positive influence on purchase intention	No
Value	**H4a.** Perceived value in travel services information from Social Media has a positive influence on customer attitude	No
	H4b. Perceived value in travel services information from Social Media has a positive influence on purchase intention	Yes
Attitude	**H5.** Customer attitude has a positive influence on purchase intention	Yes

Correlation findings help us to understand the vital role of travel product or service information in advertising. Also, customer attitude is highly correlated with the provided travel information from social media platforms. Furthermore, social media information could change the first opinion or attitude of the user about a travel product or service. User attitude, on travel information, can be affected by the comments of their social media friends or a social influencer. The findings from correlation analysis is a useful guide for tour operators or marketers in order to focus on the travel information they provide, targeting the social friends or influencers of the potential customers. Accordingly, marketers are recommended to target users who like searching or booking trips through social media, interact with, and are influenced by their social friends.

Also, the findings suggest that designers and tour operators should concentrate on adapting, factors such as pleasure, reliability, which are playing a significant role in customer attitude and purchase intention to their platforms. The factor pleasure in terms of making the user feel good, enjoyable, and pleasant when searching for travel information, must be adapted from web developers, advertising companies, and tour operators in order to increase the visibility, the interaction, and the possibility of a user's purchasing a trip. That could be achieved it, targeted advertising on social media for travel service purchased and user intention to purchase such services. Furthermore, the factors, reliability, and value, make the user feel that the platform is trustworthy and improves the perception of the price and the quality of the product or service that the platform provides.

As future research could be the investigation of more variables on SM (esp. Instagram), playing a significant role in user travel decisions. Furthermore, possible correlations between users who want to present themselves in an idealized way, designed to impress or their desire to impress social friends on travel service purchase decisions, could be investigated. Future studies similar to the current study should incorporate social media user behavior, such as purchase attitude as a dependent variable. Finally, future studies will benefit from investigating the present model in more diverse cultural contexts and lead us to deepen our understanding of user behavior in social media coming from different cultures and will lead the marketers and tour operators to the creation of more personalized products and services.

REFERENCES

[1] S. D. Kaperonis, "The Impact of Social Media on user's Travel Purchase Intention," *Data analytics 2018, the Seventh International Conference on Data Analytics*, Athens, Greece: IARIA, The International Academy, Research and Industry Association, pp. 50–54.

[2] S. Bhulai, D. Kardaras, and I.Semanjski, Eds., "*Data analytics 2018, the Seventh International Conference on Data Analytics" IARIA, The Inetrnation Academy, Research and Industry Association,* Wilmington,

USA, 2018.

[3] "Global time spent on social media daily 2017 | Statista," *Statista*, 2018. [Online]. Available: https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/. [Accessed: 19-Jul-2019].

[4] S. Choi, "An empirical study of social network service (SNS) continuance: incorporating the customer value-satisfaction-loyalty model into the IS continuance model," *Asia Pacific J. Inf. Syst.*, vol. 23, no. 4, pp. 1–28, 2013.

[5] N. Chung, H. J. Han, and C. Koo, "Mediating roles of attachment for information sharing in social media: social capital theory perspective," *Asia Pacific J. Inf. Syst.*, vol. 22, no. 4, pp. 101–123, 2012.

[6] C. Koo, Y. Wati, and J. J. Jung, "Examination of how social aspects moderate the relationship between task characteristics and usage of social communication technologies (SCTs) in organizations," *Int. J. Inf. Manage.*, vol. 31, no. 5, pp. 445–459, 2011.

[7] R. Law, R. Leung, and D. Buhalis, "Information technology applications in hospitality and tourism: a review of publications from 2005 to 2007," *J. Travel Tour. Mark.*, vol. 26, no. 5–6, pp. 599–623, 2009.

[8] M. Sigala, E. Christou, and U. Gretzel, "Social Media in Travel," *Tour. Hosp. Theory, Pract. Cases*, 2012.

[9] A. Bilgihan, A. Barreda, F. Okumus, and K. Nusair, "Consumer perception of knowledge-sharing in travel-related online social networks," *Tour. Manag.*, vol. 52, pp. 287–296, 2016.

[10] E. Parra-López, J. Bulchand-Gidumal, D. Gutiérrez-Taño, and R. Díaz-Armas, "Intentions to use social media in organizing and taking vacation trips," *Comput. Human Behav.*, vol. 27, no. 2, pp. 640–654, 2011.

[11] D. K. Kardaras, S. Kaperonis, S. Barbounaki, I. Petrounias, and K. Bithas, "An Approach to Modelling User Interests Using TF-IDF and Fuzzy Sets Qualitative Comparative Analysis," in *IFIP International Conference on Artificial Intelligence Applications and Innovations*, 2018, pp. 606–615.

[12] Z. Xiang and U. Gretzel, "Role of social media in online travel information search," *Tour. Manag.*, vol. 31, no. 2, pp. 179–188, 2010.

[13] V. A. Zeithaml, L. L. Berry, and A. Parasuraman, "The behavioral consequences of service quality," *J. Mark.*, pp. 31–46, 1996.

[14] I. Ajzen and M. Fishbein, "Understanding attitudes and predicting social behaviour," 1980.

[15] R. L. Oliver and W. O. Bearden, "Disconfirmation processes and consumer evaluations in product usage," *J. Bus. Res.*, vol. 13, no. 3, pp. 235–246, 1985.

[16] F. Buttle and B. Bok, "Hotel marketing strategy and the theory of reasoned action," *Int. J. Contemp. Hosp. Manag.*, vol. 8, no. 3, pp. 5–10, 1996.

[17] I. Ajzen and B. L. Driver, "Application of the theory of planned behavior to leisure choice," *J. Leis. Res.*, vol. 24, no. 3, pp. 207–224, 1992.

[18] H.-W. Kim, H. C. Chan, and S. Gupta, "Value-based adoption of mobile internet: an empirical investigation," *Decis. Support Syst.*, vol. 43, no. 1, pp. 111–126, 2007.

[19] H.-W. Kim, Y. Xu, and S. Gupta, "Which is more important in Internet shopping, perceived price or trust?," *Electron. Commer. Res. Appl.*, vol. 11, no. 3, pp. 241–252, 2012.

[20] S. Gupta and H. Kim, "Value-driven Internet shopping: The mental accounting theory perspective," *Psychol. Mark.*, vol. 27, no. 1, pp. 13–35, 2010.

[21] S. Kim, S. Holland, and H. Han, "A structural model for examining how destination image, perceived value, and service quality affect destination loyalty: A case study of Orlando," *Int. J. Tour. Res.*, vol. 15, no. 4, pp. 313–328, 2013.

[22] V. A. Zeithaml, "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence," *J. Mark.*, pp. 2–22, 1988.

[23] R. H. Thaler, "Mental accounting and consumer choice," *Mark. Sci.*, vol. 27, no. 1, pp. 15–25, 2008.

[24] R. Thaler, "Mental accounting and consumer choice," *Mark. Sci.*, vol. 4, no. 3, pp. 199–214, 1985.

[25] R. Thaler, "Toward a positive theory of consumer choice," *J. Econ. Behav. Organ.*, vol. 1, no. 1, pp. 39–60, 1980.

[26] R. Sánchez-Fernández and M. Á. Iniesta-Bonillo, "The concept of perceived value: a systematic review of the research," *Mark. theory*, vol. 7, no. 4, pp. 427–451, 2007.

[27] S. Kaperonis, "Virtual networks: Why do students use Instagram?," *http://dx.doi.org/10.12681/homvir.20193*, vol. 2, no. 1, pp. 43–49, 2019.

[28] S. W. Sussman and W. S. Siegal, "Informational influence in organizations: An integrated approach to knowledge adoption," *Inf. Syst. Res.*, vol. 14, no. 1, pp. 47–65, 2003.

[29] B. Mak and K. Lyytinen, "A model to assess the behavioral impacts of consultative knowledge based systems," *Inf. Process. Manag.*, vol. 33, no. 4, pp. 539–550, 1997.

[30] C. Cox, S. Burgess, C. Sellitto, and J. Buultjens, "The role of user-generated content in tourists' travel planning behavior," *J. Hosp. Mark. Manag.*, vol. 18, no. 8, pp. 743–764, 2009.

[31] J. N. Fotis, D. Buhalis, and N. Rossides, *Social media use and impact during the holiday travel planning process*. Springer-Verlag, 2012.

[32] U. Gretzel and K. H. Yoo, "Use and impact of online travel reviews," *Inf. Commun. Technol. Tour. 2008*, pp. 35–46, 2008.

[33] A. Simms, "Online user-generated content for travel planning-different for different kinds of trips?," 2012.

[34] E. McGinnies and C. D. Ward, "Better liked than right: Trustworthiness and expertise as factors in credibility," *Personal. Soc. Psychol. Bull.*, vol. 6, no. 3, pp. 467–472, 1980.

[35] C. I. Hovland, I. L. Janis, and H. H. Kelley, *Communication and persuasion: Psychological studies of opinion change*. Yale University Press, 1953.

[36] L. M. Willemsen, P. C. Neijens, F. Bronner, and J. A. De Ridder, "'Highly recommended!' The content characteristics and perceived usefulness of online consumer reviews," *J. Comput. Commun.*, vol. 17, no. 1, pp. 19–38, 2011.

[37] C. M. K. Cheung, M. K. O. Lee, and N. Rabjohn, "The impact of electronic word-of-mouth: The adoption of online opinions in online customer communities," *Internet Res.*, vol. 18, no. 3, pp. 229–247, 2008.

[38] J. K. Ayeh, N. Au, and R. Law, "'Do we believe in TripAdvisor?' Examining credibility perceptions and online travelers' attitude toward using user-generated content," *J. Travel Res.*, vol. 52, no. 4, pp. 437–452, 2013.

[39] M. Y. Cheung, C. Luo, C. L. Sia, and H. Chen, "Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations," *Int. J. Electron. Commer.*, vol. 13, no. 4, pp. 9–38, 2009.

[40] R. Filieri, S. Alguezaui, and F. McLeay, "Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth," *Tour. Manag.*, vol. 51, pp. 174–185, 2015.

[41] A. Bhattacherjee and C. Sanford, "Influence processes for information technology acceptance: An elaboration likelihood model," *MIS Q.*, pp. 805–825, 2006.

[42] S. Chaiken and D. Maheswaran, "Heuristic processing can bias systematic processing: effects of source credibility, argument ambiguity, and task importance on attitude judgment.," *J. Pers. Soc. Psychol.*, vol. 66, no. 3, p. 460, 1994.

[43] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Q.*, pp. 319–340, 1989.

[44] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "Extrinsic and intrinsic motivation to use computers in the workplace 1," *J. Appl. Soc. Psychol.*, vol. 22, no. 14, pp. 1111–1132, 1992.

[45] H. Van der Heijden, "User acceptance of hedonic information systems," *MIS Q.*, pp. 695–704, 2004.

[46] M. B. Holbrook and E. C. Hirschman, "The experiential aspects of consumption: Consumer fantasies, feelings, and fun," *J. Consum. Res.*, vol. 9, no. 2, pp. 132–140, 1982.

[47] P. Mikalef, M. Giannakos, and A. Pateli, "Shopping and word-of-mouth intentions on social media," *J. Theor. Appl. Electron. Commer. Res.*, vol. 8, no. 1, pp. 17–34, 2013.

[48] B. J. Babin, W. R. Darden, and M. Griffin, "Work and/or fun: measuring hedonic and utilitarian shopping value," *J. Consum. Res.*, vol. 20, no. 4, pp. 644–656, 1994.

[49] D.-M. Koo and S.-H. Ju, "The interactional effects of atmospherics and perceptual curiosity on emotions and online shopping intention," *Comput. Human Behav.*, vol. 26, no. 3, pp. 377–388, 2010.

[50] C.-S. Wu, F.-F. Cheng, and D. C. Yen, "The atmospheric factors of online storefront environment design: An empirical experiment in Taiwan," *Inf. Manag.*, vol. 45, no. 7, pp. 493–498, 2008.

[51] P. Björk, "Atmospherics on tour operators' websites: Website features that stimulate emotional response," *J. Vacat. Mark.*, vol. 16, no. 4, pp. 283–296, 2010.

[52] Y.-H. Hwang and D. R. Fesenmaier, "Unplanned tourist attraction visits by travellers," *Tour. Geogr.*, vol. 13, no. 3, pp. 398–416, 2011.

[53] W. H. DeLone and E. R. McLean, "Information systems success: The quest for the dependent variable," *Inf. Syst. Res.*, vol. 3, no. 1, pp. 60–95, 1992.

[54] N. Chung and C. Koo, "The use of social media in travel information search," *Telemat. Informatics*, vol. 32, no. 2, pp. 215–229, 2015.

[55] A. S. Lo and C. Y. S. Lee, "Motivations and perceived value of volunteer tourists from Hong Kong," *Tour. Manag.*, vol. 32, no. 2, pp. 326–334, 2011.

[56] T. Duman and A. S. Mattila, "The role of affective factors on perceived cruise vacation value," *Tour. Manag.*, vol. 26, no. 3, pp. 311–323, 2005.

[57] J. Sanchez, L. Callarisa, R. M. Rodriguez, and M. A. Moliner, "Perceived value of the purchase of a tourism product," *Tour. Manag.*, vol. 27, no. 3, pp. 394–409, 2006.

[58] J. E. Bigné, L. Andreu, I. Küster, and A. Blesa, "Quality market orientation: tourist agencies' perceived effects," *Ann. Tour. Res.*, vol. 32, no. 4, pp. 1022–1038, 2005.

[59] J. C. Sweeney and G. N. Soutar, "Consumer perceived value: The development of a multiple item scale," *J. Retail.*, vol. 77, no. 2, pp. 203–220, 2001.

[60] N. K. Prebensen, E. Woo, J. S. Chen, and M. Uysal, "Motivation and involvement as antecedents of the perceived value of the destination experience," *J. Travel Res.*, vol. 52, no. 2, pp. 253–264, 2013.

[61] R. N. Bolton and J. H. Drew, "A multistage model of customers' assessments of service quality and value," *J. Consum. Res.*, vol. 17, no. 4, pp. 375–384, 1991.

[62] W. J. Havlena and M. B. Holbrook, "The varieties of consumption experience: comparing two typologies of emotion in consumer behavior," *J. Consum. Res.*, vol. 13, no. 3, pp. 394–404, 1986.

[63] A. Payne and S. Holt, "Diagnosing customer value: integrating the value process and relationship marketing," *Br. J. Manag.*, vol. 12, no. 2, pp. 159–182, 2001.

[64] G. J. Bamossy and M. R. Solomon, *Consumer behaviour: A European perspective*. Pearson Education, 2016.

[65] M. Hogg, S. Askegaard, G. Bamossy, and M. Solomon, *Consumer behaviour: a European perspective.* Prentice Hall, 2006.

[66] H. Ting, E. C. de Run, and R. Thurasamy, "Young adults' attitude towards advertising: A multi-group analysis by ethnicity," *Rev. Bras. Gestão Negócios*, vol. 17, no. 54, pp. 769–787, 2015.

[67] H. Ting, W. W. P. Ming, E. C. de Run, and S. L. Y. Choo, "Beliefs about the use of Instagram: An exploratory study," *Int. J. Bus. Innov.*, vol. 2, no. 2, pp. 15–31, 2015.

[68] M.-F. Chen, "Consumer attitudes and purchase intentions in relation to organic foods in Taiwan: Moderating effects of food-related personality traits," *Food Qual. Prefer.*, vol. 18, no. 7, pp. 1008–1021, 2007.

[69] "Social media users in Greece 2019." [Online]. Available: https://napoleoncat.com/stats/social-media-users-in-greece/2019/10. [Accessed: 14-Nov-2019].

[70] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, and R. L. Tatham, "Multivariate data analysis (Vol. 6)." Upper Saddle River, NJ: Pearson Prentice Hall, 2006.

[71] R. B. Kline, "Promise and pitfalls of structural equation modeling in gifted research.," 2010.

[72] S. Ha and J. Ahn, "Why are you sharing others' tweets?: The impact of argument quality and source credibility on information sharing behavior," 2011.

[73] M. Kang, "Measuring social media credibility: A study on a Measure of Blog Credibility," *Inst. Public Relations*, pp. 59–68, 2010.

[74] K. P. Hocevar, A. J. Flanagin, and M. J. Metzger, "Social media self-efficacy and information evaluation online," *Comput. Human Behav.*, vol. 39, pp. 254–262, 2014.

[75] J.-J. Wu and Y.-S. Chang, "Towards understanding members' interactivity, trust, and flow in online travel community," *Ind. Manag. Data Syst.*, vol. 105, no. 7, pp. 937–954, 2005.

[76] N. Chung, C. K.-T. and Informatics, and undefined 2015, "The use of social media in travel information search," *Elsevier*.

[77] J. Y. Kim, J. P. Shim, and K. M. Ahn, "Social networking service: Motivation, pleasure, and behavioral intention to use," *J. Comput. Inf. Syst.*, vol. 51, no. 4, pp. 92–101, 2011.

[78] M. J. Metzger, "Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research," *J. Am. Soc. Inf. Sci. Technol.*, vol. 58, no. 13, pp. 2078–2091, 2007.

[79] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate data analysis: Pearson new international edition*. Pearson Higher Ed, 2013.

[80] K. G. Jöreskog and D. Sörbom, *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International, 1993.

[81] L.-T. Hu, P. M. Bentler, and R. H. Hoyle, "Structural equation modeling: Concepts, issues, and applications," *Eval. Model fit*, pp. 76–99, 1995.

[82] P. M. Bentler and D. G. Bonett, "Significance tests and goodness of fit in the analysis of covariance structures.," *Psychol. Bull.*, vol. 88, no. 3, p. 588, 1980.

[83] J. H. Steiger, A. Shapiro, and M. W. Browne, "On the multivariate asymptotic distribution of sequential chi-square statistics," *Psychometrika*, vol. 50, no. 3, pp. 253–263, 1985.

[84] A. Diamantopoulos and J. A. Siguaw, "Introducing LISREL: A guide for the uninitiated," *London ua*, 2000.

[85] B. M. Byrne, "Structural equation modeling with LISREL," *Prelis, and Simplis*, pp. 196–199, 1998.

[86] B. Wheaton, B. Muthen, D. F. Alwin, and G. F. Summers, "Assessing reliability and stability in panel models," *Sociol. Methodol.*, vol. 8, pp. 84–136, 1977.

[87] M. W. Browne and R. Cudeck, "Alternative ways of assessing model fit," *Sage Focus Ed.*, vol. 154, p. 136, 1993.

[88] K. G. Joreskog and D. Sorbom, "LISREL VI: Analysis of linear structural relationships by the method of maximum likelihood. Mooresville, IN: Scientific Software." Inc, 1984.

[89] J. S. Tanaka and G. J. Huba, "A fit index for covariance structure models under arbitrary GLS estimation," *Br. J. Math. Stat. Psychol.*, vol. 38, no. 2, pp. 197–201, 1985.

[90] P. M. Bentler, "Comparative fit indexes in structural models.," *Psychol. Bull.*, vol. 107, no. 2, p. 238, 1990.

[91] K. A. Bollen, "(1989b). Structural equations with latent variables. New York: John Wiley," 1989.

[92] H. W. Marsh and D. Hocevar, "Application of confirmatory factor analysis to the study of self-concept: First-and higher order factor

models and their invariance across groups.," *Psychol. Bull.*, vol. 97, no. 3, p. 562, 1985.

[93] R. B. Kline, *Principles and practice of structural equation modeling*. Guilford publications, 2015.

Digital Adoption Strategy: A Public Sector Ecosystem

Samantha Papavasiliou and Carmen Reaiche
Adelaide Business School
University of Adelaide,
Adelaide, Australia
e-mail: samantha.papavasiliou@adelaide.edu.au,
carmen.reaiche@adelaide.edu.au

***Abstract*—The shift of public sector service provision to digital first, has had a considerable impact on how individuals interact with public sector entities. Therefore, this research argues for a systems approach to explore and understand different assistance-seeking behaviours. In particular, there is a need to understand the critical points within the system at which assistance is sought and the changing behaviours correlated to post assistance-seeking outcomes. Evidence for this research is presented through a case study on the Australian Taxation Office. Observations of the digital lodgement channel were undertaken to evaluate the components in which individuals sought assistance. Through the application of systems thinking and process mapping, this research highlights the critical points in which assistance was sought within the lodgement process. The results of this research lead to the recommendations that ongoing education should be provided for four years post first lodgement, and that education should occur at change of circumstances. Through the use of strategically placed self-help assistance throughout the lodgement process, it is anticipated that individuals will be less likely to seek assistance. However, this research demonstrates the importance of maintaining human interfaces for assistance-seeking to maximise an individual's capacity to interact with the system successfully.**

Keywords- Assistance seeking; Digital Ecosystems; public sector; digital lodgement.

I. INTRODUCTION

The ongoing attention on the importance of a successful public sector service offering, and to improve and enhance digital service adoption within the public sector, is based on the provision of a better client experience. To achieve this, public sector entities are putting increasing resources into understanding their clients and providing services to meet their expectations. This has had an impact on the manner in which services are designed, the manner in which data is used and shared to increase ease of use, and how entities plan future services. With the increased attention on the services, the next stage is to understand the human interaction components, which include assistance-seeking behaviours, why individuals seek assistance, and the points in the process that are more likely to require assistance. Since the shift to digital first service policies in the public sector, increased pressure has been placed on both the service user and provider to understand both the process and the digital environment [1]. Even with high adoption rates of digital services, there are still individuals not interacting digitally and those who require ongoing support and assistance once going digital.

The aim of this paper is to identify and understand public sector service users, which includes identifying the potential barriers to digital adoption, followed by understanding why various users seek assistance, and post assistance-seeking outcomes. This paper argues that mandatory public sector services need to be inclusive, including digital and non-digital options, this extends to how assistance is provided. This research forms part of a PhD and findings and recommendations are ongoing.

Research by the Australian Digital Transformation Agency (DTA) [2] indicates the necessity of further research of this type, to proactively address digital adoption, specifically increased use. In Australia the purpose of the DTA is to improve services for the community, expecting government to protect citizens personal data and deliver digital services [3], and to ensure that digital services keep up with ongoing technological change and development. Having an increased understanding of the issues facing public sector digital service users is of increasing interest in Australia, with ongoing service delivery changes from in-person/call centre (also referred to as analogue) to digital. With the growing complexities in the environment, the factors that impact why a service user may require assistance and their post assistance-seeking outcomes are becoming more influential in their perception of the digital services provided by government generally. Therefore, it is vital that research understanding digital adoption provides a more holistic view of the various issues facing service users, particularly focusing on understanding why they seek support. This paper seeks to address knowledge gaps that have been identified in the literature, including exploring what the barriers to digital adoption are in this specific case, why assistance was sought and how the assistance sought influenced their service use.

Through building and encouraging a user centric approach to researching the various interactions between individual public sector entities and users, the barriers to digital adoption will become more apparent. This is becoming increasingly more important due to regulatory changes in the public sector space as a result of the inclusion of 'digital first' policies, which have shifted mandatory services online. These policies have fundamentally changed the manner in which service users interact with public sector services. For this research mandatory environments are classified as "Public Sector Organisations who must by legislation provide Digital Platforms for their services" [1] [4]. While mandatory interactions are defined as "Users who meet certain characteristics and must by legislation interact with the public sector service provider to meet these obligations" [1] [4]. Therefore, users must engage with providers, but under the digital first mandate expectations around how they do so has changed.

This paper focuses on a public sector case study – the Australian Taxation Office (ATO) – that was impacted by the introduction of the Australian Digital Continuity Policy 2020, which mandated the use of digital first channels for all public sector services [5]. Through the examination of previous literature, ATO corporate research, data analysis and responses from ATO staff, a gap was identified between what seems to be common knowledge about the mandatory digital service user and the profiles of the actual users who are required to use them. The impact of shifting mandatory public sector services to a digital first platform is still unknown. As digital first service provision is the way forward for all public sector organisations (especially in Australia), a holistic view of users is needed. This paper aims to provide this view. Research needs to support and assist users, improve services and inform policy to increase long-term voluntary compliance obligations in a mandatory service space.

This paper will review previous literature focusing on digital transformation, digital adoption, digital ecosystems and eGovernment to understand the background of digital transformation in Australia and the basis for digital adoption and eGovernment worldwide. At present, the standard methods used for evaluating government services are based on interviewing or surveying users about their opinions and experiences of services provided. However, this style of research often results in biased results, as users feel pressured to display expected behaviours [6-8]. As a result of this contentious and possibly flawed data collection, previous research appears to have ignored a number of factors which impact service adoption, and seems to have failed to identify the barriers to adoption within mandatory environments. There is also a gap in understanding how different experiences with digital services (both in the public and private sectors) can impact long-term adoption and the reasons for when and why users seek assistance. The focus of this research is on applying systems thinking and digital ecosystems theories to understand and validate the need for a holistic view of the users and the system, especially when seeking to understand assistance-seeking behaviours.

In this research, systems thinking, particularly a soft systems methodology, has been used as a way to understand the behaviours and actions in complex public sector environments. An important principle is the concept that each action within the system causes a reaction in the system. These reactions can lead to unintended consequences, ones which are critical to explore [9]. This approach has been used to explore the planning process the ATO undertook to minimise the barriers experienced by taxpayers when submitting their returns. Keeping in mind that a core aim of the research is to understand the connections within the system, and the way in which each part of the system influences and is influenced by other parts, systems thinking is considered the most appropriate method. Systems thinking seeks connections between solutions, systems and society, identifying components of systems and intended and unintended outputs of the system, providing a holistic view of the problem.

This research explores the assistance-seeking behaviours of individuals when they are lodging their income tax returns with the ATO. The purpose is to understand the impacts of assistance-seeking on lodgement outcomes. Of the over 3 million individuals who lodged in July 2018, 5.3% sought assistance at some point throughout the process. The aim is to understand the different drivers of assistance-seeking behaviours within a mandatory system. Therefore, this research addresses two primary research questions:

1) What are the critical points in the lodgement process/system that are causing individuals to seek assistance?
2) What are the potential policy implications of understanding assistance seeking behaviours?

In applying a systems lens to these two questions, we ultimately seek to understand how and why the system as a whole functions as it does.

We have adopted a mixed methods approach to data collection and analysis with systems thinking to support the end-to-end research. That is to say, identifying who is most likely to contact the ATO for support and the critical points of assistance-seeking, and understanding post assistance-seeking behaviours and outcomes.

The overall approach applied to this paper incorporated multiple stages to provide the greatest depth of analysis and provide a holistic understanding of the assistance-seeking behaviours within the case study. First, systems thinking was used to assist with determining links between the different components, and understanding the different elements within the system and the effects each element could have on the outcome. Second, a thematic analysis was used to outline the common themes within the assistance-seeking behaviour, and understand the points in the system requiring the most assistance. Finally, a statistical analysis was considered appropriate to determine the behaviour of the actors within the system. Specifically, summary statistical methods was used to explore the population and understand who may be more inclined to require different kinds of assistance, in an attempt to provide policy recommendations for self-help prompts provided to individuals who meet certain characteristics.

This paper is divided into six sections. Section one contains the introduction, section two presents the literature review, section three presents the conceptual model, section four outlines the research methods undertaken, section five highlights the results and addresses the research questions, and section six offers policy recommendations.

II. Literature Review

Previous research demonstrates that the purpose of undergoing a digital transformation in the public sector is to increase access to provided services, including through digital services [10]. This is achieved through a better understanding of citizens and service users to improve their outcomes, making digital services easier to access, and improving the client experience [10]. An interesting challenge for the public sector, however, is to overcome the clashing expectations over private versus public services, which are personalised, modern and responsive. Therefore, the public sector must consider the end-to-end digital services in line with the private sector. Through the application of advanced analytics, governments are able to leverage the data collected from users to improve the services provided. The purpose of transitioning to digital services is to provide public sector services more effectively and efficiently to increase public value. For this research public value is ensuring that all mandatory public sector services provided are inclusive with both digital and non-digital options, ensuring equal access for all.

It is important to consider the variety of challenges facing public sector digital services. Firstly, the public sector takes advantage of technology that is popular within other industries; however, they do not have the time or finances to compete with the services provided by private sector entities. Secondly, governments are not always able to engage with citizens and service users to provide products and services in the manner expected. Finally, there are numerous regulatory restrictions which complicate the process. Therefore, digital technologies provide an opportunity to explore new channels for service provision, to improve resource management, increase access for users, and boost accountability and trust. Digital technologies deliver benefits across the economy and society [11], however, government services need to keep pace with the opportunities that digital transformations provide (including increased value for money for the community).

Digital transformation has empowered users and providers, making it possible to choose how services are accessed or delivered, how to communicate, when to engage on policy areas or issues, which social groups to join or business areas to invest in, and how to participate more actively in local, national and global challenges [12]. Research demonstrates that governments need to understand that going digital is no longer an option, but rather an imperative maintaining their legitimacy [12]. The adoption and use of digital technologies requires applying data more efficiently as part of their strategic components to modernise the public sector. Technologies are increasingly being used to digitise existing government processes and to offer public services online [12]. There has been a shift from a government-led to a user-driven administration, which is focused on end users and citizens expectations [12]. There are numerous challenges facing digital transformation, specifically around improving the digital experience. These include citizen security, cultural barriers to engaging with digital services, regulatory and legislative barriers (including those that restrict data sharing between government agencies), resource barriers and capability barriers (both public sector employees and users).

Research shown in Figure 1 outlines the progression towards digital transformations in the public sector [12]. The process started as analogue, which focused on in-person service delivery and paper-based processes. The second stage was eGovernment, which was the first stage of digitalisation, with the progressive inclusion of digital processes and procedures, including services provided to the users. The final stage in the progression is digital government, which is predominately based on digital first service provision, maximising user-driven approaches and citizen centric designs. This iteration highlights the value of inclusion, whereby exclusion from the digital world can exacerbate other forms of social exclusion such as unemployment, low education and poverty [11]. Every Australian should benefit in the shared digital future, which means that every member of the community provides insights into how they would like the service designed. This includes incorporating the user's views in the designs of processes and interfaces. The incorporation of users views, expectations and requirements into the design ensures that provided digital services take into account the different life stages and level of digital ability of users [11]. The application of systems thinking can assist in providing a different viewpoint of the potential barriers and their effect on digital adoption.

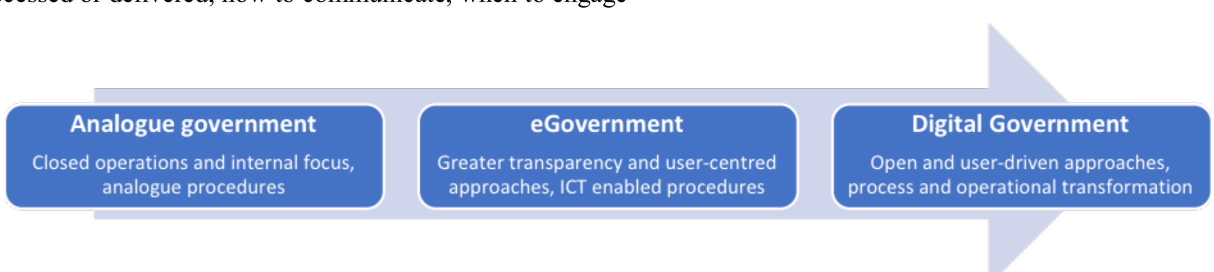


Figure 1. PROGRESSION TOWARDS THE DIGITAL TRANSFORMATION OF GOVERNMENTS [12]

A. Digitalisation

The evolution from traditional analogue services to electronic government services, to digital societies of interconnected and multichannel digital services has placed increased pressure on governments to provide high quality and easy to access public sector services [13]. This includes overcoming the challenges of meeting service user expectations, who commonly benchmark all services provided by public sector entities against those of the private sector. Users expect convenience and diverse communication channels with tailored information [13]. The development of digital services has impacted society considerably, specifically with the creation of eGovernment platforms, the shifting expectations to provide efficient, transparent and effective services, and to include open data (and transparency) and cross channel service options. With this shift to digital, the aim of public sector entities utilising these channels should be based on creating more efficient governments which offer service user's better services, enhancing the services already provided and aiming to increase all citizens' quality of life [14]. This is not a simple process, however; the use of digital processes and organisational change is the first step to ensure the successful implementation of digital first policies. It is shown that in countries that have successfully implemented digitalisation (e.g. Estonia), the government is able to provide services that meet the expectations of users, empower users, and increase their engagement with government [14].

The literature suggests that technology holds the key to shaping the world around us, as it enhances governments', businesses' and individuals' opportunities to integrate themselves within the global digital society, and enhance ongoing inclusion [15]. However, to become a high functioning digital government within the global digital society, digital adoption needs to be high, therefore, a key factor requiring consideration is the user's perspective. The most common determinants of eGovernment and digital government services adoption include the level of awareness and level of satisfaction of eGovernment services, both linked to intention to use [13]. Awareness relates to the extent users are aware of the eGovernment services [16] [17]. User satisfaction is being used to assess the continual use of eGovernment services and the success and failure of new eGovernment platforms [16] [18].

Digital or e-government adoption is most commonly defined as the continuous use of a digital service or innovation [19]. Therefore, for adoption to be achievable and sustainable, the digital service needs to be both appealing and useful [20]. This requires consideration of how a user may perceive the level of usefulness, ease of use, security level and reliability [21]. Furthermore, not all online activities are deemed to be of equal importance to an individual, especially when they are considering their personal level of human, social and financial capital [22]. Research suggests that there is a strong relationship between a person's level of education and the type and frequency of digital services utilised [22]. It is important to understand, when considering inclusivity of services, that technology access does not determine an inequality alone, whereby increased experience, exposure, digital participation and digital literacy are vital [22].

Previous research has identified four key barriers impacting digital access: (1) lack of basic digital experiences due to lack of interest, (2) no computer access, (3) lack of digital skills, and (4) lack of opportunities to use digital services [23]. Additional barriers impacting digital access include lack of access to internet, lack of awareness, language barriers, user-friendliness of websites, levels of trust and security fears [24]. Therefore, the digital divide is still an area that requires understanding and consideration when planning digital service provision in all contexts, including social, psychological, cultural and non-technological [23]. Thus, with the application of a systems view, a holistic understanding of these factors can be provided.

The challenges facing policy-makers going forward is understanding and determining the appropriate resources and functions necessary in digital services to provide a foundation for and to support positive user behaviour [25]. This requires ensuring the creation of digitally inclusive services. Digital inclusion ensures that all individuals and groups (including the disadvantaged), have access and the skills to use information and communication technology (ICT), and are thus able to participate in and benefit from the growing reliance on the digital knowledge and information society [26]. Thus, digital inclusion encourages increased access to information and communication technology, with the aim of increasing social and economic benefits. Digital inclusion links back to the digital divide, which is the gap between people who have effective access to digital technologies and those who have limited or no access. Access and cost become barriers impacting digital inclusion, however individual factors associated with engagement and confidence are affected by digital literacy, relevance, motivation, trust and safety. Therefore, the aim of eGovernment and digital government services is to provide information and public services to citizens in an easy to access manner that encourages platform participation [25]. This highlights the vital importance of understanding the users of services, the application of systems thinking to break down the different interaction points, potential barriers and so forth, so that services that meet the needs of the users can be provided. It is also important to consider the digital ecosystem impacting the users, which includes where the services fit within the broader environment, including public and private sector services. This links back to the importance of understanding different factors impacting the user's ability or willingness to utilise a digital service provided by the public sector.

B. Digital Ecosystem

When exploring the impacts of digitalisation on public service adoption, there are multiple components that need to be understood and these elements are commonly identified within the digital ecosystem. There are multiple definitions of digital ecosystems, however one of the most widely used defines a digital ecosystem as an "open, loosely coupled, domain clustered, demand driven, self-organising agents' environment, where each species is proactive and responsive for its own benefit or profit" [27, p. 3]. Therefore, each species or user who works within or utilises a digital ecosystem, is a participant who uses the system with a specific goal in mind [28 –30]. Therefore, a digital ecosystem can be characterised as consisting of organisational interactions, connected digitally, which are enabled by modularity, and are not managed by a hierarchical authority. Regardless of the definition, ecosystems are large, and they encompass numerous interactions between producers, suppliers, innovators, customers and regulators, shaping a collective outcome [31].

An ecosystem therefore emerges as a result of digitalisation, and through this process it becomes possible to connect a broad set of users together through the delivery of a singular digitalised customer solution. Thus, demonstrating the importance in understanding the creation of digital ecosystems within this research. As to ensure multiple external factors influencing user's ability are understood and incorporated into the design and planning of public sector digital services and policy.

There are two definitions applied by this research. The first defines a digital ecosystem as "an interdependent group of actors (enterprises, people, things) sharing standardised digital platforms to achieve a mutually beneficial purpose" [32, p.1]. The second definition defines a digital ecosystem as, "a network of digital communities consisting of interconnected, interrelated and interdependent digital species" [33, p. 249]. Both definitions include the stakeholders, institutions and digital devices situated within an environment, that interact as a functional unit and are linked together through actions, information and transaction flows. These definitions imply that all of the connections made by service users to achieve their specific goals within a digital platform are incorporated within the digital ecosystem. Therefore, digital ecosystems are shared communities, with scalable resources used to pursue challenges of specific goals and objectives [34]. Finally, the level of complexity within a digital ecosystem can be attributed to the differences between the participants taking part in the system and their objectives [35] [28].

The different components relevant to the inner workings of a digital ecosystem are outlined in Figure 2. At the base are the users, who are the people, businesses and entities. They consume the services from the available channels (through digital platforms, directly or through other channels) [36]. The first level is the government, which includes the federal, state and local authorities who influence policies and legislation, as well as departments, organisations and entities that implement policies and provide services to fulfil mandatory obligations [36]. The second level is the market, which refers to the non-government entities (e.g. academics, not-for-profits), intermediaries (such as health providers, tax agents), service providers (such as technology companies) and industry (such as banks, commercial entities) [36]. Finally, the third level is the environment, which includes the social norms, cultures, societal interactions and the access to the information and technology made available by the other levels [36].

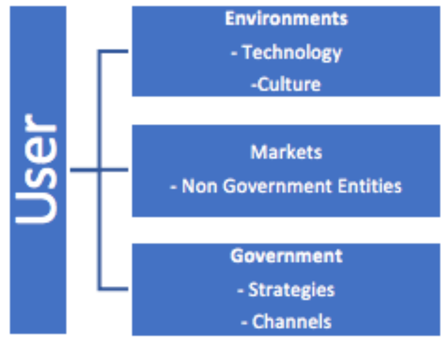


FIGURE 2. COMPONENTS OF A DIGITAL ECOSYSTEM [36]

Considering the different levels and the elements within the digital ecosystem under exploration, assists in ensuring a holistic understanding of the problem or situation under exploration. Through the exploration of the specific digital ecosystem relevant to a mandatory system, a testable conceptual model was created.

III. CONCEPTUAL MODEL

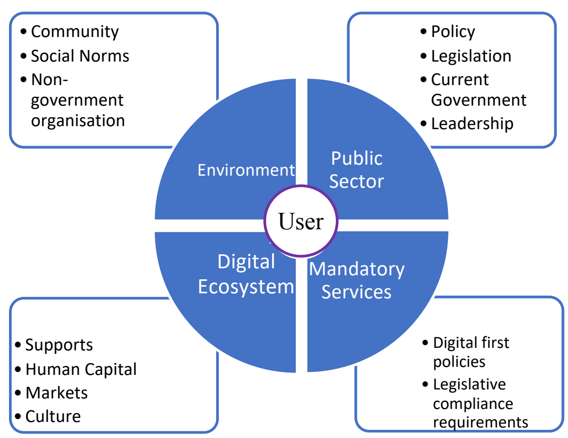


FIGURE 3. CONCEPTUAL MODEL WITH THE USER AS THE CENTRE FOCUS [3] [4]

A testable conceptual model was designed as a result of the analysis of digital, business, technology and innovation ecosystems. The purpose of the model is to provide user-centred research to guide policy and better support and understand the different users. Previous research does not focus of factors which impact a user's ability to adopt and participate within a mandatory digital ecosystem. The existing research provides minimal discussion on how digital adoption can differ in mandatory and voluntary environments. Through a thorough literature review, a number of factors within a user's environment were identified as having considerable impact on willingness or capacity to adopt mandatory digital services.

This research is focused on understanding the Digital Ecosystem quadrant in Figure 3, to explore the elements of supports and human capital within this quadrant, to determine what impact different elements components have on the Mandatory Services and Public Sector quadrants.

IV. Method

The motivation for this study originated from a desire to understand barriers to digital adoption in the public sector. A comprehensive literature review of research centred on environmental ecosystems [1] [4] has revealed that there is a lack of inclusiveness in the current mandatory system. The method outlined below outlines how the analysis provided the rationale for the proposed conceptual model: A Digital Ecosystem Quadrant (see *Figure 3*).

A. Case Study

The use of a case study method has been applied to this research to understand the different issues affecting digital adoption in the Australian Public Sector environment. The ATO case study was selected for numerous reasons. Firstly, the ATO is the lead in the public sector service provision (as the first Australian Public Sector (APS) organisation to adopt the digital first policy in Australia) [37]. The ATO collects considerable amounts of data, including interaction level data and mandatory interaction requirements data. The purpose of undertaking a case study is to explore and understand the uniqueness of a single case, while also understanding how findings from a specific case links to similar organisations or situations facing similar issues [38][39]. The use of a case study encourages greater in-depth exploration of the complex issues impacting digital adoption in that case. Therefore, by understanding the distinctiveness of the specific entity and users within the case under exploration, further research can also determine similarities to other cases (and their users and entities) [38] [39].

A mixed methods approach was applied to this research, which was appropriate for ATO case study. Therefore, the application of mixed methods encouraged the integration and interpretation of qualitative and quantitative data. This research carried out exploratory summary data analysis, text mining and thematic analysis, and hypothesis testing based on experimental design, to understand results from the data collection. This approach was applied to obtain an understanding of the key barriers impacting digital adoption and how they could be overcome. The starting point for this research analysis involved a qualitative method to explore and understand the different meanings and themes individuals or groups link to a problem [40]. An inductive approach was applied to this research, which involved a process of searching for patterns within observations, which was then used to develop explanations or theories, and from which a series of hypotheses were created [41] [42]. The application of a combination of mixed methods and inductive approaches, allowed for the emergence of research questions to assist in the identification of themes within the results. A pilot study was conducted in 2017, in order to validate the umbrella research questions and guide future research directions. This included informing feasibility and testing the research design. The purpose of the pilot study was to identify the potential problems and thus assist with designing and undertaking a larger and more informative study [1] [4]. The pilot data and analyses [1], informed a number of research questions and hypotheses as part of a larger study completed in June 2018. This data collection focused on quantitative data (both experimental and survey conditions). The application of a quantitative approach has been used for testing objective theories, through the examination of the multitude of relationships between the variables [39].

Data collection for this study occurred during a 4-week period at the beginning of 2018 Tax Time. This included a survey form which consisted of questions, both qualitative and quantitative, designed in a manner to explore different components and characteristics of a random assortment of callers (n = 3,990). The survey queried the reason for call, the caller's demographics, and why assistance was sought. To achieve randomisation, the survey collection was provided to 11 call centre operatives. All assistance-seeking phone call data was also collected. This information included what type of assistance was being sought, basic demographic information and post call outcomes. The total call centre population (n= 188,971) provided a large sample to complete further quantitative data analysis.

Post data collection, data was anonymised and categorised based on areas of interest and demographics. The first analysis utilised descriptive statistics (e.g. mean, median, proportions) to identify and understand the features within the sample population. Furthermore, through a thematic analysis of collected qualitative data, the different reasons assistance was sought were identified. The study findings address specific factors in relation to the purpose, timeframe, lodgement behaviour and number of individuals seeking assistance at various points of the lodgement process. The key questions used to collate the collected data are provided in Table 1.

TABLE 1. EXAMPLE OF THE QUESTIONS USED TO COLLATE THE DATA

Questions	Themes Factors
Why do individuals seek assistance from the ATO?	Assistance Seeking/ purpose/ intentions/ motive
Within the lodgement process where do individuals seek assistance?	Assistance seeking/ lodgement behaviour/ source of support

How many questions were asked per call? How many individuals contacted the ATO more than once?	Types of queries/ individual profile/ assistance seeking behaviour
What are the effects of assistance-seeking on lodgement timeliness?	Lodgement behaviour timeframe
How did they lodge post assistance-seeking?	Prediction adoption behaviour

B. Thematic Analysis

A thematic analysis or topic modelling was conducted on the qualitative data collected within the survey. The analysis was conducted utilising Python statistical software and the Natural Language Toolkit (NLTK) package for natural language processing. Through the use of this toolkit, the different themes or topics were identified and grouped together to create broader categories [43]. An extension of the standard approach to Gioia analysis was undertaken utilising statistical software to validate the process (see [4] for the initial Gioia analysis). When conducting the Gioia method for qualitative rigour [44] [45], the researcher categorised the accounts into three separate phases (first, second and third order), however these steps were conducted out of order. The first order, 'Concepts', is the 'voice of the user' (also known as 'voice of the customer'). The second order, 'concerns and statements', identifies specific sentences from participants which are then grouped together to discover the themes and patterns in events and accounts. These create themes that are more generalised underlying explanatory dimensions, to test consistency and patterns [44] [45]. Finally, the third order 'aggregate dimensions', identifies the generic themes encompassing all of the first and second order data [44] [45]. Significance was measured through counting occurrences of first, second and third order elements to identify themes and patterns throughout the different accounts. The patterns in the text were then linked by connections, highlighting key features and emergent concepts or themes that require further analysis.

Starting with the third order or 'aggregate dimensions', generic themes and topics were identified by word frequency through the application of Latent Semantic Analysis (LSA). LSA is based on the use of a distributional hypothesis, whereby words and expressions occurring within similar parts of text have similar meanings [46]. The significance of each of these themes was tested in the following stages of analysis, whereby counting occurrences identified themes and patterns throughout the different accounts. The second order was completed next, which seeks 'concerns and statements', which identify key sentences or phrases through the use of Latent Dirichlet Allocation (LDA) to understand themes and patterns within the accounts. LDA utilises mathematical probabilities to help define the unknown words that represent a known topic, by mapping the known elements to the unknown elements in a way that provides the probability of a word belonging to a particular topic [47]. These were used to create themes which are generalised underlying explanatory dimensions that demonstrate consistencies and patterns within the data [44] [45]. Finally, the first order 'concepts', or 'voice of the user' were identified, through the use of Text similarity Metrics. Jaccard Similarity calculates how similar two sentences are by determining the size of the intersection and the size of the union of two sets, identifying the number of words in common between sentences and providing a numeric output [48]. This identification process was simplified by the application of information obtained in the LSA and LDA processes. The patterns in the text were linked together manually to identify connections, and to highlight key features and the emergent concepts or themes not picked up by the analysis.

The results were validated by another independent researcher, who conducted their own analysis of the data provided and obtained equivalent outputs.

C. Systems Thinking

Systems Thinking analysis was applied to systematically identify and order findings into their respective components of the process [49]. This helped to identify the points within the process and system that are causing the most issues and where support can be implemented. Systems Thinking was used to visually convert the findings into simplified figures that highlight key emergent findings.

Process mapping and systems thinking principles were utilised to understand the relationships between the different steps of the lodgement process and where the different assistance-seeking behaviours sit within that process. The application of systems thinking, broken down into four key steps, was used to assist in providing a clearer understanding of the situation. Firstly, how people seek assistance to meet their mandatory ATO lodgement requirements was explored. This recognised and explored the role of digital systems, non-digital systems and assistance seeking in the process. Secondly, analysis of the assistance-seeking behaviours was conducted to understand the decisions and user in greater detail. Thirdly, a process map was drawn, which outlined the system and how assistance-seeking fits within it. Finally, the process map was used to inform the research question idea. This final part was an iterative process that was altered and added to as understanding about the users experience emerged. Process mapping is used to demonstrate, using a pictorial representation, the sequence of actions and responses between the start and end of a process [49]. This is commonly used to determine where there might be issues, inefficiencies or opportunities within the current process [49].

V. RESULTS

Results from this research demonstrate that there are multiple reasons why individuals seek assistance when undertaking their annual tax lodgement, including both digital and non-digital queries. As indicated in a previous pilot study [4] many taxpayers would have been unable or would have struggled to lodge their tax return without obtaining assistance.

A. Individuals Seeking Assistance

Firstly, there is no unique type of individual who seeks assistance – they differ in age, gender, income, occupation and even how many times they have lodged previously. As for age, as shown in Figure 4, 49% of individuals were aged between 18 and 29 years old, demonstrating that younger people are more likely to seek assistance to complete their tax return. The least likely age group to ask questions were those above 65+ years. In most of the comments from the phone contact, the individuals in both age groups, only sought assistance due to a change in their circumstance or because they were attempting an online lodgement for the first time.

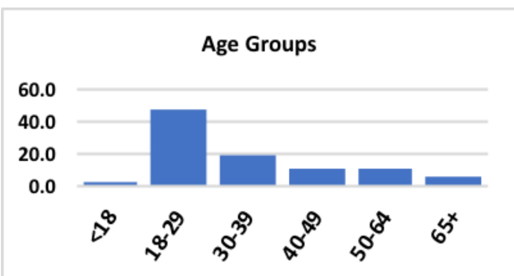


FIGURE 4. ASSISTANCE SEEKING BY AGE GROUP

Interestingly the overall the distribution of genders is relatively similar, with slightly more females (48.8%) than males (45.1%) seeking assistance overall, as shown in Table 2.

TABLE 2. AGE DISTRIBUTION

Female	Male	Undisclosed
48.8%	45.1%	6.1%

In the breakdown of the population by occupation codes in Table 3 [50], individuals who indicated being Community and Personal Service Workers (10.8%) and Labourers (13.1%) sought assistance most frequently. In comparison, Trainees, Apprentices and other related workers (3%) and Machinery Operators and Drivers (4%) sought assistance less frequently than other individuals.

TABLE 3. OCCUPATION CODES

Occupation	**%**
1 - Managers	4.7%
2 - Professionals	9.2%
3 - Technical and Trades Workers	5.2%
4 - Community and Personal Service Workers	10.8%
5 - Clerical and Administrative Workers	7.1%
6 - Sales Workers	7.8%
7 - Machinery Operators and Drivers	4%
8 - Labourers	13.1%
9 - Trainees, Apprentices and other related workers	3%
? – Not stated	35.1%

TABLE 4. INCOME RANGE

<$18,200	$18,201-$37,000	$37,001-$90,000	$90,001-$180,000	>$180,001	?
28%	21%	21%	4.4%	1%	34.1%

As per the ATO income tax brackets [51], the income range of individuals who sought assistance is demonstrated in Table 4. The majority of individuals who sought assistance earned less than $90,000 in the previous financial year.

TABLE 5. LODGEMENT

First Lodgement Year	Under 5 years	Greater than 5 years
7.7%	27.6%	64.7%

Table 5 demonstrates the number of years individuals who sought assistance had been lodging their tax return. 7.7% were undertaking their first lodgement, and 27.6% had been lodging their tax return for under five years. Interestingly, the majority of assistance was sought by individuals who had been lodging for greater than five years. The analysis was used to determine whether there was a relationship between the first year of lodgement, and how many people sought assistance.

B. Reasons For Seeking Assistance and Source of Assistance Seefiking by Individuals

Through the use of a Gioia and thematic analysis, the common themes as to why individuals sought assistance were identified. As shown in Figure 5, the results demonstrate that the majority of individuals (83.7%) sought assistance for utilising digital services (including platform support and technology support from the ATO). 10.3% of individuals sought tax advice (including system education, platform awareness and advice). Additionally, 2.9% contacted the call centre to obtain paper forms. In most cases individuals sought paper forms due to lack of computer skills, a preference for utilising paper and language barriers. Finally, 3.1% of queries were not related to lodgement of tax returns (including pre-lodgement and post-lodgement related queries). The thematic analysis demonstrates that at multiple points of the process individuals would not have been able to lodge their tax return without assistance.

Once the analysis was completed, a process map was created to outline the multiple interaction points within the system and where the different assistance-seeking points fit within it, shown in Figure 6. The diamond shaped points are decision points, rectangles are points of the process, and the oval shapes are outcomes. There are four key decision points where assistance is commonly sought. Contact point 1 is commonly where digital service support is sought. This is the first component of successfully interacting with the mandatory system. Without the appropriate access and support, some individuals who are not able to continue the digital process or who need to obtain a tax agent or request a paper form, do not lodge. Contact point 2 is where individuals contact the ATO for assistance in obtaining paper forms (publication ordering). Contact point 3 is referred to as the component when individuals require tax advice during their lodgement process.

There is a feedback loop associated with this contact point (which forms contact point 4) when an individual requires additional assistance either for the same or a different issue and contacts the ATO again for assistance. The other options for obtaining tax advice is to source the information through a google or ATO website search or obtain from a tax agent.

C. *Number of Queries Per Call*

As highlighted in Table 6, even though the majority of callers asked one question on average, in 48.4% callers had multiple questions. Interestingly 8.3% had four or more questions. For individuals who asked more than one question, in 92% of cases the questions theme changed. In 78% of these instances, the question was not one that had been prepared.

TABLE 6. NUMBER OF QUESTIONS PER PHONE CALL

1	2	3	4+
51.6%	26.8%	13.3%	8.3%

D. *Repetitive (Returning) Individuals*

Of the individuals who contacted the ATO seeking assistance, 38.7% did so on more than one occasion within the one-month period of data collection. Of those individuals, 69.9% contacted the ATO twice, 19% three times and 11.1% four or more times. Interestingly only 1.2% proceeded to lodge through paper forms. Of those remaining, 29.2% shifted to the use of a tax agent or intermediary, 32% lodged digitally and 37.6% had not yet lodged by October 31st 2018.

E. *Impact and Effect of Lodgement Timeliness and Lodgement Process Post Assistance-Seeking*

TABLE 7. POST ASSISTANCE SEEKING INTERACTION/LODGEMENT CHANNEL

	Total
Not lodged	23.7%
Digital lodgement	51.7%
Tax agent or intermediary	22.7%
Paper	1.9%

TABLE 8. POST ASSISTANCE SEEKING INTERACTION/LODGEMENT CHANNEL BY ASSISTANCE TYPE

	Digital	**Advice**	**Other**	**Publication ordering**
Not lodged	11%	24%	80.7%	32.5%
Digital lodgement	67.7%	31.1%	11.7%	9.4%
Tax agent or intermediary	20%	43.5%	6.6%	23.2%
Paper	1.3%	1.4%	1%	34.9%
Total	100%	100%	100%	100%

Individuals may lodges their tax return in Australia via a number of avenues. If an individual chooses to self-prepare they can lodge at no cost, through the Digital myTax channel, via a paper form and in some cases over the phone. If they choose, however, an individual can obtain an intermediary or tax agent to lodge on their behalf. Individuals who choose to complete their lodgement this way, pay for the service. The results of the thematic analysis demonstrate that assistance-seeking fit into four key categories: Digital Services, Tax Advice, Publication Ordering and Other non-lodgement queries. As shown in Table 7, of those who had sought assistance 76.3% had lodged within the expected timeframe, and 23.7% had not lodged on time. On-time lodgement indicates that an individual has lodged their income tax return prior to October 31st 2018, as per legislative requirements. The results of the thematic analysis demonstrate key elements outlining why individuals sought assistance.

The majority of those individuals who had not lodged on time were seeking support on other non-lodgement related queries (including obtaining a tax file number, superannuation queries and deceased estates), and in many cases those individuals were not required to lodge a tax return at all. As highlighted in Table 8, the majority of assistance seekers (51.7%) lodged through the digital platform, and the majority of these individuals sought assistance for digital issues followed by tax technical advice. Interestingly, only 11% of individuals seeking advice on digital matters had not lodged on time; however, 21.3% lodged through non digital means, which could imply that they were not able to obtain a solution. In contrast, 24% of individuals seeking tax technical advice had not lodged on time, and the largest portion of these individuals lodged through a tax agent or intermediary. This could demonstrate a lack of understanding of the system and the confidence they obtained by seeking additional support. For those requesting paper forms, 34.9% lodged through that method, whereas 32.5% of those individuals requesting paper forms had not lodged on time. This could be attributed to the additional processing time required for a paper form (approximately 50 days, as opposed to 14 days for digital and tax agent lodgements).

VI. Policy Implications and Recommendations

There are a number of potential policy implications and recommendations that have been determined within this research. However, it is important to add that this research highlights that assistance provided, no matter how well presented and in what format, may not provide the information in a manner that resonates with the individual seeking it. Therefore, it is important to provide a number of different channels for information, including different languages, the use of visualisations, and a combination of over the phone and in person assistance where needed. Furthermore, of all individuals who utilise the digital system, the number of individuals who sought assistance as part of digital system use is low, however we argue that this population is still important to understand and explore.

Our research found that the majority of assistance-seeking individuals were those who had been lodging for greater than five years. Therefore, the education provided to the individual regarding system use needs to be on a long-term basis, on an average of four to five years based on the evidence above. Training and policy design should also consider providing training to individuals when they have a change in circumstances or role in the system to ensure they have all information needed to successfully interact with the system. This falls into the concept of predictive adoption and assistance-seeking which will be explored in future research.

From the thematic analysis the majority of individuals sought assistance for digital advice, including system education and platform awareness. While the minority sought non-digital means for lodgement due to lack of computer skills and a personal preference towards utilising paper means of lodgement. This indicates that individuals require education systems that are accessible and relevant. Furthermore, incremental digital adoption requires system education at the point in time that the user requires it. The thematic analysis demonstrates that there are multiple points at which, without assistance, individuals would not have been able to lodge. Therefore, a critical aspect to consider for policy design is the identification and full systems integration of these key assistance trigger points. For policy-makers it is important to consider stronger investment of resources in support mechanisms at various points of lodgement (i.e. points identified with the system approach and thematic analysis. The process map shows where the different assistance seeking points fit within the system. Through this visual mapping process, it is possible to identify the correlations between where assistance was sought and the different stages of the lodgement process. These correlations highlighted the key points where self-help options or additional education could be provided, and this could assist in minimising the number of individuals who are required to contact the ATO for digital service issues.

Profiling individuals based on query behaviour will not be effective, as there is strong evidence to suggest that variations of queries and unpredicted types of questions emerge through the dialogue. This demonstrates the value of human interface (via phone, in person and chat windows) as these support services provide ongoing support and education. This research provides support for the notion that not everyone will adopt a fully digital system from public sector service providers. As highlighted, many individuals who sought assistance were able to determine multiple areas within the process that they required greater understanding, which was possible as a result of a dialogue with a human interface (or person).

This was demonstrated through the analysis of individuals who called numerous times, to obtain additional support from a human interface, even though additional information was available from other digital means. As a result of these human interactions, 32% of individuals who sought assistance more than once, shifted to digital lodgement from non-digital means the year before. Therefore, repetitive advice can be designed to target individuals needing reassurance of processes. This component of the research is still underway and will form part of future research. Finally, when understanding the individual's post assistance-seeking outcomes, the summary statistics demonstrate that additional research and profiling is required to understand in depth the different needs of the population. Future research will explore predictive adoption and assistance-seeking in post assistance-seeking outcomes.

The framework being developed within this research, which will result from more testing of the conceptual model, can be applied more broadly within the public sector digital services space. This research has explored in detail the Digital Ecosystem quadrant of the conceptual model, which outlines the importance of supports and human capital on digital service use. By providing public sector service designers with more information about the service users, the greater the level of understanding they will have which should lead to more inclusive services. This research framework has been applied more broadly to the adoption of digital health platforms, specifically exploring the different user views and support requirements to utilise and engage with these style services. This includes exploring how different types of assistance provided by an organisation, supports or hinders the use of a service, including available information, call operatives providing advice and other support services available outside of government. Understanding the broader implications of this research on other public sector services is underway and ongoing. Further research is being conducted to understand how the model fits within the social services space.

This research and the policy recommendations and implications are aimed at providing advice that encourages mandatory public sector services to be inclusive, utilising both digital and non-digital options, including the provision of assistance.

The policy implications from this study are fundamental and rich with possibilities for future longitudinal research. The most basic finding is that perceived human support is the main predictor of the intent to adopt a digital mandatory system, with relevance as the major constituent driver of perceived understanding of the digital government system.

VII. Limitations

This study has potential limitations. The first limitation is that this research is being conducted as part of a PhD. Therefore, the research is ongoing and progressive. The second potential limitation is the use of qualitative data and analysis may limit the generalisability of the specific findings within the research to other areas. The third potential limitation is the data collection was conducted within the period tax lodgements were due for the Australian Taxation Office, therefore some of the results are specific to lodgement of income tax returns. Lastly, the results of this research may not be completely generalisable. Additional research is underway to explore different public sector services and build a framework that is relevant across multiple channels and public sector services.

Acknowledgements

The results and views presented in this paper do not represent the view of the Australian Taxation Office or the Australian Government.

References

[1] S Papavasiliou, C Reaiche, and P Ricci, 'Digital Interactions Strategy: A Public Sector Case', The Thirteenth International Conference on Digital Society and eGovernments. pp. 19-23. 2019. ISBN: 978-1-61208-685-9

[2] L. Reichelt, Gov.AU is a 'Mental Model' of Government [Online]. Available from: https://www.dta.gov.au/blog/gov-au-is-a-mental-model-for-government/ 2016. Retrieved 10 November 2018.

[3] Digital Transformation Agency, Digital Delivery of Government Services: Digital Transformation Agency Submission to the Finance and Public Administration Committee [Online]. Available from: https://www.aph.gov.au/DocumentStore.ashx?id=9d695f6a-2354-4cc5-be0d-913de41b25de&subId=516630 2017. Retrieved 10 November 2018.

[4] S. Papavasiliou, C Reaiche and P Ricci, 'Digital Adoption: The Need for Truly Inclusive e-Government Services', Proceeding of the 2019 International Conference on E-Learning, E-Business, Enterprise Information Systems & E-Government. pp. 43-49, 2019. ISBN: 1-60132-495-2

[5] National Archives of Australia 2015, Digital Continuity 2020 [Online]. Available from: www.naa.gov.au/Information.management /Digital.Continuity.2020. Retrieved 10 November 2018.

[6] A. Furnham, 'Response Bias, Social Desirability and Dissimulation', Personality and Individual Differences. Vol. 7 (3), pp. 385–400, 1986. doi:10.1016/0191-8869(86)90014-0

[7] A. J. Nederhof, 'Methods of Coping with Social Desirability Bias: A Review', European Journal of Social Psychology. Vol. 15 (3), pp. 263–280, 1985. doi:10.1002/ejsp.2420150303

[8] M. T. Orne, 'On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications', American Psychologist. Vol. 17 (11), pp. 776–783, 1962. doi:10.1037/h0043424

[9] M Brimble and A Jones, 'Using Systems Thinking in Patient Safety: A Case Study on Medicines Management', *Nursing Management (Harrow)*. Vol. 24(4) pp. 28-33, 2017. doi: 10.7748/nm.2017.e1621.

[10] A Bertrand, How Does Digital Government Become Better Government [Online]. Available from: https://www.ey.com/en_gl/government-public-sector/how-does-digital-government-become-better-government 2019. Retrieved 10 July 2019.

[11] Department of Industry, Australia's Tech Future: Delivering a Strong, Safe and Inclusive Digital Economy [Online], Available from: https://www.industry.gov.au/sites/default/files/2018-12/australias-tech-future.pdf 2018. Retrieved: 10 July 2019.

[12] OECD, Strengthening Digital Government [Online]. Available from: https://www.oecd.org/going-digital/strengthening-digital-government.pdf 2019. Retrieved: 10 July 2019.

[13] Y Yang, 'Towards a New Digital Era: Observing E-Government Services Adoption in a Chinese Municipality', Future Internet, Vol. 9(53) pp. 21-17, 2017. doi: 10.3390/fi9030053

[14] B Corydon, V Ganesan, and M Lundqvist, Digital by Default: A Guide to Transforming Government [Online]. Available from: https://www.mckinsey.com/~/media/mckinsey/industries/public%20sector/our%20insights/transforming%20government%20through%20digitization/digital-by-default-a-guide-to-transforming-government.ashx 2016. Retrieved: 10 November 2018.

[15] World Economic Forum, Our Shared Digital Future; Building an Inclusive, Trustworthy and Sustainable Digital Society [Online]. Available from: http://www3.weforum.org/docs/WEF_Our_Shared_Digital_Future_Report_2018.pdf 2018. Retrieved: 10 July 2019.

[16] S Al-Jaghoub, H Al-Yaseen and M Al-Hourani, 'Evaluation of Awareness and Acceptability of Using e-Government Services in Developing Countries: the Case of Jordan', The Electronic Journal Information Systems Evaluation, vol. 13(10) pp. 1-8, 2010. ISSN 1566-6379

[17] A Charbaji and T Mikdashi, 'A Path Analytic Study of Attitude Toward e-Government in Lebanon', Corporate Governance, Vol 1(3), pp. 76-82, 2003. DOI 10.1108/14720700310459872

[18] A Alawneh, H Al-Refai and K Baitha, 'Measuring User Satisfaction from eGovernment Services: Lessons from Jordan', Government Information Quarterly, vol. 30, pp. 277-288, 2013. doi: 10.1016/j.giq.2013.03.001

[19] E.M. Rogers, Diffusion of Innovation, New York: Free Press, 1995.

[20] E. Ziemba, 'The Contribution of ICT Adoption to Sustainability: Households' Perspective', Information Technology and People, vol. 32(3) pp. 731-753, 2018. doi: 10.110 8/ITP-02-2018- 0090

[21] M. A. Shareef, Y. K. Dwivedi, S. Laumer and N. Archer, 'Citizens' Adoption Behaviour of Mobile Government (mGov): A Cross-Cultural Study', Information Systems Management, vol.33(3), pp. 268-283, 2016. doi: 10.1080/10580530.2016.1188573

[22] E. Hargitti and A Hinnant, 'Digital Inequality: Differences in Young Adults' Use of the Internet', Communication Research, vol.35(5), pp. 602-621, 2008. doi: 10.1177/0093650208321782

[23] J. Van Dijk, The Network Society, Social aspects of New Media, Sage: Thousand Oaks, 1999.

[24] E. Ziemba, T. Papaj and R. Zelazny, 'A Model of Success Factors for E-Government Adoption – The Case of Poland', Issues in Information Systems, vol.14(2), pp. 87-100, 2013. doi:10.18517/ijaseit.7.4.2518

[25] S. Sawalha, M. Al-Jamal and E. Abu-Shanab, 'The Influence of Utilising Facebook on e-government Adoption', Electronic Government, vol.15(1), pp. 1-20, 2019. doi: 10.1504/EG.2019.096573

[26] Institute of Museum and Library Services, University of Washington Technology and Social Change Group, International City/County Management Association, Building Digitally Inclusive Communities: A Guide to the Proposed Framework [Online]. Available from: https://www.imls.gov/assets/1/AssetManager/DIC-FrameworkGuide.pdf 2011. Retrieved 10 July 2019.

[27] E. Chang and M. West, 'Digital Ecosystem - A Next Generation of the Collaborative Environment', Proceedings of iiWAS 2006, Yogyakarta, 2006.

[28] M Hadzic, T Dillon, and E Chang, 'Use of digital ecosystem and ontology technology for standardization of medical records', Digital Ecosystems and Technologies Conference, 2007. DEST '07. Inaugural IEEE-EIS, 21-23 Feb 2007, pp. 595-601. Cairns, Australia: IEEE.

[29] LD Serbanati, F Ricci, G Mercurio, and A Vasilateanu. 'Steps Towards a Digital Health Ecosystem', Journal of Biomedical Informatics. Vol. 44(4) 201. doi: 10.1016/j.jbi.2011.02.011.

[30] GE Itawa, M Herselman and A Botha, 'Digital Health Innovation Ecosystems: From Systematic Literature review to Conceptual Framework', Procedia Computer Science, vol. 100 2016. doi: 10.1016/j.procs.2016.09.149

[31] World Economic Forum, Platforms and Ecosystems: Enabling the Digital Economy [Online]. Available from: http://www3.weforum.org/docs/WEF_Digital_Platforms_and_Ecosystems_2019.pdf 2019. Retrieved: 10 July 2019.

[32] Gartner, Insights from the 2017 CIO Agenda Report: Seize the Digital Ecosystem Opportunity [Online]. Available from: https://www.gartner.com/imagesrv/cio/pdf/Gartner_CIO_Agenda_2017.pdf 2017. Retrieved 10 November 2018.

[33] J Hausberg, K Liere-Netheler, S Packmohr, S Pakura, and K Vogelsang. Digital Transformation in Business Research: A systematic literature review and analysis. (April 26, 2018). DRUID18, Copenhagen Business School, Copenhagen, Denmark. Available from: https://ssrn.com/abstract=3169203 2018. Retrieved 10 November 2018.

[34] A Rudolph, Digital Transformation is About Ecosystems and Thinking Differently. Available from: https://www.digitalistmag.com/digital-economy/2017/07/12/digital-transformation-about-ecosystems-thinking-differently-05209209 2017. Retrieved 07 August 2017.

[35] G Briscoe, 'Digital Ecosystems,' Dissertation Department of Electrical and Electronic Engineering, 2009, URL: https://arxiv.org/pdf/0909.3423.pdf

[36] Australian Taxation Office, ATO Leads Digital by Default [Online]. Available from https://www.ato.gov.au/media-centre/media-releases/ato-leads-digital-by-default/ 2015. Retrieved 05 August 2017.

[37] Australian Digital Transformation Agency, Digital Services Platform [Online]. Available from: https://www.dta.gov.au/book/export/html/769 2017. Retrieved 05 August 2017.

[38] H Simons, Case Study Research in Practice, Sage Publication: LA 2009.

[39] H Harrison, M Birks, R Franklin R and J Mills, 'Case Study Research: Foundations and Methodological Orientations [34 paragraphs]', Forum Qualitative Sozialforschung/ forum: Qualitative Social Research, vol. 18(1), pp. 1-17. 2017. doi: 10.17169/fqs-18.1.2655

[40] JW Creswell JW, Research Design: Qualitative, Quantitative and Mixed Methods Approach (3rd Edition), Sage: LA, 2009.

[41] W Goddard and S Melville, Research Methodology: An Introduction 2nd Edition, Juta & Co. Ltd: Lansdowe, 2004.

[42] HR Bernard, Research Methods in Anthropology 5th Edition, AltaMira Press: United Kingdom 2011.

[43] S Bird, E Klein and E Loper, Natural Language Processing with Python: Analysing Text with the Natural Language Toolkit. O'Reilly Media: United Kingdom 2009.

[44] D. A. Gioia, K. G. Corley and A. L. Hamilton, 'Seeking Qualiative Rigour in Inductive Research; Notes on the Gioia Methodology', Organisational Research Methods, vol. 16(1), pp. 15-31, 2012. doi:10.1177/1094428112452151

[45] D. A. Gioia and .K Chittipeddi, 'Sensemaking and Sensegiving in Strategic Change Initation', Strategic Management Journal, vol. 12(6), pp. 433-448, 1991. Doi: 10.1002/smj.4250120604

[46] TK Landauer, PW Foltz and D Laham, 'An Introduction to Latent Semantic Analysis', Discourse Processes vol. 25, pp. 259-284, 1998. doi: 10.1080/01638539809545028

[47] DM Blei, AY Ng and MI Jordan, 'Latent Dirichlet Allocation', Journal of Machine Learning Research vol. 3, pp. 993-1022. 2003. ISSN: 1532-4435.

[48] A Ritter, M Etzioni and O Etzioni, 'A Latent Dirichlet Allocation Method for Sectional Preferences', Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11-16 July 2010.

[49] Damaleo R, The Basics of Process Mapping (2nd Edition), Boca Raton Florida: CRC Press 2011.

[50] Australian Bureau of Statistics, Australian and New Zealand Standard Classification of Occupations, Available from: https://www.abs.gov.au/ausstats/abs@.nsf/mf/1220.0 2013. Retrieved: 10 July 2018.

[51] Australian Taxation Office, Individual Income Tax Rates. Available from: https://www.ato.gov.au/Rates/Individual-income-tax-rates/ 2019. Retrieved: 07 August 2019.

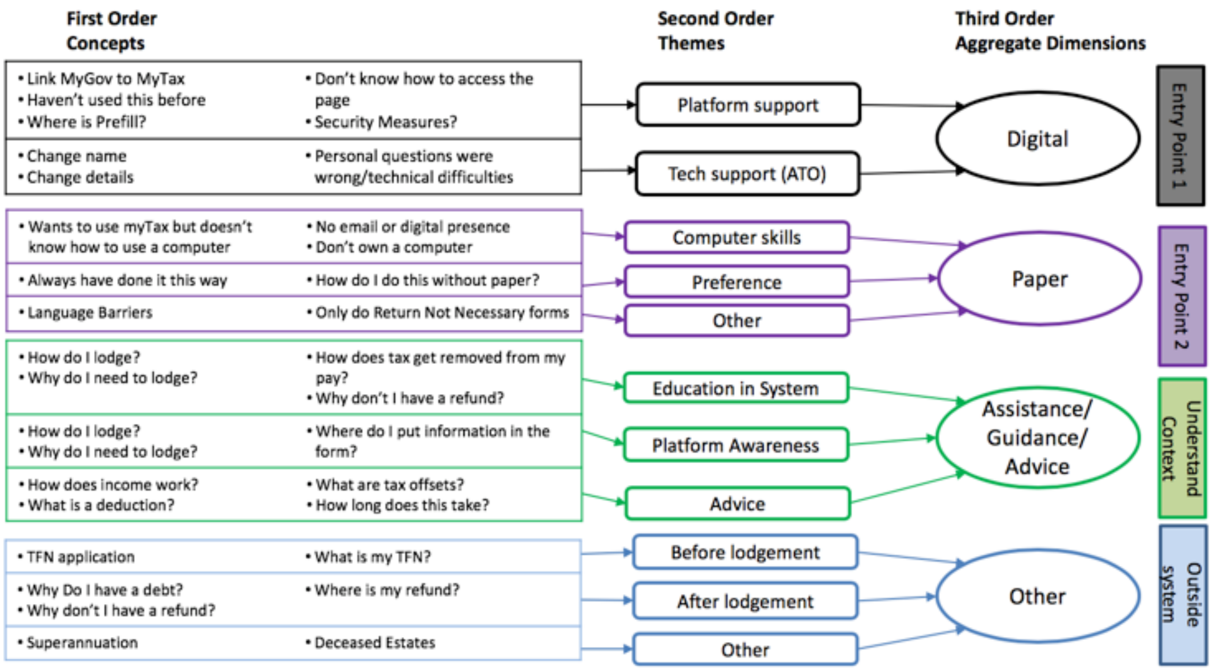


FIGURE 5. RESULTS OF THE GIOIA AND THEMATIC ANALYSIS

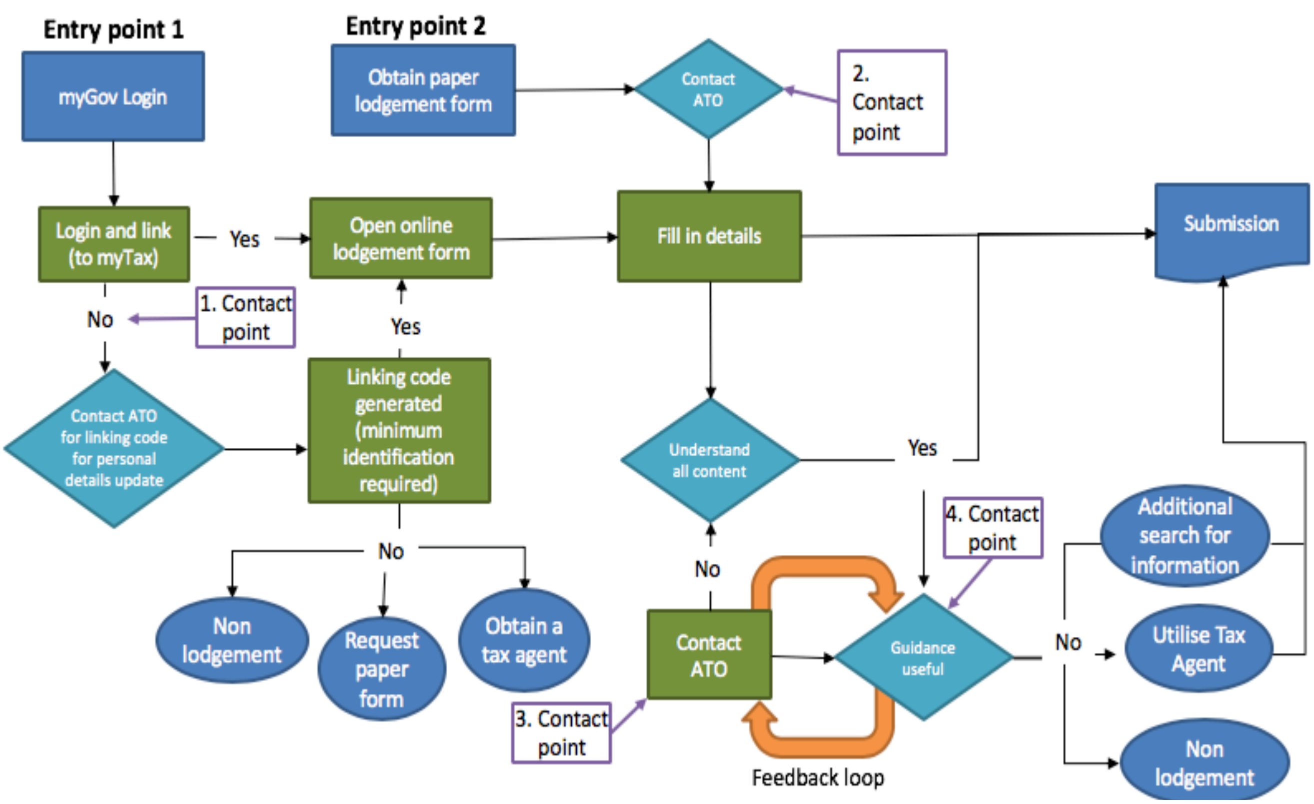


FIGURE 6. PROCESS MAP OF LODGEMENT PROCESS

Smart Shopping Cart Learning Agents

Dilyana Budakova
Department Of Computer Systems and Technologies
Technical University of Sofia, Branch Plovdiv, Plovdiv, Bulgaria
email: dilyana_budakova@yahoo.com

Lyudmil Dakovski
Department of Communication & Computer Technologies
European Polytechnic University
Pernik, Bulgaria
email: lyudmil.dakovski@epu.bg

Veselka Petrova-Dimitrova
Department of Computer Systems and Technologies
Technical University of Sofia, Branch Plovdiv, Plovdiv, Bulgaria
email: vesi_s_petrova@yahoo.com

***Abstract*—The paper describes the design, implementation and user evaluation of utility-based and goal-based intelligent learning agents for smart shopping cart. In keeping user's shopping habits or user's shopping list, they guide visitors through the shops and the goods in the shopping center or according to new promotions in the shops, respectively. It is envisaged that concrete implementation of the shopping agents will be running on each shopping cart in the shopping centers or on holographic displays. The k-d decision tree, the best identification tree, and reinforcement-learning algorithm are used for agents learning. The task environment is partially observable, cooperative, deterministic, and a multi - agent environment, with some stochastic and uncertainty elements. It incorporates text-to-speech and speech recognizing technology, Bluetooth low energy technology, holographic technology, picture exchange communication system. Machine learning techniques are used for agents modeling. This kind of intelligent system enables people with different communication capabilities to navigate in large buildings and in particular to shop in the large shopping centers and maximize user comfort. Some initial user opinions of the shopping cart agents are presented. Different embodiments of the shopping agents are discussed like holographic agent embodiment, embodied virtual agent or social robot embodiment. Some approaches of realization of a smart robotic shopping cart that can follow the user are discussed too. The performance of the Q learning algorithm with an introduced environment measures model (a model of the environment criteria) is proposed and explored. Study of the learning parameter is presented. Smart Shopping Cart Learning Agents modeling and development task allows for applying and improving the learning algorithms.**

Keywords-smart shopping cart learning agent; machine learning; reinforcement learning; Q learning; decision tree; identification tree; ambient intelligence; holographic technology; beacon-based technology; assistive technologies.

I. INTRODUCTION

In big and unfamiliar indoor spaces, such as shopping centers, airports, stadiums, hotels, office buildings, people may have difficulties with finding the desired destination. Many categories of people – the elderly, the children, the people with visual or hearing impairment, with difficulties in communication etc. – need specialized ways of communication [1][2][3][4]. This paper presents modeling, implementation and user evaluation of three intelligent learning agents for smart shopping cart. They guide visitors through the shops and the goods in the shopping center according to users' shopping habits, user's shopping list or according to new promotions in the shops, respectively [1]. It is envisaged that concrete implementation of the shopping agents will be running on each shopping cart in the shopping centers or on holographic displays [1]. The paper is inspired by [1] presented at the Cognitive 2019 conference.

The task environment incorporates text-to-speech and speech recognizing technology, Bluetooth low energy technology, holographic technology, information kiosks, picture exchange communication system.

The rest of the paper is structured as it follows: in Section II the technologies that the task environment incorporates are briefly discussed; in Section III the task environment specification, including performance measure, properties, environment actuators and sensors description is presented; the agent programs realization of the goal-based learning agent, personal utility-based learning agent and utility-based learning agent by means of a decision k-d tree, identification tree and reinforcement-learning are explained in Section IV; the degree of development of the proposed cognitive architecture components is explained in Section V; an empirical survey about the interest of end customers to the used technologies; a survey about the way the customers perceive the three developed agents; a survey about users' opinion and possibilities for shopping agents embodiments and user following smart shopping cart realizations; a survey of the performance of the Q learning algorithm with an introduced environment measures model (a model of the environment criteria) and study of the learning parameter ϒ are considered in Section VI; A section for future work is Section VII; in the VIIIth Section a number of conclusions are drawn.

II. BACKGROUND TECHNOLOGY USED FOR TASK ENVIRONMENT

Beacons are used to mark the location of objects and navigate people in indoor spaces [5][6][7][8][9]. They work on the principle of lighthouses by emitting signals at short intervals based on Bluetooth Low Energy (BLE) technology. The distance to the Beacon can be defined depending on the signal strength [10]. In addition to emitting advertising or other types of announcements, it is also possible to locate beacons [5][7][9].

Holograms are made of light and sound, appear in the around space and reply to gestures, voice and gaze commands [11]. A hologram can be placed and integrated in the real world or can tag along with user as an active part of user's world helping for navigation in indoor spaces.

Another possible solution to the problem of orientating people in indoor spaces is the use of embodied conversational information kiosks [12][13]. These systems use the information they have both about their own location and about the layout of the building and give instructions to the users how to find the desired place in the building.

The information kiosks are a collection of different technologies such as video processing from face detection, speaker-independent speech recognition, array microphone for noise cancellation, a database system, and a dynamic question answering system [13][14].

The Picture Exchange Communication System (PECS) [15][16] allows people with little or no communication abilities to communicate using pictures. People using PECS are taught to approach another person and give them a picture of a desired item in exchange for that item [17][18].

Screen readers [19][20][21] and text-to-speech (TTS) systems [22] enable blind and vision impaired people to use computers and provide the key to education and employment.

According to [23], the first step in designing an agent must always be to specify the task environment as fully as possible. That includes performance measure, environment actuators and sensors description. That's why we will consider smart shopping problems, task environment specifying and shopping learning agents modeling in the next section.

III. Specifying the Task Environment

It is envisaged that shopping agents will be implemented on the shopping cart. The consumers will run their cart following the directions given by the agents. In the future, the shopping agents can be implemented on a robotic shopping cart like an autonomous Kuka robot that can be controlled by gestures [24][25][26][27][28] to follow the user. Then, the environment will become very complex and similar to the environment of the automated driver.

The modulus of the system prototype is given in Figure 6. The task environment consists of four main blocks: input, output, shopping, and navigation.

The technologies, used in the input block, are face detection and speech recognition. The equipment comprises a camera, a microphone, a keyboard, a mouse, and a touch screen. The general object detection algorithm consisting of a cascade of classifiers proposed by Viola and Jones [29] is used to detect faces. For video processing, C# and Intel OpenCV library [30] is used.

The output block uses speech synthesis and virtual character visualization for giving information to the user.

The shopping block includes: drag and drop pictures for creating the shopping list (Figure 4); pictures-to-speech convertor;

The navigation block includes: Beacons/iBeacons or/and Holograms for smart buildings, smart shopping mall navigation. Using of Google Beacon Platform or/and Microsoft HoloLens respectively is needed.

Agent programs include goal-based learning agent, utility-based learning agent and personal utility-based learning agent realization by means of reinforcement-learning, decision k-d tree and identification tree building.

A. Performance Measure

The performance measure, to which the shopping agents are aspired include getting to the correct shop in the shopping mall; getting to the new promotion in the shopping mall; minimizing the path when going through the shops from the shopping list; maximizing passenger comfort; maximizing purchases; and enabling people with different communication possibilities to navigate in big buildings and in particular to shop in the big shopping centers.

B. Environment

Any shopping agent deals with a variety of shops in the shopping malls; the newest promotion could be in each and any of the shops in a mall; the agents can recommend visiting the shops in a mall in various sequences. An option is to visit all desired shops following the shortest possible way. Another option is to go around the shops in accordance with the arrangement of the items on the shopping list or in accordance with shopping habits of the user. A third option is to go to the shops in accordance with the availability of sales or new promotions. The location of the shops in an exemplary Mall is given in Figure 1. The model of the environment in Figure 3 is presented in the form of a graph. The nodes are the shops and the edges are the connecting corridors.

Figure 1. Exemplar location of eight shops in a shopping center.

C. Actuators

The shopping agents are visualized on the display screen. Only the head of an agent is modeled by means of the program Crazy Talk. Face animation includes synchronization of the lip movement with the pronounced text and expressed emotions. The agents' faces normally express friendliness and calmness and when a new promotion or sale is announced they express excitement and joy. The emotions of elevation are realized through changing the strength and the height of the speech and by visualizing a model of the emotion "joy" on the face.

The shopping agents can be realized as holographic agents. In this case they will be visualized on holographic displays. Then not only the heads of the agents will be modeled, but also their bodies, their gestures, clothes, as well as the way of keeping the appropriate distance to the consumer;

D. Interaction and Sensors

For interaction both with the intelligent agents and the consumers are used: Keyboard entry; Microphone; Touch Screen; Camera – OpenCV, Face Detecting; Natural Language Understanding; Speech recognizing; drag-and-drop pictures, pictures to speech convertor; Beacons/iBeacons or/and Holograms for smart shopping mall navigation.

E. Properties of a Task Environment

The behavior of the three intelligent agents is mutually complementary. They aim at facilitating the user access to the desired commodities and increasing the number of purchases, made by him/her, as well as at offering information about promotions and sales, in which he/she is interested.

The agent does not know when a new promotion or a new customer will appear. Therefore he/she periodically checks on the site of the mall if there are files, containing information about new promotions or sales and reads them if available. Then, he/she transmits this information to the customers, planning to visit the corresponding shops. Whenever a new customer appears, the agent receives his/her shopping list and defines the sequence for visiting the shops in the mall. That’s why task environment is partially observable, cooperative and a multi-agent environment.

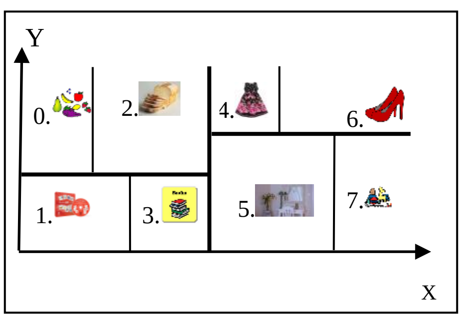


Figure 2. Shop k-d decision tree.

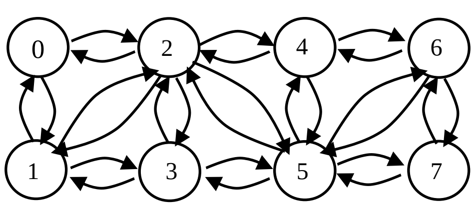


Figure 3. Presentation of the location of the shops in an exemplary mall by an orientated graph.

The shopping world is also deterministic with some stochastic elements and contains elements of uncertainty of the environment. The task environment is episodic and it can be realized either as static or as semi-dynamic environment. The environment can be regarded as static as the location of the shops in the mall is known. The agent receives the whole shopping list and suggests a certain path around the shops. Whenever there is information about a new promotion or sale appearing during the shopping, the agent can dynamically recommend a change in that shopping sequence.

The environment can be regarded as both known and sequential because every next shop to visit is determined by the current location of the user and by the items he/she has pointed at as important to buy.

IV. Smart Shopping Learning Agents Modeling

Three software agents have been realized. The first one is a utility-based learning agent, the second is a goal-based learning agent, while the third is personal utility-based learning agent.

A. Utility-Based Learning Agent

One of the agents can be regarded as a Utility-based agent. That is because it feels happy when discovering that there is a promotion or a sale in a shop, in which the customer is interested to go.

The utility-based agent uses a decision k-d tree to quickly find where, (in which shop) the customer is located according to his/her coordinates. The theory of building and implementing a decision k-d tree is given in [31]. The customer is depicted in Figure 1 by means of an emoticon, which can be moved using the mouse and placed everywhere on the shown map of the shops in the shopping center. Another way of finding the location of the customer is by using estimate beacons sensors or holograms.

The Utility-based agent checks if there are new files about promotions or sales published on the site of the shopping center. In case there are such files, it withdraws them and informs the customer about those of them, which are related to the shops the customer intends to visit.

The information about promotions and sales is given to the customer also in the case when it can be seen from the shopping list that the customer has planned to visit a particular shop where there is a promotion or a sale.

The customer receives notifications about promotions/sales when he/she goes past a beacon as well.

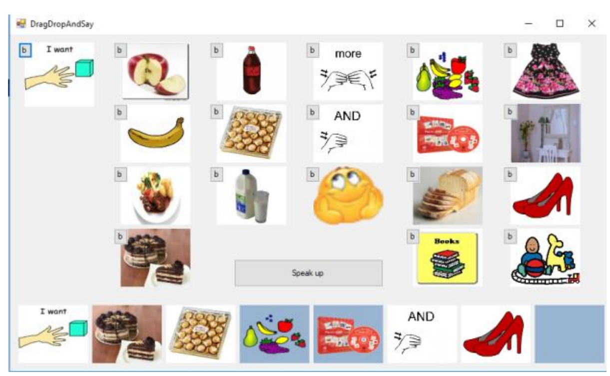


Figure 4. Making a shopping list by dragging and dropping pictures.

1) Decision k-d Tree Realization

In order to build the decision tree, the location of the eight shops in the exemplary shopping mall, given in Figure 1, is considered. As it is described in [31][32] all shops are divided first by width alone into two sets, each with an equal number of shops. Next each of the two sets is divided by heights alone. Finally, each of those four sets is divided by width alone, producing eight sets of just one block each. The shop sets are divided horizontally and vertically until only one block remains in each set as it is shown in Figure 2. The overall result is called a k-d tree, where the term k-d is used to emphasize that the distances are measured in k-dimensions.

Finding the nearest block is really just a matter of following a path through a decision tree that reflects the way the objects are divided up into sets. As the decision tree in Figure 2 shows, only three one–axis comparisons are required to guess the shop, in which the user is positioned.

In general [31], the decision tree with branching factor k=2 and depth d=3 will give 2^3= 8 leaves (shops in our task). Accordingly, if there are n shops (or goods, or users) to be identified, d will have to be large enough to ensure that $2^d \geq n$. Then, the number of comparisons required, which corresponds to the depth of the tree, will be of the order of $\log_2 n$.

B. Goal-Based Learning Agent

According to [33][34][35], Reinforcement learning is a method of learning, by which what to do is taught, i.e., how to match a situation to an action, so that a numerical reward received as a signal, is maximized. The teacher does not point at the actions to be undertaken. Instead, the trainee has to find out those, leading to the greatest reward and try to realize them. In the most interesting and challenging cases, not only the immediate reward could be taken into account when choosing an action, but also the further situations and the future rewards.

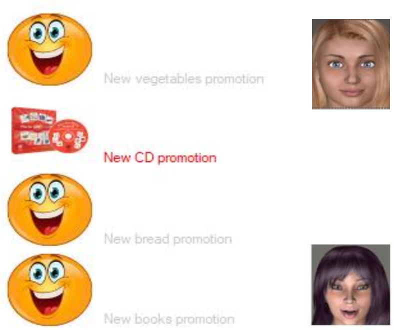


Figure 5. Message about a promotion of a new Compact Disk.

All reinforcement learning agents have explicit goals, can sense aspects in their environment and choose actions, which influence it. The agent is realized by a program, matching the way the agent perceives reality and the actions it undertakes.

A reinforcement learning algorithms is used for the second Goal-based learning agent. The agent receives the shopping list from the customer (this is what the agent perceives) and informs the customer about the sequence of shops he/she can visit in order to buy all the goods needed (these are the actions the agent undertakes). The shortest possible route is suggested, in accordance with the particular shopping list.

Since the goal is to visit all the shops from the shopping list, the particular shopping list can be regarded as a plan or a sequence of goals to achieve in order to fulfill the task completely.

It is also possible for the agent to get the exact location of a customer and a particular shop to get to. The shortest possible path to the desired shop is suggested in this case as well.

In order to realize the agent's learning process the following is to be developed: Environment model; Rewards model; Agent's memory model; Agent's behavior function; Value of the training parameter.

The environment model is a graph (Figure 3) of the different environment conditions. The nodes in the graph (Figure 1) are the shops in the exemplary shopping mall. The edges point at the shops, between which there is a transition. Then, this graph is presented by an adjacency matrix. The number of rows and columns in this matrix is equal to the number of shops in the mall. Zero is put in the matrix in a place where there is a connection between the number of a shop, set by a number of a row, and the number of a shop, given by a number of a column. Values of -1 are placed in the other positions of the adjacency matrix.

The rewards model is needed to set a goal for the agent. Reaching every shop from the customer's shopping list is such a goal. Since the agent is a goal-based one, it behavior can be changed by just setting a new goal, changing the rewards model [33]. A reward is only given when the agent gets to a particular shop.

The agent's memory is modeled by presenting it with the help of an M-matrix (Memory of the agent). The rows in the M-matrix represent the current location of the customer, while the columns are the shops, where he/she can go. It is assumed at the beginning that the agent does not have any knowledge and therefore all elements in the M-matrix are zeros.

The rule for calculating the current location of the customer at the moment of choosing the next shop to visit is as it follows:

M (current location of the customer, chosen shop to visit next) = **R**(current location of the customer, next shop) + Υ**.** **Max**[**M**(next shop, all possible shops where the customer could go from the next shop)].

The following is taken into account in the above formula: The immediate reward, obtained when the customer decides from the current location to go to a next shop: **R**(current position, chosen shop to go next); The biggest possible future reward. This is the biggest reward, chosen from among the rewards, which would have been obtained when the customer goes out of the next shop and enters any possible shop: **Max**[**M**(next shop, all possible shops where it is possible to go from the next shop)]. The value of the learning parameter Υ defines the extent, to which the agent will take into account the value of the future reward. The value of the learning parameter ϒ is within 0 to 1 (0 ≤ϒ< 1). If ϒ is closer to zero, then the agent will prefer to consider only the immediate reward. Experiments have shown that in this case it is impossible to teach the agent to achieve the goal. If ϒ is closer to one, then the agent will consider the future reward to a greater extent. This is the better option for successful training of the agent. The value of the learning parameter was experimentally chosen to be ϒ=0.8. At this value, the obtained weights for all possible actions are clearly identifiable and the process of training is reliable. A random initial position is chosen for the customer in the algorithm for training the agent. The following steps are realized until the target shop is reached:

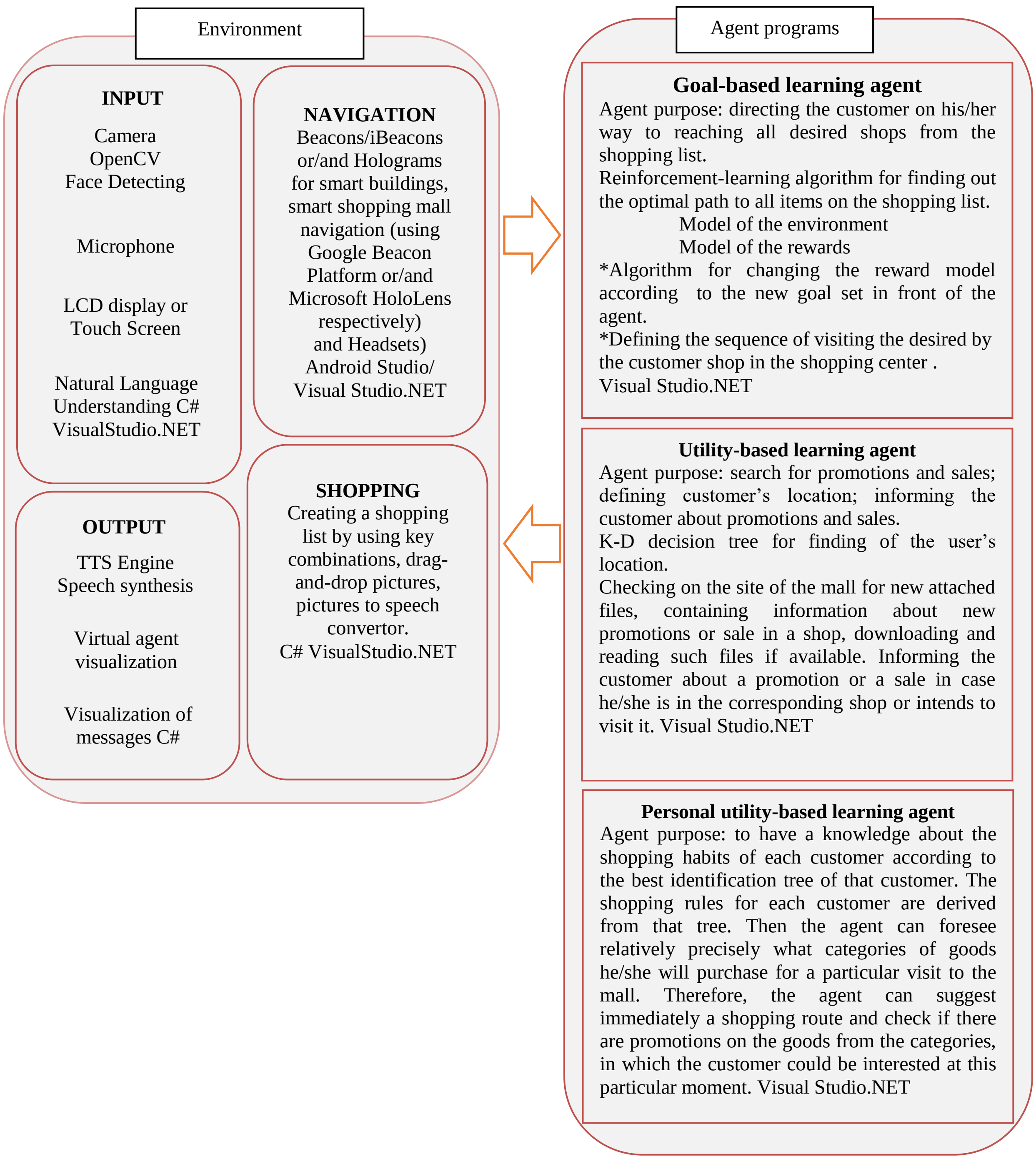


Figure 6. Specifying the task environment and smart shopping learning agents modeling.

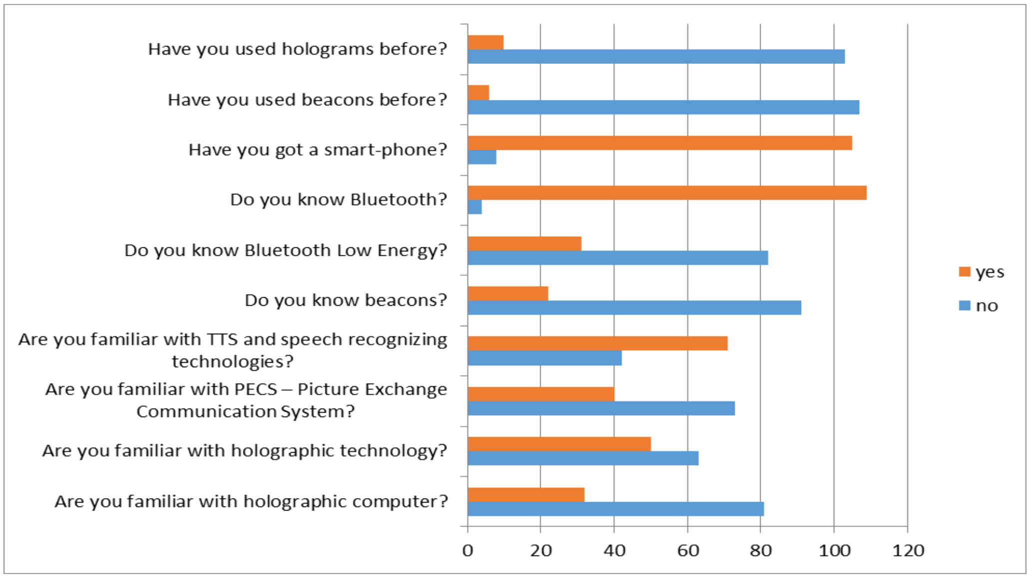


Figure 7. A survey about the interest of end customers in beacon-based services, holographic technology, PECS, TTS and Speech Synthesis technologies.

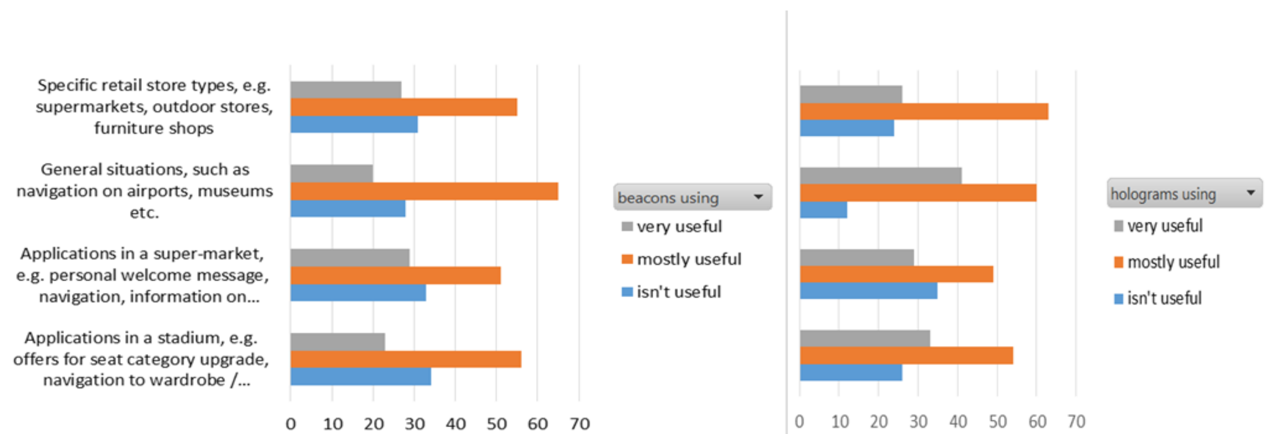


Figure 8. A survey about rediness of the user to use beacon-based services and holographic technology.

One of all possible shops is chosen, where it is possible to go from the current position. The shop, to which the customer would go next is considered. For this next position now all the shops, to which it is possible to go further are considered. The value of the highest reward is taken. The next position is then set as a current one.

C. Personal shopping Utility-Based Learning agent

The purpose of the personal shopping utility-based learning agent is to have a knowledge about the shopping habits of each customer. The personal shopping utility-based learning agent recognizes the customer at the moment he/she registers himself/herself by means of his/her shopping card.

All the purchases are grouped in categories in accordance with the shops in the shopping mall. A characteristic table with positive and negative examples is created for each category of goods. When a product from a considered category is purchased during a given visit of a customer to the mall, this is a positive example for this category of goods. For each realized purchase the following characteristics are saved in the characteristic table: season (spring, summer, autumn, winter); month (1-12); number of the week in the month (1-4); day (workday, weekend, public holiday, birthday); purchased good from another category (the categories, related to the other available shops in the mall are considered); if there were promotions of goods from the considered category at the moment of the purchase.

These characteristics are used for tests, allowing to make a classification of the examples of purchases for each category of goods. The most important characteristic for each customer is found out. This is the characteristic, for which the biggest number of examples fall into a subset of either positive or negative examples.

A subset, in which the examples are either positive or negative is called homogeneous. The examples, fallen into a non-homogeneous subset should be classified once again by using another characteristic, i.e., another criterion for classification. The procedure is repeated until all examples are classified, i.e., until all examples fall into homogeneous subsets. Thus the smallest identification tree is built. This is the tree, which classifies the examples in the best way. The shopping rules for each customer are derived from that tree. Each test (characteristic) is regarded as a prerequisite, while the result from the test, i.e., the fact if a given category of goods is purchased or not, is regarded as a conclusion. Here is an example of an obtained shopping rule for a customer: If the season is summer and it is a weekend, then the customer buys a product from the category of desserts.

When the agent has at disposal the shopping rules of a customer, it can foresee relatively precisely what categories of goods he/she will purchase for a particular visit to the mall. Therefore, the agent can suggest immediately a shopping route and check if there are promotions on the goods from the categories, in which the customer could be interested at this particular moment.

V. Degree of Development of the Proposed Cognitive Architecture Components

The goal-based learning agent, the utility-based learning agent and personal utility-based learning agent are fully developed. The head of each agent is modeled and visualized. We have used Crazy Talk 6 for emotion modeling. The decision k-d tree, identification tree algorithm and reinforcement-learning algorithm are completed and used for agents function realization. The program for creating a shopping list by using key combinations and drag and drop pictures is ready. The picture to speech converter program can pronounce all the existing pictures and the created shopping list. The agents can recognize and react to a few speech commands. They start communication with the users when detecting a face in front of themselves. A number of experiments are conducted with some Estimote beacons and a notification program [36]. The complete beacon based navigation system and the corresponding software are not ready yet, however. The holograms and holographic computer have not been used for now. We have just obtained the holographic computer and we have got start to use it now. The holographic agent's visualization are planned. Experiments in a real shopping center and with real users shopping data are planned as well. User following smart shopping cart realizations is forthcoming too.

VI. Empirical Survey

The survey was conducted at the university. The total number of 115 students were offered the questionnaire. All of the participants were between the ages of 18 to 23 years old.

A. A survey about the interest of end customers in beacon-based services, holographic technology, PECS, TTS and Speech Synthesis technologies

To investigate people's mindset towards the use of beacons, the use of holograms, drag-drop pictures, pictures to speech, TTS and Speech Synthesis an empirical study was conducted. The survey's purpose was to explore the interest of end customers in beacon-based services, holographic technology, PECS, TTS and Speech Synthesis technologies and the willingness to use them. As a base we use [9] but append some questions about new technologies.

With this end in view, we designed a questionnaire with the following tree sections. The participants were asked (Figure 7), whether they (1) own a smart-phone, (2) know Bluetooth, (3) know Bluetooth Low Energy, (4) know Holographic computer, (5) know Holographic technologies, (6) know beacons, (7) have used beacons before, (8) have used holograms before, (9) are familiar with PECS, (10) are familiar with TTS and speech recognizing technologies.

This helps to understand, whether consumers are aware of beacons. Then, they were given a short introduction of the beacon technology, holographic technology, PECS, TTS and speech recognizing technologies, as a preparation for the remaining questions. Participants were asked to assess the usefulness of typical applications, which were based on already existing scenarios by using beacon-based or holographic realizations (Figure 8): General situations, such as navigation on airports, coupons in stores, information on exhibits in museums, etc.; Specific retail store types, e.g., supermarkets, outdoor stores, furniture shops; Applications in a super-market, e.g., personal welcome message, navigation to products on the shopping list, information on products, special offers, and electronic payment at the checkout; Applications in a stadium, e.g., offers for seat category upgrade, navigation to wardrobe/restrooms, special offers for drinks and snacks.

Beacon-based technology and holographic technology are little known and the services based on them are not used widely yet, but the respondents declared willingness and readiness to use them.

Five blind people aged 45-65 also took part in the survey. These respondents were not familiar with the described technologies and had not used them before. However, they do know and use in their daily routine Internet, Skype, smartphones, e-mail, all TTS programs and desktop reading programs. They showed great enthusiasm and willingness to get acquainted with beacon-based services and holographic services for navigation in buildings.

B. A survey about the way the customers perceive the two developed agents and whether they consider their purpose useful

The capabilities of the three agents were demonstrated in front of the students. The idea of Smart Shopping and Smart Shopping Cart Agents was presented. Then, the students were asked to evaluate usefulness of the three agents, to compare their functionality and to consider the services, of which agent prefer; to say their opinion about shopping with Shopping Cart Smart Agents. Some of the questions were: Would you use a shopping cart with Intelligent Virtual Agents installed on it; would you use a shopping cart with Personal Shopping Utility Based Learning Agent installed on it; is it useful for you to be informed by Smart Shopping Cart Agents (SSCA) about the latest promotion in the shops you visit; is it useful for you if SSCA explain to you how to get to a given shop in the mall; do you think that shopping is more comfortable when communicating with SSCA; do you think the sales will go up as a result of the communication between the customers and SSCA during shopping.

It can be seen from (Figure 9) that, the customers would use SSCA and they think the agents will be useful and their presence would make the shopping practice more comfortable. The personal shopping utility agent and utility-based learning agent that search for promotions and sales are the preferred agents.

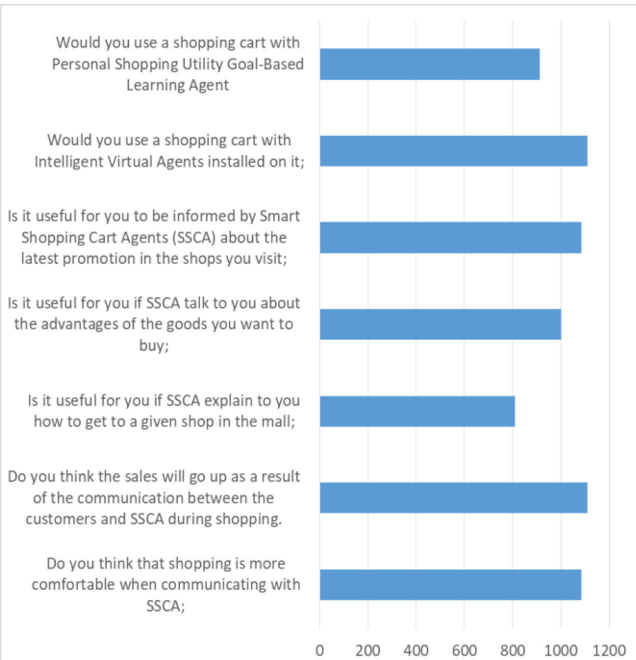


Figure 9. A survey about the way the customers perceive the three developed agents and whether they consider their purpose useful (values 1-10).

C. A survey about user's opinion and possibilities for shopping agents embodiments and user following smart shopping cart realizations

During the third survey the students were asked to guess, which of six possible shopping carts they would use. The options are: 1) a standard shopping cart with a tablet attached to it, on which the virtual shopping agents are visualized; 2) a robotized shopping cart, capable of following the user and controllable by means of gestures; 3) a robotized shopping cart with a robotized hand/arm, capable of following the user and taking/giving objects, and controllable by gestures; 4) a standard shopping cart, pushed by a human-like robot; 5) a standard shopping cart, equipped with a holographic screen, on which a hologram of a human-like shopping robot is visualized; 6) a standard shopping cart, equipped with a holographic screen, on which a 3D scene with a hologram of a human – shopping assistant is visualized.

The results from the survey are given in Figure10. The students have the strongest preference to the following two options: 1) a robotized shopping cart, capable of following the user and controllable by means of gestures; and 2) a standard shopping cart, equipped with a holographic screen, on which a hologram of a human-like shopping robot is visualized;

Having in mind the obtained results, the possibilities for realization of these two shopping cart types will be discussed.

1) A robotized shopping cart, capable of following the user and controllable by means of gestures

It is characteristic of this task that the velocities are very low; all the other shopping carts and users are regarded as dynamic obstacles to be avoided. Commodity shelves are static obstacles, which should be avoided, too. The shopping cart needs to recognize the user to be followed. It has to recognize the user's gestures and have a reaction to them. The distance between the shopping cart and the user should be kept the same as well.

Consequently, it can be summarized, that the autonomous user following is inherently a complex cognitive process associated with object recognition, static or dynamic obstacles avoidance in the environment, in which humans and robots work together; position determining; predicting the direction of movement; decision making, context understanding; working with uncertain and incomplete knowledge; vehicle communication and targeted actions and movements.

User following achievement approaches are Motion planning approaches [37]; Robust motion control approaches [38][39][40]; Game theory approach and Connectionist approach [41][42][43].

2) A standard shopping cart, equipped with a holographic screen, on which a hologram of a human-like shopping robot is visualized

It turns out that the users prefer the holographic shopping agents to look like a human, but not to be modeled as people at 100%. The users know that they are not people actually and reject them, perceiving this overly realistic model as a deception. That is why we are working to find a suitable model for the appearance of a humanoid robot that can be perceived by people well.

Modeling of the whole body of the humanoid robot is envisaged, as well as choosing a 3D scene that fits in the context of the mall.

The holographic robot must adhere to the social norms of communicating with consumers. Not to enter their personal space. Not to intrude. To keep distance when communicating with the user. The robot must behave as a mall officer. Communication and negotiation skills are needed [44], related to directing viewer to the speaker; awaiting for the speaker to finish; taking the turn; answering a question; asking a question; waiting for the answer; analysis of the answer and continuing the conversation; disengagement from the user. Skills for grasping holographic goods and presenting them in front of the user will also be needed. Such skills a holographic agent can learn by using deep neural networks (DNNs) and recurrent neural networks (RNNs) [41][42][43].

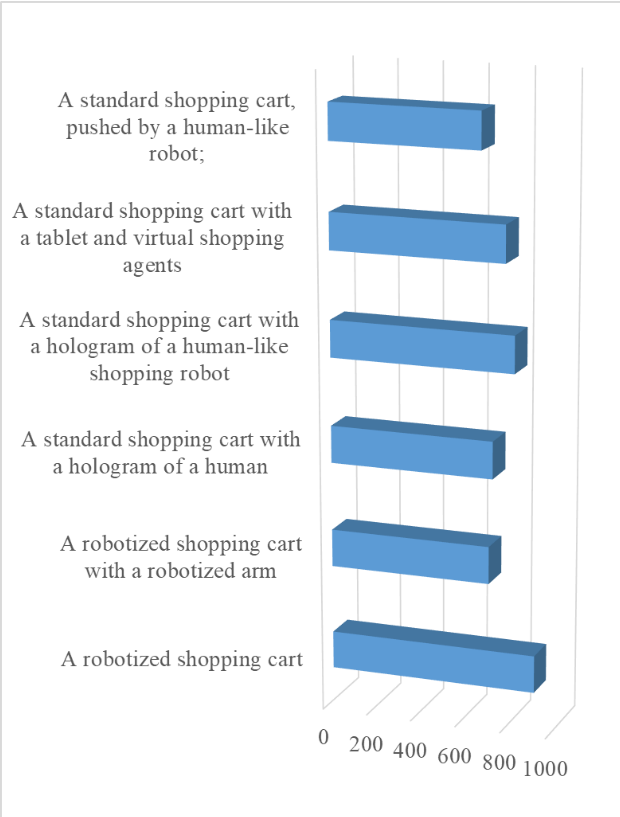


Figure 10. A survey about user's opinion for shopping agents embodiments and user following smart shopping cart realizations. (values 1-10).

The holographic agent has to understand where the user's attention is directed, to understand the user's reaction toward its actions or words. It means that scenarios have to be developed for the human-holographic agent and both reinforcement learning and Learning from Demonstration (LfD) can be used for the purpose [39][40].

In multi-agent scenarios, Learning from Demonstration paradigm already allows the robot to observe the partner's reaction that matches its movement. On the one hand, LfD speeds up reinforcement learning algorithms and can be used to re-demonstrate policies as well as to build new policies. LfD develops with the introduction and improvement of ways to obtain teacher feedback. In the latter algorithms the feedback may even be in the form of corrective advice. Therefore these algorithms could be successfully used to develop algorithms for acquiring social behavior by the holographic agents. Training can also be done by encoding the action-reaction patterns in a Hidden Markov Model (HMM) using [38].

D. A survey of the performance of the Q learning algorithm with an introduced environment measures model (a model of the environment criteria), presented as Matrix K

Q learning that is simplification of reinforcement learning is an extremely flexible method. It allows to easily find an optimal path from each position in the modeled environment to the goal state.

It is known that Imitation Learning is a way to optimize the Reinforcement learning and Q learning in particular. The task considered here is related to smart shopping realization but it does not allow a teacher to show how to reach the goal in order to achieve better results.

This is due to the availability of lots of ways for achieving a particular goal. Besides, the goal is different every time. The shops and stands in the Shopping Center that consumers want to reach are different. Some consumers look for promotional goods; others need artwork. Some customers use shopping as a therapy and want to reach the most frequently visited shops and to go through the busiest lobbies and hallways; others want to avoid the crowded zones. So in this task it is important not only to reach the goal. It is of the same level of importance how it is reached and what criteria a certain path meets.

In order to make the Q learning agent find the optimal sequence of lobbies or hallways, meeting a specific criterion, the use of environment measures model represented as matrix K is introduced.

For the purposes of the experiment the Shopping Center is represented by a graph with 17 nodes and 36 edges between them as shown in Figure 11. Every shop in the considered Shopping Center is represented as a graph. Every lobby or a hallway, connecting the shops, is represented by an edge. The busiest and most wanted to go through lobbies or hallways are marked in orange color and have a measure of 1 in the K matrix (Figure12). The secondary, distant, non-desirable pathways are marked in blue and have a measure of 2 in the K matrix (Figure12). The environment measures model presented by matrix K is similar to the environment reward model, presented by matrix R (Figure12). The values of a given criterion, to which corresponds each edge in the graph, are kept in the K matrix. The minus one (-1) in the matrix R and in the matrix K says that has no edge on this place in the graph. Now the learning algorithm is changed. The agent has to go only through those edges in the graph, which have a specific measure value.

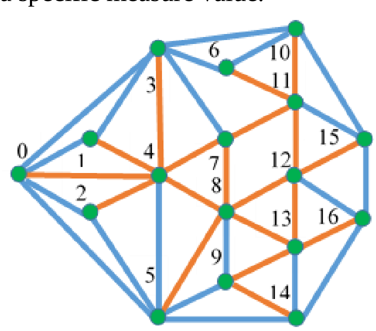


Figure 11. A Shopping Center with 17 shops and lobbies or hallways between them, presented by an undirected graph.

The experiment is conducted in the following way: a goal is set in front of the agent to reach node 15 in the graph; a reward of 100 for going through the edge, connecting nodes 11 and 15 is announced as well in reward matrix R. Other edges have zero reward (Figure 12). The black dot line denotes the optimal path found from node 0 to node 15.

First stage. No criterion is set, which the desired sequence of edges should meet in order to reach the goal.

```
//R - reward matrix
public static int [,] R = new int[,] {
    { -1,  0,  0,  0,  0,  0, -1, -1,  -1, -1, -1, -1, -1, -1, -1, -1, -1 },
    {  0, -1,  -1, 0,  0, -1, -1, -1,  -1, -1, -1, -1, -1, -1, -1, -1, -1 },
    {  0, -1, -1, -1,  0,  0, -1, -1,  -1, -1, -1, -1, -1, -1, -1, -1, -1 },
    {  0,  0, -1, -1,  0, -1,  0,  0,  -1, -1, 0, -1, -1, -1, -1, -1, -1 },
    {  0,  0,  0,  0, -1,  0, -1,  0,   0, -1, -1, -1, -1, -1, -1, -1, -1 },
    {  0, -1,  0, -1,  0, -1, -1, -1,   0,  0, -1, -1, -1, -1,  0, -1, -1 },
    { -1, -1, -1,  0, -1, -1, -1, -1,  -1, -1,  0,  0, -1, -1, -1, -1, -1 },
    { -1, -1, -1,  0,  0, -1, -1, -1,   0, -1, -1,  0, -1, -1, -1, -1, -1 },
    { -1, -1, -1, -1,  0,  0, -1,  0,  -1,  0, -1, -1,  0,  0, -1, -1, -1 },
    { -1, -1, -1, -1, -1,  0, -1, -1,   0, -1, -1, -1, -1,  0,  0, -1, -1 },
    { -1, -1, -1,  0, -1, -1,  0, -1,  -1, -1, -1,  0, -1, -1, -1, 0, -1 },
    { -1, -1, -1, -1, -1, -1,  0,  0,  -1, -1,  0, -1,  0, -1, -1,  100, -1 },
    { -1, -1, -1, -1, -1, -1, -1, -1,   0, -1, -1,  0, -1,  0, -1,  0,  0 },
    { -1, -1, -1, -1, -1, -1, -1, -1,   0,  0, -1, -1,  0, -1,  0, -1,  0 },
    { -1, -1, -1, -1, -1,  0, -1, -1,  -1,  0, -1, -1, -1,  0, -1, -1,  0 },
    { -1, -1, -1, -1, -1, -1, -1, -1,  -1, -1,  0,  0,  0, -1, -1, -1,  0 },
    { -1, -1, -1, -1, -1, -1, -1, -1,  -1, -1, -1, -1,  0,  0,  0,  0, -1 }};

//K- measure matrix
public static int[,] K = new int[,] {
    { -1,  2,  2,  2,  1,  2, -1, -1,  -1, -1, -1, -1, -1, -1, -1, -1, -1 },
    {  2, -1,  -1, 2,  1, -1, -1, -1,  -1, -1, -1, -1, -1, -1, -1, -1, -1 },
    {  2, -1, -1, -1,  1,  2, -1, -1,  -1, -1, -1, -1, -1, -1, -1, -1, -1 },
    {  2,  2, -1, -1,  1, -1,  2,  2,  -1, -1,  2, -1, -1, -1, -1, -1, -1 },
    {  1,  1,  1,  1, -1,  2, -1,  1,   1, -1, -1, -1, -1, -1, -1, -1, -1 },
    {  2, -1,  2, -1,  2, -1, -1, -1,   1,  2, -1, -1, -1, -1,  2, -1, -1 },
    { -1, -1, -1,  2, -1, -1, -1, -1,  -1, -1,  2,  1, -1, -1, -1, -1, -1 },
    { -1, -1, -1,  2,  1, -1, -1, -1,   1, -1, -1,  1, -1, -1, -1, -1, -1 },
    { -1, -1, -1, -1,  1,  1, -1,  1,  -1,  2, -1, -1,  1,  1, -1, -1, -1 },
    { -1, -1, -1, -1, -1,  2, -1, -1,   2, -1, -1, -1, -1,  1,  1, -1, -1 },
    { -1, -1, -1,  2, -1, -1,  2, -1,  -1, -1, -1,  1, -1, -1, -1,  2, -1 },
    { -1, -1, -1, -1, -1, -1,  1,  1,  -1, -1,  1, -1,  1, -1, -1,  2, -1 },
    { -1, -1, -1, -1, -1, -1, -1, -1,   1, -1, -1,  1, -1,  1, -1,  1,  2 },
    { -1, -1, -1, -1, -1, -1, -1, -1,   1,  1, -1, -1,  1, -1,  2, -1,  1 },
    { -1, -1, -1, -1, -1,  2, -1, -1,  -1,  1, -1, -1, -1,  2, -1, -1,  2 },
    { -1, -1, -1, -1, -1, -1, -1, -1,  -1, -1,  2,  2,  1, -1, -1, -1,  2 },
    { -1, -1, -1, -1, -1, -1, -1, -1,  -1, -1, -1, -1,  2,  1,  2,  2, -1 }};
```

Figure 12. Reward matrix R and Measure matrix K.

The optimal path found from node 0 to the goal is given in Figure 13. It can be seen that the path goes through edges with a different value of the criterion, set in the K matrix.

Second stage. The agent receives a requirement to reach the goal by going only through edges with a measure value of 1. The optimal path found from node 0 to the goal is given in Figure 14. As it can be seen, the path goes only through edges with a measure value of 1 for the criterion, set in the K matrix.

Third stage. The agent has to reach the goal by going only through edges, having a measure value of 2. The optimal path for this case, starting from node 0 and going to the goal, is shown in Figure 15. As it shows, the path goes only through the edges with a measure value of 2 for the criterion, set in the K matrix.

In the example under consideration, there are primary and secondary paths that connect all locations in the example Shopping Center. There might be a situation, in which a primary or a secondary path to a given place is missing. Then the algorithm can be modified by allowing the agent to go through a certain number of edges, which do not correspond to the value of the criterion "measure" in the K matrix.

An advantage of the proposed modification of the Q learning algorithm is that it allows the agent to give explanation of the reasons why a given path to the goal has been chosen. In addition, the proposed modification allows to introduce various criteria for a path choice. If the criteria from Maslow's theory of personality motivation are used [45], a model of a system of values could be developed using different scenarios.

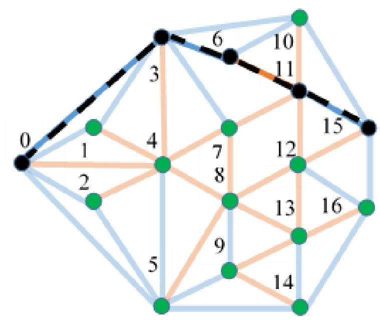


Figure 13. Optimal path from node 0 to node 15. Requirement for measure value not set.

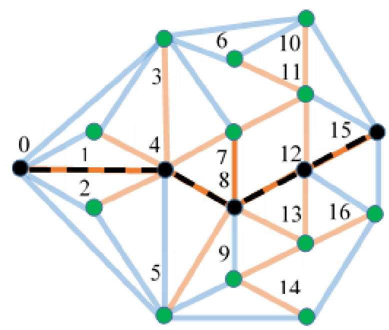


Figure 14. Optimal path from node 0 to node 15. Requirement for measure value: to be equal to 1.

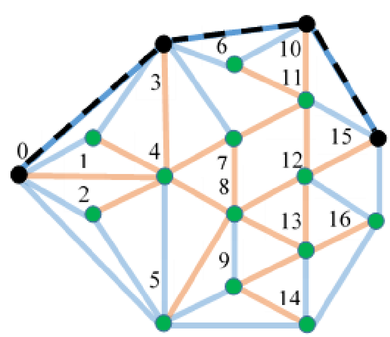


Figure 15. Optimal path from node 0 to node 15. Requirement for measure value: to be equal to 2.

E. Study of the learning parameter ϒ

It is known that the value of the learning parameter ϒ is within 0 to 1 (0 ≤ϒ< 1). The aim is to study if any of these values is more appropriate to be preferred during a Q learning agent training.

The experiment is conducted in the following way: the graph in Figure 11 is considered; a goal is given to the agent and the training gets started. Twelve experiments are carried out actually with each value of the learning parameter ϒ - from 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 to 1, respectively.

First stage: The number of episodes required to train the agent to reach the goal using the optimal path from each node in the graph is considered. The exploration by the agent of each node-to-node path until it reaches the goal node is called an episode.

The training is considered completed when any further change of the assessment of each edge in the graph does not lead to a change in the optimal paths found. The results from this stage are shown in Figure 16. It is obvious that values from 0.3 to 0.9 of the learning parameter ϒ are appropriate for realizing the process of training. When these values are 0, 0.1, 0.2 and 1, no paths to reach the goal are found.

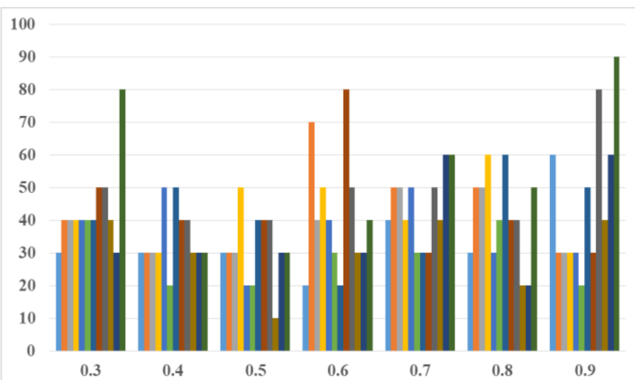


Figure 16. Number of episodes needed to find the optimal paths at values of the learning parameter from 0 to 1 during a Reinforcement learning agent training.

Second stage. The number of episodes required to train the agent to reach the maximum value in the assessment of each edge in the graph is considered. During the training process the assessment of each edge in the graph increases until it reaches its maximum value. When the maximum value is reached, the further training no longer changes the assessments of the edges in the graph. The results from this stage are presented in Figure 17.

Comparing the results of the two stages of the experiment, it can be seen that finding the optimal paths to the goal requires much fewer episodes than reaching the maximum value of each edge in the graph. Besides, the values of the learning parameter of 0.8 and 0.9 offer the greatest possibilities both for training and for finding optimal paths. As it can be seen in Figure 17, these learning parameter values require an average of 800 and 1600 episodes, respectively, until the maximum value in the assessment of each edge in the graph is reached.

VII. FUTURE WORK

Many tasks remain to be solved. The work on the development of the Reinforcement learning algorithm will continue in the first place; opportunities for modeling the training agent's value system will be looked for; efforts will be put to modeling a system for generating explanations by the trained agent. Using the holographic computer, it is now possible to model and visualize a virtual advertising agent. It is assumed that the communication with such an agent will be engaging and helpful to consumers. As mentioned in this article, there is a lot of interest in modeling a robotic shopping cart to follow the consumer. Efforts will therefore be made to this end. For example, it is important to combine and share intelligent behaviors such as: wander behavior; path following; collision avoidance; obstacle and wall avoidance; patrol between a set of points; flee behavior.

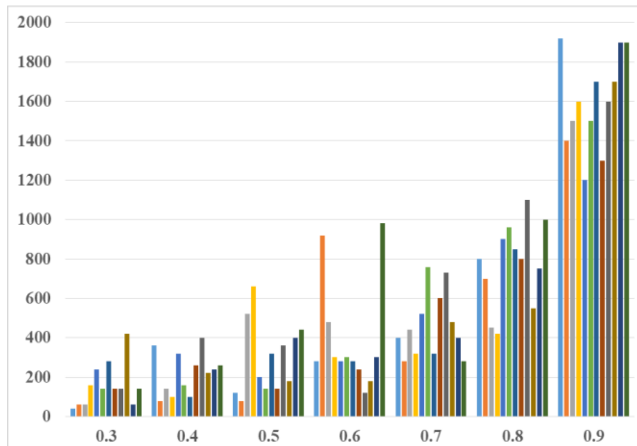


Figure 17. Number of episodes needed to reach the maximum value of the assessment for each edge in the graph under consideration.

VIII. CONCLUSION

The paper describes the design and implementation of an intelligent Smart Shopping Cart Learning Agents prototype and their environment. The system differs from other intelligent systems by the combination of machine learning techniques, beacon-based navigation and/or hologram-based navigation in the mall, the integration of Picture Exchange Communication System in it and by its language understanding and speech synthesis capabilities, drag-and-drop techniques and keyboard button combinations enabled access.

The task environment is partially observable, cooperative and a multi-agent environment. The shopping world is deterministic with some stochastic and uncertainty elements. The task environment is episodic and can be realized either as static or as semi-dynamic.

The utility-based agent uses a decision k-d tree to quickly find where, in which shop the customer is located according to his/her coordinates. It getting to the new promotion in the shopping mall according to user‘s shopping list and inform them.

Reinforcement learning algorithm is used for the other Goal-based learning agent. The agent gets the shopping list from the customer and informs the customer about the sequence, in which he/she can visit the shops to buy all needed goods.

The personal utility-based agent makes the best identification tree according to the shopping data of each user. That way the agent knows their shopping habits and can suggest a shopping route and check if there are promotions on the goods from the categories, in which the customer could be interested.

The performance measure, to which the shopping agents are aspired includes getting to the correct shop in the shopping mall; getting to the new promotion in the shopping mall according to user‘s shopping list or according to user’s shopping habits; minimizing the path when going through the shops from the shopping list; maximizing customer comfort; maximizing purchases; and enabling people with different communication capabilities to navigate in big buildings and in particular to shop in big shopping centers.

The empirical survey conducted with a limited number of users showed their positive mindset for using such Smart Shopping Cart Learning Agents in indoor spaces. The utility-based learning agent that search and informs for promotions and sales is the preferred one.

The performance of the Q learning algorithm with an introduced environment measures model (a model of the environment criteria) is proposed and explored. Study of the learning parameter ϒ is presented.

In the future work, it is intended to realize user following smart shopping cart and holographic visualization of the shopping agents.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support provided within the Ministry of Education and Science of Bulgaria Research Fund Project FNI DN17/11.

REFERENCES

[1] Dilyana Budakova, Lyudmil Dakovski, Veselka Petrova-Dimitrova, “Smart Shopping Cart Learning Agents Modeling and Evaluation,” The Eleventh International Conference on Advanced Cognitive Technologies and Applications (COGNITIVE 2019) IARIA, 05-09 May 2019, Venice, Italy, pp. 12-19, ISSN: 2308-4197, ISBN: 978-1-61208-705-4.

[2] Y. J. Chang, S. M. Peng, T. Y. Wang, S. F. Chen, Y. R. Chen, and H. C. Chen, “Autonomous indoor wayfinding for individuals with cognitive impairments,” Journal of NeuroEngineering and Rehabilitation, 7(1), pp. 1– 13, 2010.

[3] V. Kulyukin, C. Gharpure, J. Nicholson, and G. Osborne, “Robot-assisted wayfinding for the visually impaired in structured indoor environments,” Autonomous Robots, Springer, 21(1), pp. 29–41, 2006.

[4] L. Niua and Y. Song, “A schema for extraction of indoor pedestrian navigation grid network from floor plans,” In The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, volume XLI-B4, Prague, Czech Republic, pp. 325-330, 2016.

[5] My Dream Companion project – YGA. [Online]. Available from: https://www.yga.org.tr/en/visually-impaired-techologies [retrieved: 12, 2019].

[6] Y. Zhuang, J. Yang, Y. Li, L. Qi, and N. El-Sheimy, “Smartphone-based indoor localization with Bluetooth low energy beacons,” Sensors, April 2016. 16(5), pp. 596, doi: 10.3390/s16050596.

[7] D. Ahmetovic, et al. “Navcog: A navigational cognitive assistant for the blind,” In International Conference on Human Computer Interaction with Mobile Devices and Services, ACM, 2016, pp. 90-99.

[8] A. Ch. Seyed, N. Vinod, and S. Kaushik, “IBeaconMap: Automated Indoor Space Representation for Beacon-Based Wayfinding,” Human-Computer Interaction arXiv: 1802.05735v1 [cs.HC], USA, 2018.

[9] A. Thamm, J. Anke, S. Haugk, and D. Radic, “Towards the Omni-Channel: Beacon-based Services in Retail,” International Conference on Business Information Systems (BIS 2016), Springer, Leipzig, Germany, July 2016, pp. 181-192. July 2016. DOI: 10.1007/978-3-319-39426-8_15.

[10] Google Beacon Platform. [Online]. Available from: https://www.youtube.com/watch?v=0QeY9FueMow, [retrieved: 12, 2019].

[11] Microsoft HoloLens holographic computer. [Online]. Available from: https://developer.microsoft.com/en-us/windows/mixed-reality/, [retrieved: 12, 2019].

[12] J. Cassell, et al. “MACK: Media lab Autonomous Conversational Kiosk,” Imagina’02, Monte Carlo, 2002 vol. 2, pp. 12–15. [Online]. Available from: https://www.media.mit.edu/gnl/pubs/imagina02.pdf, [retrieved: 12, 2019].

[13] L. McCauley and S. D’Mello, “MIKI: A speech enabled intelligent kiosk,” IVA 2006, LNAI 4133, Springer, 2006, pp.132-144.

[14] Tomorrow‘s digital signage, today with 3d holographic kiosks, Intel IoT Digital kiosks, Solution brief, [Online], Available from: https://d1io3yog0oux5.cloudfront.net/_35cb5a593466915af09244e7df3c153e/provision/db/255/661/pdf/Intel+Provision+3D+Brief+2016.pdf, [retrieved 12, 2019].

[15] A. Bondy and L. Frost, PECS - Picture Exchange Communication System. [Online]. Available from: https://pecsusa.com/apps/, [retrieved: 12, 2019].

[16] C. Lord, M. Rutter, P. C. Dilavore, and S. Risi, “ADOS - Autism Diagnostic Observation Schedule,” Western Psychological Services, 2002.

[17] National Autism Resources Inc. [Online]. Available from: https://www.nationalautismresources.com/the-picture-exchange-communication-system-pecs/, [retrieved: 12, 2019].

[18] What is PECS. [Online]. Available from: https://pecs-unitedkingdom.com/pecs/, [retrieved: 12, 2019].

[19] JAWS. [Online]. Available from: http://www.freedomscientific.com/Products/Blindness/JAWS [retrieved: 12, 2019].

[20] NVDA. [Online]. Available from: https://www.nvaccess.org/, [retrieved: 12, 2019].

[21] Vocalizer Android Dariya Bulgarian voice. [Online]. Available from: https://play.google.com/store/apps/details?id=es.codefactory.vocalizertts&hl=bg, [retrieved: 12, 2019].

[22] Innoetics TTS Reader Female Bulgarian Voice IRINA. [Online]. Available from: https://www.innoetics.com/, [retrieved: 12, 2019].

[23] S. Russell and P. Norvig, “Artificial Intelligence A Modern Approach”, Prentice Hall, Third Edition, 2010, ISBN-13 978-0-13-604259-4, ISBN-10 0-13-604259-7.

[24] Kuka robot. [Online]. Available form: https://cyberbotics.com/doc/guide/robots, [retrieved: 12, 2019].

[25] N. Ç. Kılıboz and U. Güdükbay, “A hand gesture recognition technique for human–computer interaction,” Journal of Visual Communication and Image Representation, Elsevier, Volume 28, pp. 97-104, April 2015.

[26] N. G. Shakev, S. A. Ahmed, A. V. Topalov, V. L. Popov, and K. B. Shiev, “Autonomous Flight Control and Precise Gestural Positioning of a Small Quadrotor,” Learning Systems: From Theory to Practice, Springer, pp. 179-197, 2018. |

[27] E. Coupeté, et al. “New Challenges for Human-Robot Collaboration in an Industrial Context: Acceptability and Natural Collaboration,” Fifth workshop "Towards a Framework for Joint Action", IEEE RO-MAN, 2016, pp. 1-4.

[28] S. R. Fletcher and P. Webb, “Industrial robot ethics: facing the challenges of human-robot collaboration in future manufacturing systems,” A world with robots: international conference on robot ethics (ICRE 2015) Springer, 2015, pp. 159-169.

[29] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” Computer Vision and Pattern Recognition (CVPR 2001) IEEE, Vol.1, 2001, pp. I-511-I-518. doi: 10.1109/CVPR.2001.990517.

[30] The Intel Open Source Computer Vision Library, vol. 2006. [Online]. Available from: https://software.intel.com, [retrieved: 12, 2019].

[31] P. H. Winston, “Artificial Intelligence,” Addison-Wesley Publishing Company, ISBN-13: 978-0201533774 ISBN-10

[32] Y. Y. Song and Y. Lu, “Decision tree methods: applications for classification and prediction,” Shanghai Archives of Psychiatry, Vol. 27, No.2, pp. 130-135, 2015. doi: 10.11919/j.issn.1002-0829.215044, PMID: 26120265.

[33] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, The MIT Press, Cambridge, London, England, 2014. [Online]. Available from: http://incompleteideas.net/book/ebook/the-book.html, [retrieved: 12, 2019].

[34] A. Gosavi, “Reinforcement Learning: A Tutorial Survey and Recent Advances,” INFORMS Journal on Computing, Vol. 21 No.2, pp. 178-192, 2008.

[35] R. R. Torrado, P. Bontrager, J. Togelius, J. Liu, D. Perez-Liebana, “Deep Reinforcement Learning for General Video Game AI,” IEEE Conference on Computatonal Intelligence and Games, CIG. 2018-August, 10.1109/CIG.2018.8490422

[36] Estimote Beacons. [Online]. Available from: https://estimote.com/, [retrieved: 12, 2019].

[37] A. Best, S. Narang, D. Barber, and D. Manocha, “AutonoVi: Autonomous Vehicle Planning with Dynamic Maneuvers and Traffic Constraints”, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2017) IEEE, 2017, DOI: 10.1109/IROS.2017.8206087.

[38] D. Lee, C. Ott, and Y. Nakamura, “Mimetic communication model with compliant physical contact in human-humanoid interaction”, International Journal of Robotics Research, SAGE, Volume 29 issue13, pp. 1684–1704, November 2010.

[39] B. Argall, “Learning Mobile Robot Motion Control from Demonstration and Corrective Feedback”, Robotics Institute Carnegie Mellon University Pittsburgh, PA 15213, March 2009.

[40] H. B. Amor, D. Vogt, M. Ewerton, E. Berger, B. Jung, J. Peters, “Learning Responsive Robot Behavior by Imitation,” IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013) IEEE, Tokyo, Japan, November 3-7, 2013, pp. 3257-3264.

[41] K. Takahashi, K. Kim, T. Ogata, S. Sugano, “Tool-body assimilation model considering grasping motion through deep learning,” Robotics and Autonomous Systems, Elsevier, Volume 91, pp. 115–127, 2017.

[42] G. I. Parisi, J. Tani, C. Weber, S. Wermter, “Emergence of multimodal action representations from neural network self-organization,” Cognitive Systems Research, Elsevier, Volume 43, pp. 208-221, June 2017.

[43] K. Noda, H. Arie, Y. Suga, and T. Ogata, “Multimodal integration learning of robot behavior using deep neural networks,” Robotics and Autonomous Systems, Elsevier, Volume 62, pp. 721–736, 2014.

[44] S. Monahan, et al., “Autonomous Agent that Provides Automated Feedback Improves Negotiation Skills,” Chapter in Artificial Intelligence in Education, Springer International Publishing, volume 10948, 2018.

[45] A. Maslow, Motivation and Personality, Harper & Row Publishers, 1970. [Online]. Available from: https://www.academia.edu/29491165/Motivation_and_Personality_BY_Abraham_Maslow, [retrieved: 12, 2019].

Performance Catalogs for Just-In-Time Delivery of Web-based Natural Language Processing Services

Soheila Sahami

Natural Language Processing Group
University of Leipzig, Germany
Email: sahami@informatik.uni-leipzig.de

Thomas Eckart

Natural Language Processing Group
University of Leipzig, Germany
Email: teckart@informatik.uni-leipzig.de

Abstract—**The modern production of industrial goods is often based on Just-In-Time delivery of required resources. This paradigm makes a variety of demands on the infrastructure in which production takes place or in which services are provided and consumed. The reasons for introducing this strategy and its potential benefits are in large parts applicable to the area of Web-based Natural Language Processing services. This contribution focuses on prerequisites and potential outcomes of a Just-In-Time-capable infrastructure of Natural Language Processing services using examples in the context of real-world research projects. The benefits of this endeavor are sketched with a focus on the ongoing development of large scale service delivery platforms like the European Open Science Cloud, CLARIN, or similar projects. As a major outcome, the creation of "performance catalogs" containing extensive information about the run-times and performance of every involved tool is seen as an essential precondition for these environments.**

Keywords*–*Just-In-Time Delivery; Performance Catalog; Natural Language Processing Services; Research Infrastructure; Cluster Computing.

I. INTRODUCTION

This article is an extended version of the conference publication "Just-In-Time Delivery for NLP Services in a Web-Service-Based IT Infrastructure" presented at the *Adaptive Workflow Management for NLP Data Processing workshop* at ADAPTIVE 2019 [1]. This article provides a more detailed discussion of relevant parameters defining performance profiles of services. These parameters, which may impact the performance of services, such as language and quality of input material are defined and an extended set of experiments depending on different hardware configurations and tool parameters are conducted.

In industrial production environments, providing resources immediately before they are required in the context of a larger production chain – typically called *Just-in-Time Delivery* (JIT delivery) – is a standard procedure for many decades now. The transfer of this concept into the area of information technology offers a new competitive opportunity that promises significant advancements, such as faster responses, improved quality, flexibility, and reduced storage space [2].

The use of Natural Language Processing (NLP) applications – i.e., tools for preprocessing, annotation, and evaluation of text material – is an integral part for a variety of applications in scientific and commercial contexts. Many of those tools are nowadays available and actively used in service-oriented environments, where – often complex – hardware and software configuration is hidden from the user. In the context of large research infrastructures, like CLARIN [3] or DARIAH [4], or cross-domain projects, like the European Open Science Cloud (EOSC) [5], one of the key goals is to facilitate the use of services which, are seen as integral and indispensable building blocks of a modern scientific landscape.

These research infrastructure projects can be seen as driving forces for current trends in the dissemination and delivery of tools and services. However, in many respects, they are undergoing a similar development as already completed in many commercial areas where delivery and use of services are performed in an industrial scale. Systematic assessment and improvement of quality, measurement of throughput times or other criteria are prerequisites for the use of services even for time-critical applications [6].

One of the potential outcomes and goals of a more "industrialized" infrastructure could be a just-in-time delivery of services, providing the benefits – while requiring comparable prerequisites – already accustomed in the industrial production of goods. They include reduced response times, reduction of required storage facilities [7], and more. However, those topics are hardly addressed in today's text-oriented research infrastructures. Some of the missing preliminary work that is required to offer JIT delivery of linguistic services – like the transparency of the process and its sub-processes, deep knowledge about required resources and execution times – are addressed in this contribution.

One of the important challenges in JIT delivery is the applied strategy to address the reliability and predictability of services [8]. In IT infrastructures, utilizing fault-tolerant techniques is one of the solutions to improve the reliability of an application. Parallelised implementations using cluster-based processing architectures are technologies that are utilized to decrease run-times and to enable the processing of large scale resources. Furthermore, they provide a helpful means to configure processes in a dynamic manner. This allows suggesting several configurations based on the available resources of the service provider or temporal requirements of the user. Clear information about potential expenses and the estimated delivery time for each configuration gives users a means to select a suitable service (or service chain) or service configuration that fits their needs best.

This also helps users to have a clear strategy for data storage, duration of data retention, and delivery time. These features have the potential to enhance the user's satisfaction and provide added values that lead to a stronger position in competitive industrialized IT infrastructures.

In this contribution, we present examples of Natural Language Processing services with a focus on their transparency regarding execution times and required resources. As a result, valid resource configurations can be chosen considering available resources and expected delivery times. It should be pointed out that multiple NLP tools have been implemented to prove the suggested approach and more tools and the other methods – such as machine-learning-based or hybrid approaches – can be contemplated as the future extension of this research.

The following Section II gives more details about the parallelism of just-in-time delivery of IT services and their industrial counterpart. Section III describes the used methodology and its characteristics, technical approaches and tools. Section IV describes the research infrastructures that can be used as the base for these tools and compares them with commercial counterparts. Section V explains the chosen parameters followed by Section VI showing the outcomes of the experiments that are performed by assigning various cluster and tool configurations varying every individual parameter. Sections VII and VIII illustrate and discuss the outcomes and results and are followed by a brief conclusion of this contribution and a short outlook in Section IX.

II. Just-in-time Delivery of IT-Services

Just-in-time delivery (or just-in-time manufacturing) is a management concept that was introduced in the Japanese automotive industry [9] and was adopted for many other areas of production and delivery of goods since. Based on experiences and best practices, catalogues were developed that contain extensive lists of requirements that make the usage of JIT delivery chains manageable and trustworthy.

Established requirements deal with all kinds of aspects in the organizational, legal and technical environment of companies and organizations that are involved in the overall process. At least a subset of those requirements is directly transferable to activities in IT processes and infrastructures [2], including the more recent deployment, provisioning, and use of services in complex Service-Oriented Architectures (SOAs). This contains procedures and guidelines like the strict use of a "pull-based" system, process management principles with a focus on flow management, adequate throughput, and continuous assessment of quality and fitness of used processes and outcomes. Its obvious benefits have made the underlying policies also a cornerstone of modern agile management principles (c.f. [10]).

There is some research about transferring the JIT concept and its principles to service-oriented environments, like the ones gaining momentum in the area of NLP applications. In the context of such IT services chains, the term *just-in-time* can be understood in different ways. It is often referring to the specific decision for a set (or chain) of services – out of a potentially large inventory of compatible services from different providers – as part of the typical discovery/bind-process *at run-time*, i.e., without a fixed decision for specific providers or even prior knowledge about the current inventory of available services. This is sometimes called "just-in-time integration" of services (for example in IBM's developer documentation [11]).

Many essential requirements for a JIT integration are already handled in existing frameworks – for example CLARIN's WebLicht [12] –, like compatibility-checks of all services regarding their input parameters and generated output, a systematic monitoring of all participating service providers of the federation, or – in parts – even adherence to legal constraints.

A different approach for services-based JIT delivery focuses on the estimated time of arrival (ETA) of the required results for a specific service chain. This is especially important considering the growing amount of text material that is required to be processed. Most academic providers of NLP services are not able to guarantee acceptable processing times – or the completion of large processing jobs at all – with their current architectures for (very) large data sets in the context of SOAs. However, this kind of functionality is required to reach new user groups and to make them competitive offerings in comparison with the other (including commercial) service providers. This aspect is hardly addressed in previous and current projects of the field but gains significance in current attempts to make scientific working environments more reliable and trustworthy with a strong focus on cloud-based solutions (like the European Open Science Cloud EOSC).

A key idea is the incremental creation and adaptation of "performance profiles" for all elements of a provider's service catalog. This contains the identification of all relevant parameters of a tool and well-founded empirical knowledge about their effects on the run-time of every single NLP task for all kinds of plausible inputs. This requires a processing architecture that is able to dynamically allocate resources for each assigned job while minimizing (or even eliminating) the effects of other jobs that are executed in parallel.

In the following, we will describe a concrete example of such a service performance profile depending on the assigned hardware configuration and workflow arrangement and sketch its benefits.

III. Approach

In this section, the essential features of a JIT delivery in IT systems are explained. This is followed by a discussion of the chosen technologies and their specific features relevant to the context of this contribution. Finally, the implemented NLP tools and utilized resources are described.

A. Essential Features

Generally speaking, the degree of user acceptance in regard to JIT delivery in IT systems relies on several parameters such as diversity of the tools, state-of-the-art technologies, support for complex requests, and quality of the services. However, reliability and predictability of services [8] are two of the most important parameters that have large effects on the success of JIT delivery in IT systems. Accordingly, it is essential to choose technologies that support these two parameters.

1) Reliability: The international standard ISO/IEC/IEEE 24765:2017 has defined reliability as the "degree to which an object or an object's services provide agreed or expected functionality during a defined time period under specified conditions." A highly reliable application performs the expected functionality and is able to avoid, detect, and repair the failures in a way that users do not notice them. In other words, the application should be enabled to continue normal operation in the presence of faults, or be fault-tolerant. Fault tolerance is defined in ISO/IEC/IEEE 24765:2017 as the "degree to which a system, product or component operates as intended despite the presence of hardware or software faults." [13]

In the context of this article, fault tolerance refers only to the ability of a system to detect a hardware fault and immediately switch to a redundant hardware component. To increase the degree of fault tolerance and provide more reliable applications, distributed systems are implemented using cluster-based processing architectures [14] [15].

2) Predictability: In general, a predictable system is a system that has agreed or expected functionality for the possible states. Predictability is the degree to which a system or a component of the system behaves as expected in different situations. In the field of information technology and computer science, this criterion is known as the availability of an application. Availability refers to the probability that the system will operate continuously without failure during a specific time period [16]. High availability, as a critical feature, is addressed by cluster computation technology. This technology uses redundancy in order to improve the degree of availability [17].

B. Technical Approach

For services, where response times are critical, used technology should support technical features like fault tolerance and high availability. In addition, the technology should be able to process large scale data in a feasible time to satisfy the demand for the processing and analyzing of a rapidly growing amount of text data.

For this contribution, Apache Hadoop clusters and the Apache Spark execution engine are used to address the topics of fault tolerance and high availability. This approach supports the distribution and parallelization of tasks and significantly improves execution and response times.

Apache Hadoop is a popular framework to store and process large-scale data in a distributed computing environment that – with its built-in mechanisms – provides a high degree of parallelism, robustness, reliability, and scalability. One beneficial approach in this distributed processing technology is to run processes wherever the data is located. This means Hadoop initially distributes the data to multiple machines in the cluster and then assigns computation tasks based on the locality of data. This location-based assignment reduces the communication overhead between participating nodes [18]. Apache Hadoop's large ecosystem consists of the Hadoop Kernel, MapReduce, Hadoop Distributed File System (HDFS), Apache Spark, and some other components [19].

In the context of Big Data, partitioning the data across a number of separate machines is obligatory. HDFS, as the storage layer of Apache Hadoop, is a distributed file system that provides access to the data across the Hadoop clusters. HDFS is a highly fault-tolerant distributed storage system that is able to handle the failure of storage infrastructure without losing the data. The application data – the file contents – is split into large blocks with a default size of 128 MB, and each block of the file is independently replicated at multiple machines, where the size of the data blocks and the number of the replicas can be defined by the user on a file-by-file basis. Given the importance of the file system namespace, HDFS makes it resilient to failure by providing a backup of the file system and also periodically creating a copy of the list of blocks belonging to each file. Duplicating the copies of data blocks multiple times in the individual data nodes increases the reliability, allowing the system to tolerate node failure without suffering data loss. In the event of failure, this information facilitates the recovery of data [20].

Apache Spark is also a general-purpose cluster computing framework for big data analysis with an advanced In-Memory programming model. It uses a multi-threaded model where splitting the tasks on several executors improves processing times and fault tolerance. Apache Spark uses the Resilient Distributed Dataset (RDD), a data-sharing abstraction that is designed as a fault-tolerant collection and is capable of recovering lost data after the failure using the lineage approach. In this approach, Apache Spark keeps a graph of the data transformations during the construction of an RDD. In the event of failure, it reruns all failed operations to rebuild the lost results. The RDDs are persisted and executed completely in RAM – In-Memory Databases (IMDB) –, therefore generating and rewriting the recovered data are performed as fast processes [21] [22].

C. NLP Tools and Resources

Using Hadoop-based cluster computing architecture, a variety of typical NLP tools were implemented, including sentence segmentation, pattern-based text cleaning, tokenizing, language identification, and named entity recognition [23]. These tools use Apache Hadoop as their framework, Apache Spark as execution engine and HDFS as file system and storage manager. Furthermore, their atomic design facilitates to integrate them into SOA-based annotation environments.

In order to have an accurate estimation of execution times, a variety of benchmarks were carried out for the implemented tools. For these benchmarks, we have used a cluster provided by Leipzig University Computing Center [24]. Table I illustrates the characteristics of this cluster [25]. In this cluster, Apache Hadoop 2.7.3 is used as framework to process the data in the distributed environment and store the data on HDFS. Apache Spark 2.3.0 is used as the execution engine and YARN is configured as the resource manager.

TABLE I. Cluster Characteristics

Number of nodes	CPUs	Hard drives	RAM	Network
90	6 Cores per Node	>2PB in total	128GB per Node	10Gbit/s Ethernet

As an example, one of these benchmarks evaluates the duration of sentence segmentation for datasets of German documents with sizes from 1 to 10 Gigabytes using different cluster configurations. The cluster configuration varies in the number of assigned executors (1 to 32 nodes) and allocated memory per executor (8 or 16 GB). Each test was repeated three times; average execution time over all three runs was used for the following statistics. Figures 1 and 2 show these execution times for sentence segmentation from 1 to 10 GB of input data with different resource configurations.

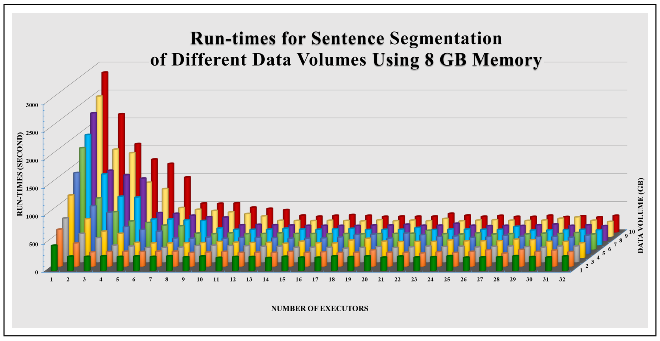


Figure 1. Run-times for segmenting 1 to 10 GB text materials using 1 to 32 executors and 8 GB memory per executor.

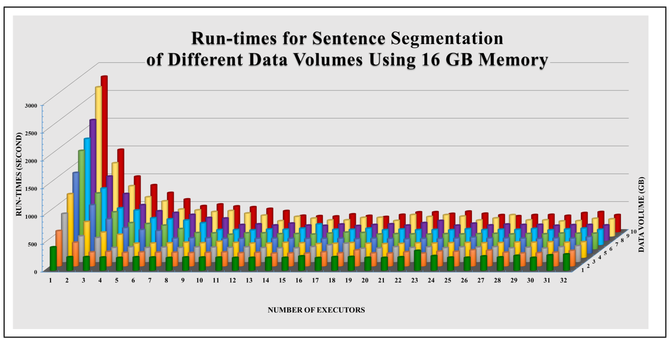


Figure 2. Run-times for segmenting 1 to 10 GB text materials using 1 to 32 executors and 16 GB memory per executor.

A brief explanation is given below for each of the tools.

1) Text Cleaner: This tool is a parallel implementation of a pattern-based text cleaner that uses sets of rules to detect invalid character patterns in text documents. *General rules* – being language-independent – contain generic unlikely patterns in written language. *Language-based rules* consist of individual rules for each of the defined languages to detect language specific ill-formed sentences. *Genre-based rules* are defined to specify additional patterns typically occurring in different sources of origin, including Web, Newspapers and Wikipedia.

2) Sentence Segmentizer: The implemented sentence segmentizer is a distributed version of a rule-based sentence segmentation tool that uses multiple of rule sets and lists of tokens to identify the sentence boundaries in a text. Disambiguation rules consider preceding and succeeding tokens of any potential sentence boundary symbol for their decision. The lists of tokens include *typical sentence boundary marks*, *abbreviations* for different languages, and a set of tokens which are not indicators of a sentence boundary, like file extensions (.com, .pdf) or top-level domains in URLs (.com, .de, .org).

3) Tokenizer: The implemented tokenizer identifies token boundaries using a set of rules and tokens. The tool relies on lists of typical token boundaries and punctuation characters, lists of abbreviations, and a list of known phrases and multi-word units that should be treated as single tokens. The implementation relies on sentences as input material. In cases where the input is not already separated into sentences, the aforementioned sentence segmentizer is called in advance.

4) Language Identification: The language identification tool utilizes sets of high-frequency words in different languages and their frequency to calculate the probability of a sentence belonging to a specific language.

5) Named Entity Recognition: The named entity recognition (NER) tool provides two main functions: the training of new NER models and the actual annotation task. New machine-learning-based models can be created and trained using fully annotated documents as input. The trained models are used to recognize instances of named entities in new documents. This functionality can be run on a cluster to annotate documents in parallel.

IV. Application

This section gives a short overview of research and infrastructure projects relevant to the context of the presented work. The sketched applications are already in use in some of them. Furthermore, a short overview of commercial providers of NLP services gives a broader picture of the current NLP-services landscape.

A. Leipzig Corpora Collection

The *Leipzig Corpora Collection* (including its subproject *Deutscher Wortschatz* focusing on the German language) uses a complex crawling infrastructure to continuously gather text material based on freely-available Web resources. This can include the acquisition of more than one terabyte of raw data per day (mostly based on HTML documents). The processing pipeline to convert this vast size of input documents to statistically and linguistically annotated text corpora contains a variety of tools with language and – sometimes – genre-specific configurations. Figure 3 gives a short summary of the used toolchain, that includes many of the aforementioned tools. For more details, cf. [26].

When starting the processing of Web material in the early 2000s, a single-threaded processing pipeline with mostly single-threaded tools was sufficient to convert raw material. Over time, the amount of gathered material, and thus requirements for the processing of this data has increased significantly.

As a consequence, most parts of the processing pipeline were parallelized and different approaches and system architectures were tested. Nevertheless, adequate handling of crawled material is still problematic and can lead to unprocessed heaps of data when incoming bandwidth exceeds the processing capability of subsequent processing and storage solutions.

B. Commercial and Academic Cloud-based NLP Frameworks

As mentioned before, the increase in the importance of NLP and NLU-based tools in both commercial and scientific contexts, has strengthen the demand for easy-to-use interfaces suitable for result-oriented users without deep knowledge of the underlying algorithms. As a consequence, a variety of platforms were created that target both user groups and focus on their specific demands.

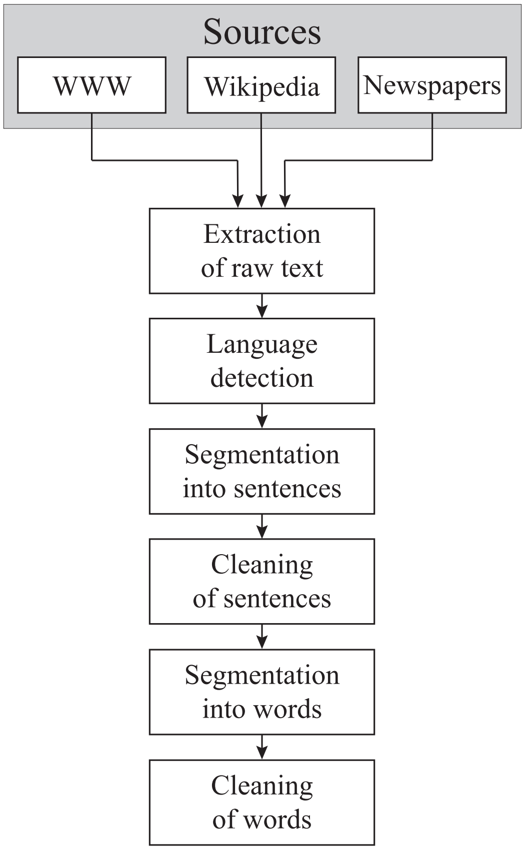


Figure 3. General overview of the processing pipeline of the Leipzig Corpora Collection.

Noteworthy examples of commercial platforms are *Amazon Comprehend* [27], *Google Cloud Natural Language Processing* [28], and IBM's *Watson Natural Language Understanding* [29]. Relevant academic or academic-oriented platforms include CLARIN's *WebLicht* [30](which will be discussed in more detail in the next subsection), *Language Applications (LAPPS) Grid* [31] or *GATE Cloud* [32].

Comparisons between these competitors may be based on a variety of evaluation criteria [33], like :

- Quality of documentation
- Openness of the platform
- Scalability
- Responsiveness
- Extent of supported tools
- Supported programming languages
- Supported natural languages (including specific dialects or language registers)
- Quality of results
- Costs

Evaluation of available platforms shows significant differences especially when comparing commercial systems with their academic counterparts. Typical advantages of the latter include support of more languages (including languages having lesser commercial interests), a broader landscape of provided tools, more options to participate as a service provider, and a lesser focus on financial gain. On the other hand, commercial platforms typically excel in most of the usability-related and technical aspects: high-quality documentation, easy-to-use (Web) interfaces, support of a variety of programming interfaces and, furthermore, often a built-in strategy for scalability aspects that allows the processing of large and very large data sets inside their system. This is especially obvious for companies like Amazon and Google being global providers of cloud computing services (*Amazon Web Services* AWS, *Google Cloud*).

In the context of this paper, the aspect of scalability obviously stands out as one the most relevant issues and is seen by the authors as one of the reasons that hinder the wide application and general success of open platforms in the field. This is especially problematic considering the sub-optimal availability of unique tools for a variety of languages which do not have a commercial focus because of rather small speaker groups or low commercial relevance. This can be seen as problematic considering today's diverse and broad cultural landscape. Closing the gap regarding scalability aspects can be helpful to reduce the identified divergence.

C. CLARIN WebLicht

A concrete example of an academic-based annotation framework is the WebLicht platform [12] of the CLARIN project [3]. WebLicht is an execution environment for automatic annotation of text corpora and provides currently more than 400 services in a federated and service-oriented environment. The number of supported tools and languages exceeds those of commercial alternatives and its restriction for scientific applications makes its use free of charge for the targeted user group. Its openness towards new service providers makes it a suitable candidate to evaluate potential benefits of improved scalability in services endpoints of such a federated infrastructure.

A WebLicht-compatible endpoint was created that uses the sketched processing architecture as a back-end (for more details, cf. [34]). Figure 4 shows the structure of this endpoint in the WebLicht environment. User input is handled by WebLicht's Web interface where input documents are uploaded and pipelines are built and started by the user. The second layer contains the implemented tools and shows an example of a user-defined workflow including the reading, annotating, and returning of processed documents. The format conversion is also performed in this layer if required. All actual processing is based on the execution framework providing a file system, resource manager, and execution engine as depicted above.

V. PARAMETERS

In any services-based environment, service catalogs are defined and developed based on demand, experiences, and best practices. These service catalogs contain – among many other aspects – comprehensive lists of requirements that are needed to provide services to the customer. As part of a service catalog management, detailed information about the performance of each item of the service catalog is required. It is an essential part of this task to include and consider all relevant parameters when modifying and improving this data continuously and

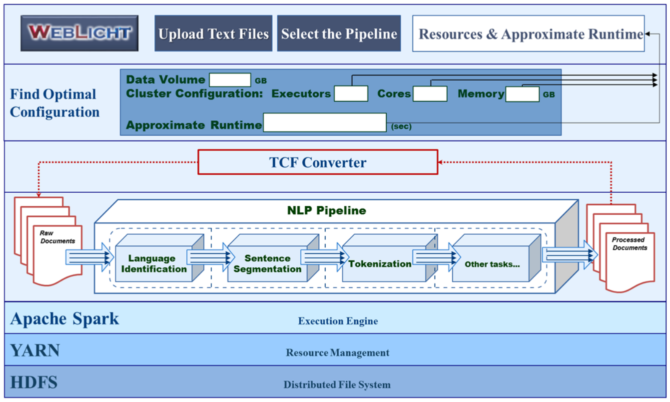


Figure 4. Service-oriented Architecture in WebLicht with Cluster Computation Technology as Backend.

to create meaningful "performance catalogs" for all provided services.

In this contribution, we considered several parameters relevant for performance evaluation and carried out several benchmarks to measure the effect of these parameters on the execution time of the tools. These parameters are categorized into two main categories: *cluster configuration* and *tool configuration*.

1) Cluster Configuration Parameters: These parameters consist of the following hardware resources:

- Executor: The experiments are performed using 8 or 16 worker nodes as executors. On each executor, 5 cores are available allowing 5 concurrent threads on each executor node. The number of cores multiplied by the number of executors defines the degree of parallelization or the maximum number of parallel tasks.
- RAM: Each machine in the cluster is equipped with 128 GB RAM; the allocated memory per executor is set to 8, 16, and 32 GB for most experiments. In general, 1 GB of memory is reserved for Hadoop and its related applications and 7% of the memory is considered for overhead. The rest will be used to process the data.

2) Tool Configuration Parameters: These parameters are selected based on the input documents and the specific workflow and include the following:

- Source of data: The input documents for these experiments are taken from the Leipzig Corpora Collection [26] that are collected from different sources comprising of Wikipedia, Newspapers, and general Web documents.
- Language: The documents are in English or German.
- Data volume: The input document collections are in sizes of 1 to 7 GB.
- Workflow: Workflows in this contribution are based either on the sentence segmentation or tokenization tool, or their combination into a joint workflow.

VI. Experiments

Various experiments were performed to analyze the effects of different cluster and tool configuration parameters on the execution times of the tools. The cluster is configured individually for each experiment to show different degrees of parallelism for all of the tools and origin of documents. Each experiment was repeated three times and the final execution time was measured as the average of these three run-times.

The following depicts the outcomes in diagrams which will be discussed in Section VII.

- Figures 5 to 8 present the run-times for the sentence segmentation task for different data volumes, text types, and cluster configurations.
- Figures 9 to 12 present the run-times for the word tokenization task for different data volumes, text types, and cluster configurations.
- Figures 13 to 16 present the run-times for the combined sentence segmentation and word tokenization task as a joint workflow for different data volumes, text types, and cluster configurations.

VII. Results

As Figures 1 and 2 illustrate, run-times vary for different job configurations significantly. As expected, using only a single executor – therefore, executing the job without any parallelization on the cluster – results in the maximum run-time for every data volume. The outcomes of all tests comply with the expected behaviour of parallel processing: a sharp decrease in execution time by increasing the assigned resources (i.e., executors), followed by a smoother reduction and finally no significant improvement when adding more resources to the job does not improve run-times anymore. The results show consistent behaviour for different data volumes using various cluster configurations.

The statistics also depict other relations: figures 5 to 16 also show that the execution times differ for various tool configuration parameters. Furthermore, the outcomes illustrate that in general, English documents need less time to be processed using these tools. Regarding the source of the documents, newspaper texts are segmented faster than other text sources in both English and German language, where this parameter does not affect the run-times of tokenization tasks significantly.

Another area worth analyzing is the negative impact of overheads in the processing. Figures 17 and 18 depict the execution times of sentence segmentation and tokenizing of 7 GB documents in two scenarios: applying the tools separately or performing the workflow as a combined task. Creating a joint workflow of both tasks reduces some of the redundant steps, as well as some I/O tasks and results in an improved execution time. For instance, processing 7 GB of German newspaper documents using 16 executors and 32 GB memory required 665 seconds when each tool applied separately. This decreases to 552 seconds for a combined toolchain. For English newspaper material and an identical cluster configuration, this leads to a decrease from 492 to 376 seconds.

Figure 19 gives an overview of run-times for data sets from 1 to 10 GB using 1 to 32 executors and 16 GB RAM per executor. It represents a significant reduction in execution times by providing more executors that is followed by a steady state.

Figure 20 shows the results for the sentence segmentation task of 10 GB text material which required 2860 seconds using

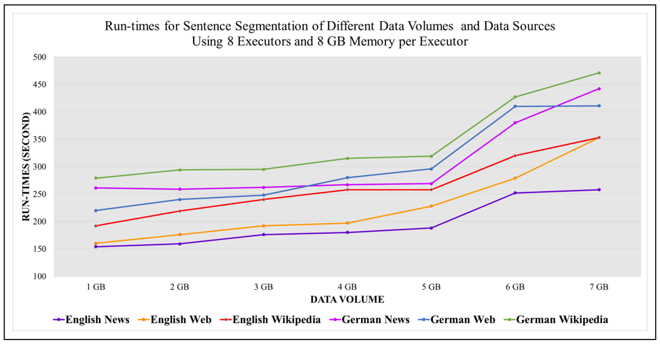


Figure 5. Run-times for segmenting 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 8 executors and 8 GB memory per executor.

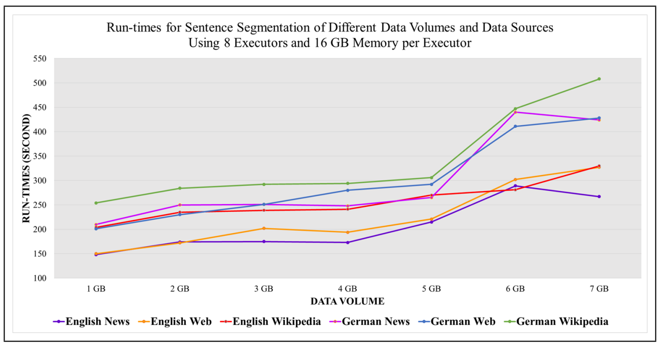


Figure 6. Run-times for segmenting 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 8 executors and 16 GB memory per executor.

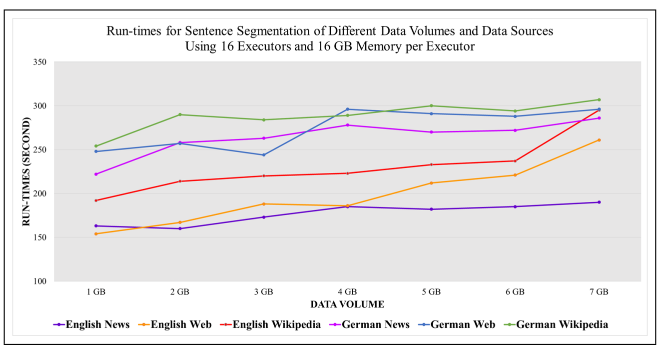


Figure 7. Run-times for segmenting 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 16 executors and 16 GB memory per executor.

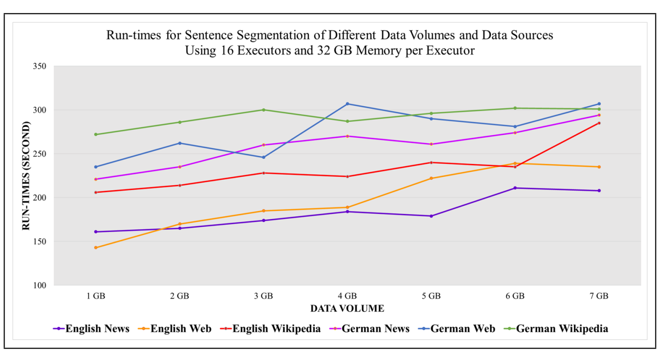


Figure 8. Run-times for segmenting 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 16 executors and 32 GB memory per executor.

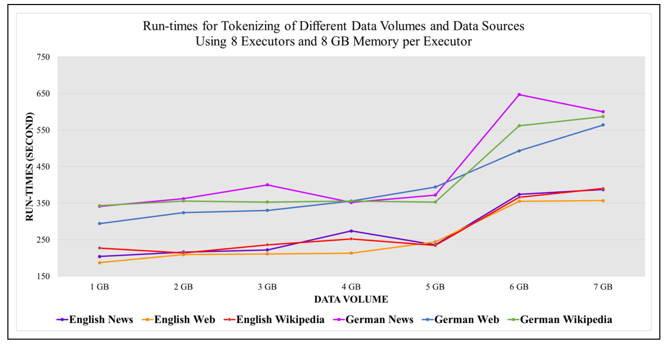


Figure 9. Run-times for tokenizing 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 8 executors and 8 GB memory per executor.

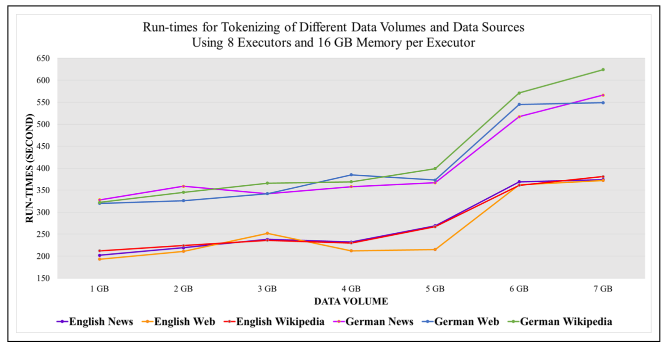


Figure 10. Run-times for tokenizing 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 8 executors and 16 GB memory per executor.

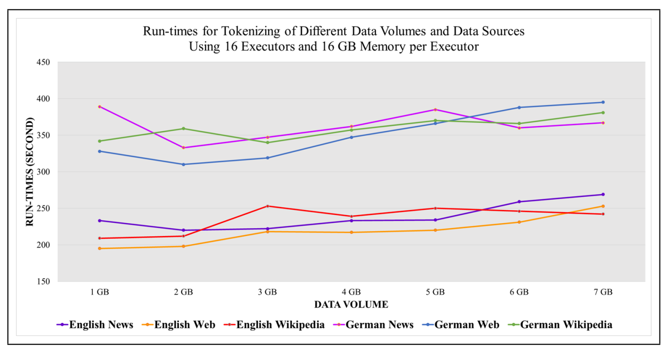


Figure 11. Run-times for tokenizing 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 16 executors and 16 GB memory per executor.

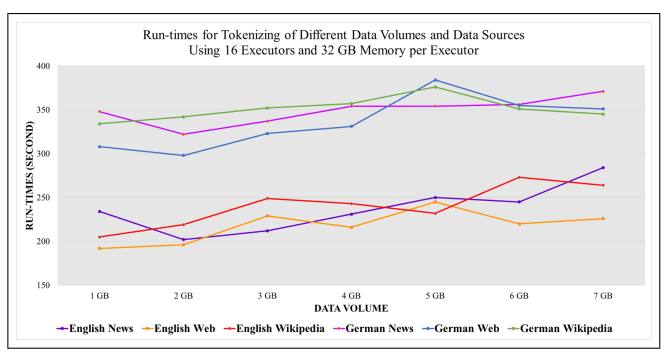


Figure 12. Run-times for tokenizing 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 16 executors and 32 GB memory per executor.

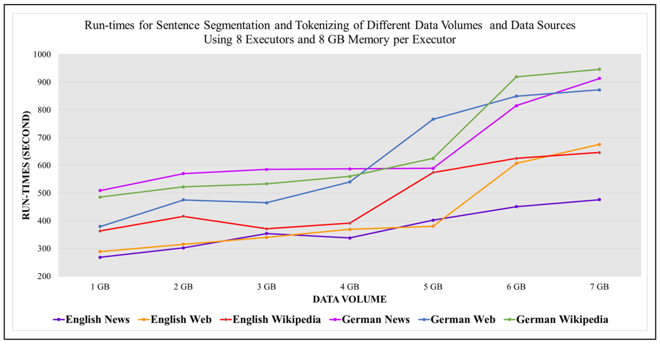


Figure 13. Run-times for segmenting and tokenizing 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 8 executors and 8 GB memory per executor.

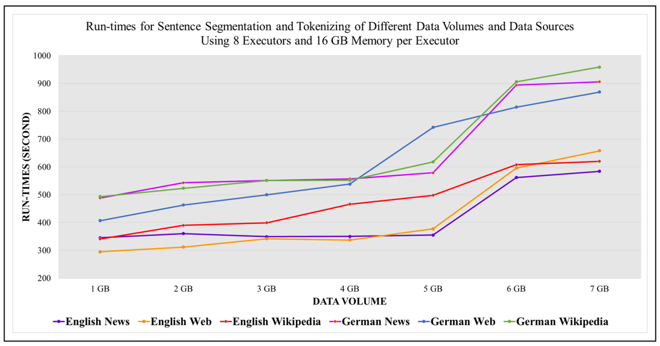


Figure 14. Run-times for segmenting and tokenizing 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 6 executors and 16 GB memory per executor.

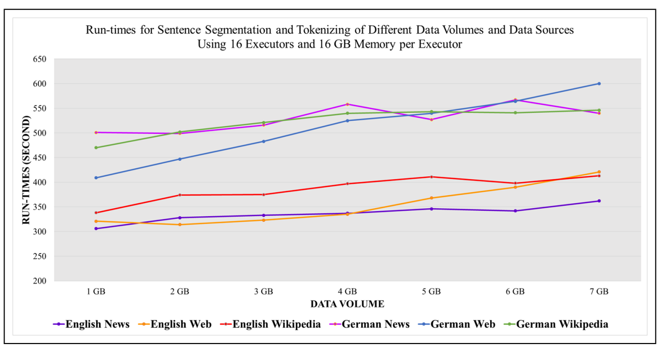


Figure 15. Run-times for segmenting and tokenizing 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 16 executors and 16 GB memory per executor.

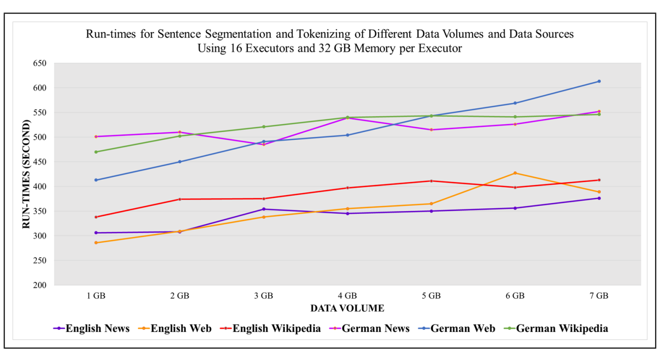


Figure 16. Run-times for segmenting and tokenizing 1 to 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 16 executors and 32 GB memory per executor.

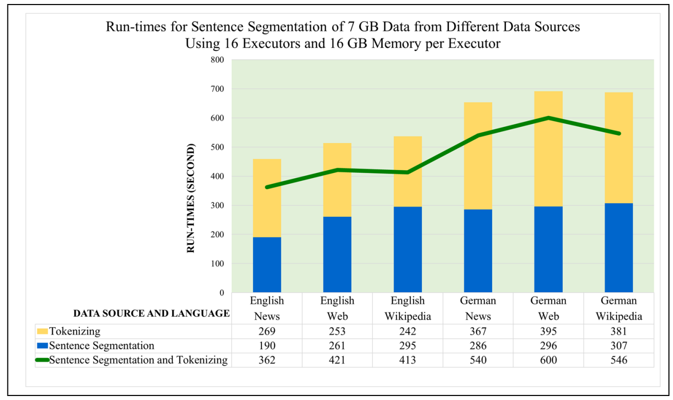


Figure 17. Run-times for applying segmentation, tokenization, and segmentation and tokenization combined as a toolchain on 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 16 executors and 16 GB memory per executor.

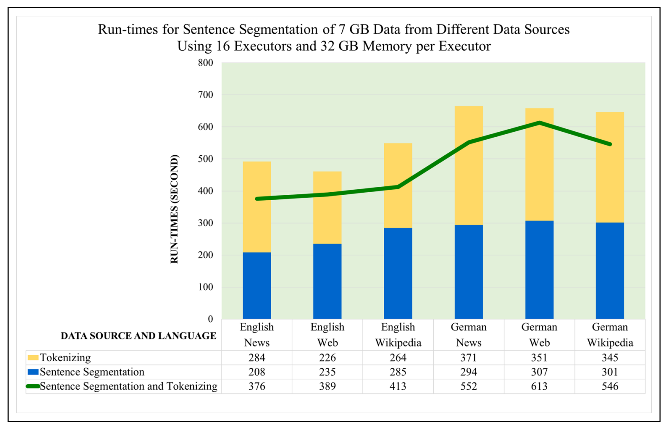


Figure 18. Run-times for applying segmentation, tokenization, and segmentation and tokenization combined as a toolchain on 7 GB text materials from Newspapers, Wikipedia, and General Web documents in English and German, using 16 executors and 32 GB memory per executor.

8 GB RAM and 2795 seconds using 16 GB RAM on a single node. Adding a second executor decreases the run-time to 2115 respectively 1480 seconds.

The typical trend can be seen again where run-times decrease significantly up to (around) 7 assigned executors, and with no improvements when allocating 14 executors or more.

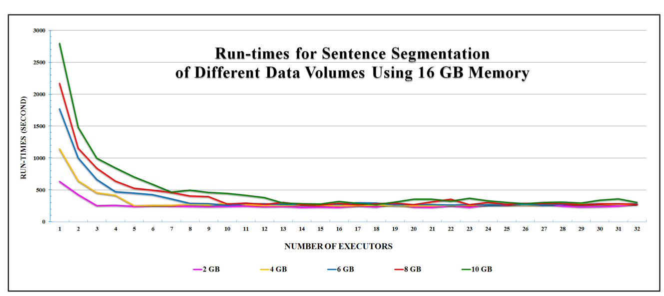
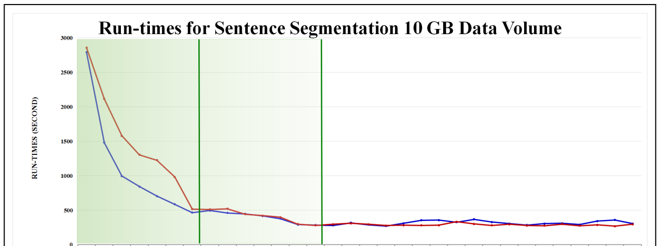


Figure 19. Run-times for different number of executors and data volumes using 16 GB memory per executor for sentence segmentation.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
16 GB Memory	2795	1480	997	843	704	587	467	496	461	447	417	378	293	286	280	318	291	271	310	355	357	327	368	330	305	285	306	310	292	342	360	306
8 GB Memory	2860	2115	1581	1304	1226	983	517	513	521	443	421	398	300	281	297	310	297	279	283	281	283	333	301	280	296	278	277	299	275	288	269	297

Figure 20. Run-times for different numbers of executors, illustrating different ”speedup areas”.

VIII. Discussion

Execution times are valuable information that can be utilized for the estimation of times of arrival for annotation jobs in NLP toolchains. Measured execution times give the opportunity to configure a cluster dynamically based on expected response times, available resources and the current load by a varying number of parallel users or jobs. For instance, if there are x free resources available on the cluster and a processing job requires $x+y$ resources, the new job may be scheduled to be executed after finishing the first running job which, has allocated at least y resources.

Furthermore, execution times are relevant for estimating an ”optimal” resource allocation for every individual tool. In the context of this contribution, these resources include the number of executors and the amount of memory which, can be assigned to each task. Obviously, the term ”optimal” is a very ambiguous one: it depends on the context of which value should be actually optimized. In this context, this may be the overall run-time of a job (i.e., a user-oriented view), the amount of allocated resources (i.e., a cost-oriented view) or a combination of both (by finding some balance between both).

By allocating more executors, execution times can be decreased. At a certain point (which may depend on a variety of parameters), assigning more resources will have no positive effect on execution times anymore. This point can be seen as the optimal configuration for the particular task in respect of optimized run-times, and contains the amount of resources which, are required to generate a result in the shortest possible execution time. In this situation, it is also feasible to generate results by assigning fewer resources – with the drawback of extended processing times – but it is obviously not reasonable to assign more resources to the job. As an example, in Figure 20 the fastest configuration for sentence segmentation of 10 GB text data consists of 14 executors with 16 GB RAM per executor where assigning more resources generates more costs without providing faster execution.

The extracted information helps to provide different resource configurations in accordance with the available hardware resources and desired response times for the user's requested service and input material. For instance, if a user wants to segment 10 GB text material in less than 25 minutes, 3 executors with 16 GB RAM or 4 executors with 8GB RAM would be both suitable configurations. In contrast, for a response time of up to 5 minutes, a configuration consisting of at least 14 executors with 8 or 16 GB RAM would suffice. In an environment where accounting of actual expenses is included, the balance between technical or financial costs and acceptable run-times can also be delegated to the user. In such an environment, a user can choose the desired configuration considering estimated run-times and incurred expenses.

The presented diagrams also show that for particular configuration changes resulting improvements of run-time are only marginal. Especially in case of limited available resources or unexpected usage peaks, these configurations do not have to be available anymore as their effect from the user's perspective are small. For instance, in Figure 20 assigning 7 executors with 16 GB RAM generates the expected result in 467 seconds whereas doubling the number of executors leads only to an execution time of 286 seconds (i.e., a 39% run-time reduction).

The diagrams also illustrate the impact of the language and source of the documents on the response time. In general, for the considered tools and inputs, German texts required more processing time than same-sized English texts. As a more specific outcome, segmenting newspaper texts took less time for both English and German documents compared with other sources of origin.

The reasons are related to different aspects. For example, texts that are published in newspapers typically follow high standards of writing rules using the correct punctuation marks which, are the base for segmenting the data, whereas general web documents and Wikipedia text are less likely to use adequate punctuation marks. Furthermore, the applied rule-based approaches use different numbers of more or less complex rules for different languages and the amount of considered external resources (like the number of multi-word units for a tokenization task) also varies between languages. Language-specific invariants (like the average length of sentences, average number of words per sentence etc.) must also be considered [35].

In addition, the results illustrate that the local combination of tools can decrease execution times significantly. This reduction is achieved by eliminating redundant steps such as reading common rule sets and abbreviation files and, more importantly, removing read/write operations between each of the tools. In our experiments, combining the sentence segmentation and tokenization tasks can decrease the whole execution times by about 25% (Figures 17 and 18). Obviously, this effect is much larger in a distributed environment where format conversion and network transmission times have to be taken into account as well.

IX. Conclusion and Outlook

In this contribution, we described some prerequisites for providing JIT delivery in service-oriented research infrastructures using typical NLP tasks as an example. We have utilized Apache Spark as execution engine on an Apache Hadoop cluster to allow parallel processing of large text collections and to increase the reliability and predictability of the services. An evaluation of required resources for processing different amounts of text offers information about possible hardware configurations in form of a ”performance catalog” that is useful for estimating delivery times and potential expenses for each task.

For more meaningful results, the sketched experiments have to be extended for more languages. In the context of this article German and English documents were used as an example but processing documents in other languages will provide more clear information about the effect of specific languages (or dialects) on resulting response times.

Naturally, providing and maintaining such resources and tools lead to actual financial costs. In commercial platforms, like Amazon Comprehend [27] or Google Cloud NLP [28] these costs are covered by contracts with costumers based on defined parameters (kind of service, required availability, costs of data storage, CPU cycles, etc.). The selected configuration and execution time can be used as a basis for an accounting system which, relies on well-founded expenses for every individual NLP job.

The presented run-times in this abstract can only be a part of a qualified assessment of NLP tasks. Performance profiles require a variety of training cycles to be meaningful and to cover all kinds of input material and their effects on the assessed tool. Furthermore, measuring actual response times for larger toolchains in text-oriented research infrastructures is more complex and needs to take more parameters into account. This is especially relevant for toolchains where multiple service providers are used. Other relevant parameters, like data transfer times between user and service provider or between different services, required format conversions, or similar tasks were not considered here.

Acknowledgement

The evaluation for this paper was performed using resources of Leipzig University Computing Center.

References

[1] S. Sahami and T. Eckart, "Just-In-Time Delivery for NLP Services in a Web-Service-Based IT Infrastructure," in Workshop NLP-adapt: Adaptive Workflow Management for NLP Data Processing at ADAPTIVE 2019: The Eleventh International Conference on Adaptive and Self-Adaptive Systems and Applications, Venice, Italy, 2019, pp. 103–106. [Online]. Available: https://www.thinkmind.org/index.php?view=article&articleid=adaptive_2019_7_20_58002

[2] F. W. McFarlane, Information technology changes the way you compete. Harvard Business Review, Reprint Service, 1984.

[3] E. Hinrichs and S. Krauwer, "The CLARIN Research Infrastructure: Resources and Tools for e-Humanities Scholars," Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014), May 2014, pp. 1525–1531. [Online]. Available: http://dspace.library.uu.nl/handle/1874/307981

[4] J. Edmond, F. Fischer, M. Mertens, and L. Romary, "The DARIAH ERIC: Redefining Research Infrastructure for the Arts and Humanities in the Digital Age," ERCIM News, no. 111, Oct. 2017. [Online]. Available: https://hal.inria.fr/hal-01588665

[5] EOSC, "EOSC European Open Science Cloud," Online, 2019, Date Accessed: 4 Dec 2019. URL https://www.eosc-portal.eu.

[6] C. Kuras, T. Eckart, U. Quasthoff, and D. Goldhahn, "Automation, management and improvement of text corpus production," in 6th Workshop on the Challenges in the Management of Large Corpora at the 11th Language Resources and Evaluation Conference (LREC 2018), Miyazaki (Japan), 2018.

[7] H. Wildemann, Das Just-in-time-Konzept: Produktion und Zulieferung auf Abruf, 5th ed., ser. TCW. Mnchen: TCW Transfer-Centrum fr Produktions-Logistik und Technologie-Management, 2001, vol. 4.

[8] P. Blais, "How the information revolution is shaping our communities," Planning Commissioners Journal, vol. 24, 1996, pp. 16–20.

[9] T. Ohno, Toyota Production System: Beyond Large-Scale Production. Taylor & Francis, 1988.

[10] P. Heck and A. Zaidman, "Quality criteria for just-in-time requirements: just enough, just-in-time?" in 2015 IEEE Workshop on Just-In-Time Requirements Engineering (JITRE). Los Alamitos, CA, USA: IEEE Computer Society, aug 2015, pp. 1–4. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/JITRE.2015.7330170

[11] IBM, "Web services architecture overview," Online, 2019, Date Accessed: 4 Dec 2019. URL https://www.ibm.com/developerworks/web/library/w-ovr/.

[12] E. W. Hinrichs, M. Hinrichs, and T. Zastrow, "WebLicht: Web-Based LRT Services for German," in Proceedings of the ACL 2010 System Demonstrations, 2010, pp. 25–29, Date Accessed: 4 Dec 2019. URL http://www.aclweb.org/anthology/P10-4005.

[13] "ISO - Popular Standards," Online, 2019, Date Accessed: 4 Dec 2019. URL https://www.iso.org/standard/71952.html.

[14] S. G. Manikandan and S. Ravi, "Big data analysis using Apache Hadoop," in 2014 International Conference on IT Convergence and Security (ICITCS). IEEE, 2014, pp. 1–4.

[15] S. Sagiroglu and D. Sinanc, "Big data: A review," in 2013 International Conference on Collaboration Technologies and Systems (CTS). IEEE, 2013, pp. 42–47.

[16] A. S. Tanenbaum and M. Van Steen, Distributed systems: principles and paradigms. Prentice-Hall, 2007.

[17] F. Wang, J. Qiu, J. Yang, B. Dong, X. Li, and Y. Li, "Hadoop high availability through metadata replication," in Proceedings of the first international workshop on Cloud data management. ACM, 2009, pp. 37–44.

[18] V. K. Vavilapalli, A. C. Murthy, C. Douglas, S. Agarwal, M. Konar, R. Evans, T. Graves, J. Lowe, H. Shah, S. Seth et al., "Apache Hadoop Yarn: Yet another resource negotiator," in Proceedings of the 4th annual Symposium on Cloud Computing. ACM, 2013, p. 5.

[19] ApacheHadoop, "Apache Hadoop Documentation," Online, 2019, Date Accessed: 4 Dec 2019. URL http://hadoop.apache.org.

[20] K. Shvachko, H. Kuang, S. Radia, R. Chansler et al., "The Hadoop distributed file system," in MSST, vol. 10, 2010, pp. 1–10.

[21] M. Zaharia, R. S. Xin, P. Wendell, T. Das, M. Armbrust, A. Dave, X. Meng, J. Rosen, S. Venkataraman, M. J. Franklin et al., "Apache Spark: a unified engine for big data processing," Communications of the ACM, vol. 59, no. 11, 2016, pp. 56–65.

[22] M. Hamstra and M. Zaharia, Learning Spark: lightning-fast big data analytics. O'Reilly & Associates, 2013.

[23] S. Sahami and T. Eckart, "Spark WebLicht Webservices," Online, 2019, Date Accessed: 4 Dec 2019. URL http://hdl.handle.net/11022/0000-0007-CA50-B.

[24] L.-P. Meyer, J. Frenzel, E. Peukert, R. Jäkel, and S. Kühne, "Big Data Services," in Service Engineering. Springer, 2018, pp. 63–77.

[25] L.-P. Meyer, "The Galaxy Cluster," Online, 2018, Date Accessed: 4 Dec 2019. URL https://www.scads.de/de/aktuelles/blog/264-big-data-cluster-in-shared-nothing-architecture-in-leipzig.

[26] D. Goldhahn, T. Eckart, and U. Quasthoff, "Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages," in Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12), 2012.

[27] Amazon, "Amazon Comprehend - Pricing," Online, 2019, Date Accessed: 4 Dec 2019. URL https://aws.amazon.com/comprehend/pricing/.

[28] Google, "Pricing - Natural Language API - Google Cloud," Online, 2019, Date Accessed: 4 Dec 2019. URL https://cloud.google.com/natural-language/pricing.

[29] IBM, "Watson natural language understanding," Online, 2019, Date Accessed: 4 Dec 2019. URL https://www.ibm.com/cloud/watson-natural-language-understanding.

[30] "WebLicht," Online, 2019, Date Accessed: 4 Dec 2019. URL https://weblicht.sfs.uni-tuebingen.de.

[31] "The language application grid— a web service platform for natural language processing development and research," Online, 2019, Date Accessed: 4 Dec 2019. URL https://www.lappsgrid.org/.

[32] "Gate cloud: Text analytics in the cloud," Online, 2019, Date Accessed: 4 Dec 2019. URL https://cloud.gate.ac.uk/.

[33] K. Thießen, A. Al-Ali, and J. Puchta, “Funktionsumfang und Leistungsfähigkeit kommerzieller und wissenschaftlicher NLP-Plattformen im Web,” Seminar work, 2019.

[34] S. Sahami, T. Eckart, and G. Heyer, “Using Apache Spark on Hadoop Clusters as Backend for WebLicht Processing Pipelines,” in CLARIN Annual Conference 2018 in Pisa, Italy, 2018.

[35] T. Eckart and U. Quasthoff, Statistical Corpus and Language Comparison on Comparable Corpora. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 151–165. [Online]. Available: https://doi.org/10.1007/978-3-642-20128-8_8