Article

Evaluating a Reference Model for SAV in Urban Areas

Antonio Reis Pereira [1,*], Pedro Portela [2], Marta Bicho [3,4] and Miguel Mira da Silva [1,5]

1   Instituto Superior Técnico, 1049-001 Lisboa, Portugal; mms@tecnico.ulisboa.pt
2   Católica Lisbon School of Business and Economics, Universidade Católica Portuguesa, 1649-023 Lisboa, Portugal; pedroportela99@gmail.com
3   Instituto Português de Administração de Marketing—IPAM Lisboa, 1500-210 Lisboa, Portugal; marta.bicho@ipam.pt
4   Business Research Unit (BRU-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), 1649-026 Lisboa, Portugal
5   INESC INOV, 1049-001 Lisboa, Portugal
*   Correspondence: reis.pereira@tecnico.ulisboa.pt

Abstract: Previous work presented a reference model for shared autonomous vehicles in urban areas supported by a systematic literature review and topic modeling. The proposed reference model was then evaluated with two real-world demonstrations: the service provided by Waymo in Phoenix and another offered by Baidu in Beijing. In this paper, we present another evaluation based on a survey conducted with a group of potential stakeholders belonging to the mobility industry who were asked about their agreement with each of the concepts in the reference model. The resulting artifact is stronger and more reliable because it reflects the feedback of mobility experts.

Keywords: shared; autonomous; vehicles; model; survey

1. Introduction

With robotaxis already on the roads in several Chinese and U.S. cities, their spread throughout other locations, namely European locations, is a matter of time, which is dependent on different factors such as customer adoption, political coverage, or technological availability, among others.

As highlighted by [1], the deployment of a shared autonomous vehicle (SAV) service in urban areas is synonymous with huge benefits, specifying that, for instance, in the Lisbon area, the mobility needs would be satisfied with just 10% of the current car parks if SAVs could replace existing passenger cars. Moreover, such replacement would mean a drop of 27% in CO_2 emissions, and all surface parking would be eliminated.

The forecasted effects in 2017 were so significant that it is not logical that SAVs are not yet ubiquitous because usually, big changes tend to be quickly adopted. The reason is the existence of obstacles or pre-conditions that do not exist yet in the deployment of SAVs.

Refs. [2,3] built a reference model with the main concepts sustaining an SAV service, demonstrating their reference model with the existing Phoenix ride hailing service operated by Waymo and Beijing ride hailing service provided by Baidu Apollo. Both real examples validated the twenty-nine concepts of the reference model, adding four more in Phoenix and another two more in Beijing, so altogether, the referred research work identified 35 concepts that should exist to have a ride hailing service in an urban area.

The reference model, demonstrated and fine-tuned with real examples, is solid because it is based on scientific research (papers) and processes (systematic literature review and topic modeling), however, even though the production of papers in this fast-growing research area as well as the two cited examples are not 100% coincident and are continuously changing, more validations are needed.

Ref. [4] stated that the demonstration of an artifact in more than one situation empowers the strength of that artifact. Ref. [5] suggested that the reference model has to be

verified in more than one way, and Ref. [6] discussed methodological field trials based on surveys for quality assurance.

In this case, specifically listening to possible stakeholders about each of the concepts that constitute the reference model of such a digital ecosystem is a step forward regarding model validation.

Therefore, a group of respondents was approached to answer a questionnaire built to validate the 35 concepts. The obtained sample comprised professionals acting directly or indirectly in the mobility industry, people who could theoretically participate in the reached digital ecosystem sustaining an SAV service in an urban area.

This new validation resulted in a much more reliable group of concepts, which were then modeled in ArchiMate to produce a revised reference model for SAVs in urban areas, which had the goal of being a decision-making tool to be adopted by SAV stakeholders in their strategical options, contributing to the quicker adoption of this new type of service.

The rest of this paper is as follows. Section 2 concerns the definition of the research background. In Section 3, we describe the research method based on the survey. Section 4 presents the results, Section 5 is the discussion, Section 6 is the updated reference model, and finally, our conclusions are presented in Section 7, which includes the main contributions, limitations, and future work.

2. Research Background

According to [7], a "survey" is often used to define a technique of obtaining data from a group of people in order to acquire information about the larger population from which the sample was obtained.

Continuing with [7], there are many types of surveys and a vast number of reasons to use surveys, although some features are common. A survey is not a census, because they obtain data from only a small number of individuals. In a bona fide survey, the sample is not designated randomly or just from individuals who offer to join; it is methodically selected in a way that everybody has the same possibility of being selected. Therefore, the outcomes can be consistently predicted for a bigger community. Data are obtained by using homogenous inquiries in order to have all surveyed respondents answer the same question.

The goal of a survey is not related to the description of an individual but instead to obtain the statistical shape of the population. The people who answer the survey are never recognized, and in the end, the outcomes are accessible as synopses like statistical boards or charts. The sample magnitude is based on the desired trustworthiness, so a sample size instruction is something that does not exist at all.

For [8], surveying populations is an appropriate research method, often used for non-experimental descriptive applications targeting the definition of the population. Therefore, for example, a survey research line can be utilized for the establishment of the frequency of a specific feature. Similarly, a survey approach is often utilized to obtain data regarding attitude and conduct. At times, there are themes that are better approached by a traditional trial, where the respondents are chosen from a randomized intervention group or a control group. The reality is that it is not easy to achieve an effective design. Of course, there are effective moral and real-world reasons showing why the respondents should not be randomly designated. Moreover, it is possible that a control group cannot be identified because controlling a randomization development is not an easy task.

Staying with [8], the validity of a survey might either be internal or external, well-organized, cover geographically wide samples, can present moral returns, and is flexible. Instead, surveys dependent of the sampling frame do not always present the reason why people behave, and surveys based on interviews depend heavily on the interviewers' effectiveness.

Still, according to these authors, regarding the collection of data for the survey, a survey is a sort of research design. Instead, interviews and questionnaires are techniques to collect information. There are plenty of ways to obtain people's data, where the three principal ones are: (a) direct meetings; (b) phone conferences; and (c) surveys. Choosing

the right approach varies in some aspects: contact with the respondents; learning level; theme; incentive; and existing resources.

To [8], questionnaires are not as expensive as individual interviews, and they are faster in vast and extensively distributed samples. Similar to phone meetings, a postal survey can be valuable if the respondents are dispersed. However, because there is no direct contact, prior questionnaire design and layout are critical.

Ref. [9] discussed the topic of questionnaires utilized in the sample. For the author, a questionnaire is a paradigm of consistent interrogations that operationalize the number of constructs. The objective is to show a homogenous incentive to respondents to facilitate comparable feedback from them. There is evidence proving that small variations in the inquiry wording or order are enough to considerably affect the answer, which highlights that questions should be precisely phrased, respecting the order, in order to gather analogous information. Continuing with the same author, people usually do not answer the predesign meaning of a query, but instead answer what they think it was meant to ask.

Nowadays, surveys might be web-based. Ref. [10] states that obtaining online information allows researchers to access people from a wide variety of locations in an inexpensive manner. Online surveys provide exceptional opportunities for obtaining information. These surveys are mostly suitable for collecting primary information as well as for pretesting research design and query understanding.

There are many software programs to consider when performing surveys, and Statistical Product and Service Solutions (SPSS) is a common option. SPSS is a manageable statistics software driver based on Windows, offering users a point-and-click means to produce output. This package allows innovative structures, which permits users to challenge advanced statistical procedures. Researchers may also try to use the syntax editor to produce 'code' with the goal of precise analysis, in contrast to the point-and-click scheme of producing output [11].

Taken together, given the available options, a web-based survey will be used to understand what experts think about the reached concepts, and the consequent statistical treatment will be performed by SPSS.

The ArchiMate Specification, a standard of The Open Group, is an open and autonomous modeling language for Enterprise Architecture based on diverse tool merchants and consulting firms. The ArchiMate Specification offers tools to empower Enterprise Architects to define, examine, and picture relations amongst business domains in an explicit manner [12].

3. Research Method

3.1. Sample

Sampling is to make a selection from a subset of persons that belong to a population to estimate the features of the entire population. With this method, the costs are lower and the information gathering process is faster [13].

In brief, the idea is to create a relevant stakeholder sample in terms of quantity and quality (real decision makers). This list should involve what is thought to be, at this stage, representatives of reference model elements updated with Phoenix and Beijing, where the decision makers would then be approached through a structured query with questions for a final check of the improved concepts list.

Therefore, an online survey for collecting primary data was conducted to validate the hypothesis created. The platform used for this purpose was Qualtrics, a web-based survey platform that provides tools for building and distributing surveys, analyzing responses, and generating reports, which also offers a user-friendly interface and a wide range of features for creating and conducting surveys [14].

The objective was to verify whether the twenty-nine reference model concepts added to the four new concepts found in Phoenix plus the two others identified in Beijing were confirmed by potential SAV stakeholders, who were then asked not only some general
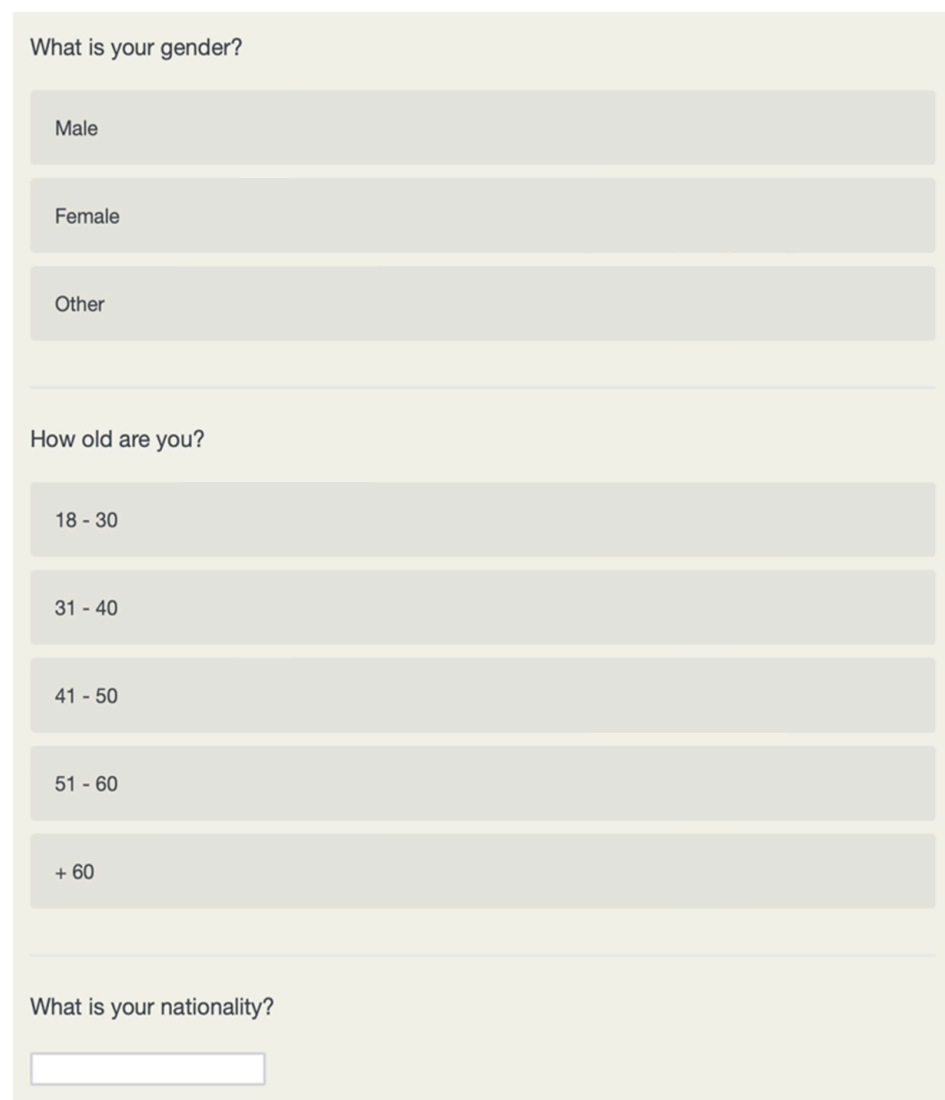
information about them (to know if they fulfilled the conditions to take part in the desired sample), but also about the relevance of each of the thirty-five identified concepts.

The population under analysis were top-tier decision makers (C-level executives, managers, directors, etc.) from various industries, ages, gender, and nationalities. The survey was sent to personal and non-personal contacts via social media (LinkedIn), direct e-mails, and WhatsApp to ensure that the targeted population was effectively reached (i.e., the sample was not randomly selected but instead the respondents were chosen based on defined criteria: people with decision-making capacity, with origin in industries that potentially support SAVs, from several countries, various ages, men and women).

Participants were asked to express their level of agreement with each statement by using a rating scale from 1 (strongly disagree) to 7 (strongly agree). Through the survey, the goal was to answer 35 hypotheses, each regarding a concept, namely whether those concepts are considered relevant for the implementation of SAVs in urban areas.

3.2. Online Questionnaire

The survey outlook in Qualtrix was as follows (Figures 1 and 2).

Figure 1. Survey—Respondents' data.

Please rate your level of agreement with the following concepts, on a scale from 1 (strongly disagree) to 7 (strongly agree).

When implementing a Shared Autonomous Vehicle Service in an Urban Area, I think that this concept is relevant

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
Automakers	O	O	O	O	O	O	O
Car	O	O	O	O	O	O	O
Autonomy	O	O	O	O	O	O	O
Vehicle detection and positioning hardware	O	O	O	O	O	O	O
Car-pool	O	O	O	O	O	O	O
Carsharing	O	O	O	O	O	O	O
Ridesharing	O	O	O	O	O	O	O
Client	O	O	O	O	O	O	O
IT Support Infrastructure	O	O	O	O	O	O	O
V2X (Vehicle-to-Everything)	O	O	O	O	O	O	O
Data	O	O	O	O	O	O	O
e-Mobility	O	O	O	O	O	O	O
Fare	O	O	O	O	O	O	O
Firms	O	O	O	O	O	O	O
Fleet	O	O	O	O	O	O	O
Fraud	O	O	O	O	O	O	O
Physical Infrastructure	O	O	O	O	O	O	O
Insurance	O	O	O	O	O	O	O
Metropolitan	O	O	O	O	O	O	O
Norms	O	O	O	O	O	O	O
Researcher	O	O	O	O	O	O	O
Privacy	O	O	O	O	O	O	O
Vehicle protocols	O	O	O	O	O	O	O
Security	O	O	O	O	O	O	O
Telecom	O	O	O	O	O	O	O
Traffic	O	O	O	O	O	O	O
Mobility	O	O	O	O	O	O	O
Vehicle AI platform	O	O	O	O	O	O	O
WI-FI	O	O	O	O	O	O	O
Client interface	O	O	O	O	O	O	O
Customer's feedback	O	O	O	O	O	O	O
Mapping	O	O	O	O	O	O	O
Customer perks	O	O	O	O	O	O	O
Remote Driving Service	O	O	O	O	O	O	O
QR Code with sanitary test	O	O	O	O	O	O	O

Figure 2. Survey—Concept validation.

3.3. Respondents' Data

The survey was answered by 122 respondents, but 34 did not finish the study, leaving 88 valid responses (participants were forced to respond to all items to avoid losing observations). Hundreds of people were approached, but the difficulty of obtaining responses was immense, especially because the survey was sent to top-tier managers.

Among the valid participants, 87.5% were male and 86.3% had an age gap between 41 and 60 years. Regarding the nationality of the participants, 68.2% were from Portugal, and a total of 12 nationalities (USA, Brazil, UK, France, Germany, Italy, Mexico, Poland, Portugal, Slovakia, Spain and Turkey) were included in this study. Of the 88 valid participants, 87.5%

were in the top three levels of decision making on a scale from 1 (lowest) to 7 (highest). For this study, to ensure that the respondents were undoubtedly people with a strong decision-making capacity, only the ones (77) who confirmed having a level of decision making equal to 5, 6, or 7 (on a Likert Scale of 1 (lowest) to 7 (highest)) were considered (Figure 3).
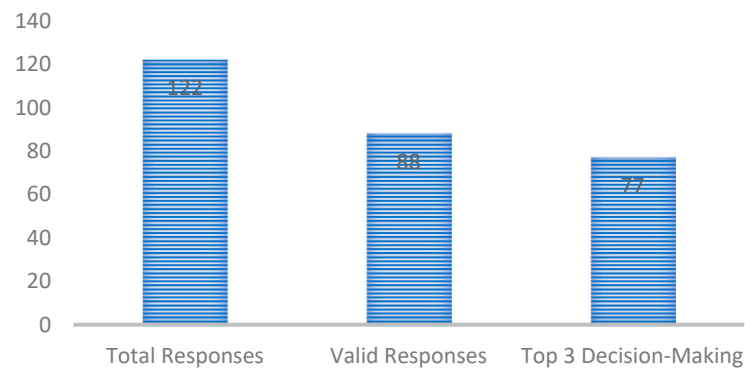
Figure 3. Survey respondents.

In other words, requests were only sent to people with recognized decision-making capacity, and on top of that, only the answers from respondents who were recognized to have that decision-making capacity were considered. With this extra measure, the obtained sample was more likely to be the targeted sample, avoiding, for instance, a person who can apparently make decisions but in fact cannot having answered the survey, thus reducing the possibility of delegation to a person without decision-making capacity.

Table 1 represents the frequency of the level of decision making of the respondents.

Table 1. Survey respondents' decision-making level.

	Frequency	Frequency (%)
6	36	40.9
5	24	27.3
7	17	19.3
Total	77	

Respondents were also questioned about the industry in which they worked. From the 77 responses considered for this study, 33.8% worked in the automotive industry, 33.8% worked in the banking/finance services industry, and 15.6% worked in the mobility industry. Approaching people from the banking/financial services industry does make sense because it is often through captive finance companies that automakers can promote new mobility concepts. The respondents in insurance, law, consulting, and research were also people specializing specifically in the automotive/mobility industry.

In general, the people approached were decision makers in areas forecasted to be around SAVs. Overall, the idea behind the present survey was to approach potential SAV stakeholders, and the sample attained that goal.

Table 2 shows the industry frequency where the respondents worked.

In addition to their level of decision making and industry, the participants were asked about their job position, which resulted in several types of answers (Table 3).

Table 2. Survey of the respondents' industries.

Industry	Frequency	Frequency (%)
Automotive	26	33.8%
Banking/Finance	26	33.8%
Mobility	12	15.6%
Insurance	5	6.5%
Law	3	3.9%
Research	3	3.9%
Political	1	1.3%
Consulting	1	1.3%
Total	77	100.0%

Table 3. Survey of the respondents' positions.

Position	Frequency	Frequency (%)
C-Level	19	24.7%
Director	13	16.9%
Managing Director	10	13.0%
Marketing Director	4	5.2%
Manager	3	3.9%
Partner	3	3.9%
Country Manager	2	2.6%
Sales Director	2	2.6%
Sales Manager	2	2.6%
ARAC	1	1.3%
Branch Manager	1	1.3%
Branch Manager FO	1	1.3%
CDO	1	1.3%
Digital Transformation/New Business Development Manager	1	1.3%
Equity Partner	1	1.3%
Fleet Sales Director	1	1.3%
FO	1	1.3%
Head of Fleet	1	1.3%
Head of Legal and Compliance	1	1.3%
Mayor	1	1.3%
New Mobility and Connected Services	1	1.3%
PhD Candidate	1	1.3%
Regional Manager	1	1.3%
Secretary-General	1	1.3%
Senior Head of Department	1	1.3%
Senior Manager	1	1.3%
Unit Head	1	1.3%
Vice-President	1	1.3%
Total	77	100.0%

For the purpose of this study, the positions were aggregated into three categories: C-level (top management), directors, and middle managers.

C-level executives, also known as C-suite executives, are high-ranking executives within an organization who typically hold titles beginning with the letter "C" such as CEO, COO, CFO, CMO, or CIO.

These individuals are part of the top leadership team and are responsible for making strategic decisions that impact the organization as a whole [15].

Directors are individuals who are appointed or elected to serve on the board. The board of directors oversees the company's administration and planned strategy. Directors have fiduciary responsibilities to make decisions respecting the best interest of the company and the respective owners. They are involved in making high-level decisions, setting company policies, and providing guidance to the executive management team [15].

A middle manager is an individual within an organization who holds a management position and is typically responsible for overseeing a team or department within the company [15].

From the selected 77 responses considered for this study, 51.9% were C-level executives, 44.2% were directors, and 3.9% were middle managers. Table 4 shows the frequency of the positions of the respondents.

Table 4. Aggregated survey of the respondents' positions.

Position	Frequency	Frequency %
C-Level	40	51.9%
Director	34	44.2%
Middle Manager	3	3.9%
Grand Total	77	100%

In general, the reached sample included relevant people for the implementation of SAVs like a Rent-a-Car association leader, the mayor of a European city capital (that dominates an urban area), and the global chief digital officer of an automaker captive finance company.

Regarding gender, taking into consideration that women CEOs run 10.4% of Fortune 500 companies, and taking that as a population feature, the 12.5% of female respondents, although far from being a positive figure, in practice, this is in line with the population. Additionally, according to the consulting firm Deloitte, the female gender represents only 20% of the automotive workforce, decreasing to an amount that does not reach 10% at the executive level, which means that a sample with 12.5% of females is aligned with what exists in the top-management executive market concerning gender.

Regarding the sample age, taking figures from consulting firm Korn Ferry, in 2019, the U.S. market had an average age of 56 years old for C-members compared with the obtained average sample of 49.3 years. The sample average was 6.3 years younger than the population, but considering that C-members will make SAV decisions in an uncertain future, the obtained sample will be older when those decisions are made, so it is likely that the age gap is not that material.

It is also fair to question and discuss whether the reached sample size is large enough to extrapolate statistical conclusions, especially because, among other factors, the sample size depends on the population dimension, in this case, the global decision-making executives from potential SAV stakeholders, which is a highly subjective number that is impossible to calculate.

The available literature is usually conservative about recommending a number for a sample size [14], instead, it has pointed out that if the type of statistics is descriptive, like mean or frequencies (like this present case), then nearly any sample size will suffice, suggesting that a minimum of 100 elements are required for each major group or subgroup in the sample, and for each minor subgroup, a sample of 20 to 50 elements is enough. Therefore, in this study, our reached final sample of 77 respondents is still within the defined author boundaries.

In this way, considering the reached sample dimension, quality (40 C-level people), diversity (12 countries, 8 industries), gender, and age, to the best of our knowledge, our sample was valid to proceed with the statistical analyses.

4. Results

To test the relevance of the 35 concepts, descriptive statistics and the one-sample T-test were employed as analytical tools to assess the significance of the respondents' agreement for the validation of the 35 identified concepts. Descriptive statistics was utilized to compute the mean and standard deviation, offering a concise summary of the central tendency and variability in the participants' responses.

The one-sample T-test was then applied to compare the obtained mean with a theoretical value, specifically 4 on a 1–7 Likert scale, which represents neutral agreement. This statistical approach enables an exploration of whether the companies, as indicated by the survey respondents, exhibited a noteworthy inclination toward the validation of the concepts or not. The use of the one-sample T-test facilitates an objective assessment of whether the observed means significantly deviate from neutrality, providing a robust foundation for conclusions.

To facilitate the analysis, we decided to categorize the hypotheses into four distinct groups: Strongly Supported, Supported, Partially Supported, and Not Supported at all, based on a threshold analysis derived from the mean scores of the respondents (Table 5).

Table 5. Level of support.

Category	Mean Range	Interpretation
Strongly Supported	>6	Concepts with a mean score higher than 6, indicating robust consensus
Supported	5.5–6	Concepts with a mean score between 5.5 and 6, signifying substantial agreement
Partially Supported	5–5.5	Concepts with a mean score between 5 and 5.5, showing agreement with some diversity in perspectives
Not Supported	0–5	Concepts with a mean score between 0 and 5, suggesting a lack of consensus or lower agreement

Table 6 shows a summary of the figures obtained through descriptive analysis, using the categorization explained above for an easier understanding.

Altogether, in front of available options, a web-based survey will be used to know.

Table 7 summarizes reached results.

Table 8 shows a summary of the one-sample T-test conducted for the questions.

The initial analysis of the descriptive statistics provided insights into the central tendency and variability of the responses for each concept. The subsequent one-sample T-test served as a statistical validation, supporting the initial analysis by confirming the level of support for each hypothesis. The combination of both analyses enhanced the robustness of the findings, offering a more comprehensive understanding of the respondents' attitudes and perceptions related to SAV services in urban areas. In this case, the one-sample T-test corroborated and reinforced the conclusions drawn from the descriptive statistics, providing a more robust basis for the support levels of each hypothesis.

Table 6. Descriptive analysis.

Variables	Concept	N	Minimum	Maximum	Mean	Std. Deviation	Result
6	Automakers	77	1	7	5.61	1.309	Supported
7	Autonomy	77	2	7	5.96	1.069	Supported
8	Car	77	2	7	5.66	1.231	Supported
9	Car-pool	77	3	7	5.79	1.092	Supported
10	Carsharing	77	1	7	5.44	1.419	Partially supported
11	Client	77	2	7	5.49	1.314	Partially supported
12	Client interface	77	4	7	6.06	0.848	Strongly supported
13	Customer perks	77	2	7	5.36	1.256	Partially supported
14	Customer's feedback	77	2	7	6.00	1.039	Strongly supported
15	Data	77	3	7	6.23	0.985	Strongly supported
16	e-Mobility	77	3	7	6.03	0.903	Strongly supported
17	Fare	77	1	7	5.66	1.083	Supported
18	Firms	77	2	7	5.30	1.171	Partially supported
19	Fleet	77	1	7	5.52	1.199	Supported
20	Fraud	77	1	7	5.30	1.514	Partially supported
21	Insurance	77	1	7	6.13	1.128	Strongly supported
22	IT Support Infrastructure	77	4	7	6.43	0.715	Strongly supported
23	Mapping	77	2	7	5.90	1.242	Supported
24	Metropolitan	77	2	7	5.47	1.252	Partially supported
25	Mobility	77	4	7	6.23	0.776	Strongly supported
26	Norms	77	4	7	5.81	1.026	Supported
27	Physical Infrastructure	77	2	7	5.83	1.056	Supported
28	Privacy	77	2	7	5.78	1.284	Supported
29	QR Code with sanitary test	77	1	7	4.60	1.575	Not supported
30	Remote Driving Service	77	1	7	5.25	1.549	Partially supported
31	Researcher	77	1	7	5.18	1.275	Partially supported
32	Ridesharing	77	2	7	5.14	1.364	Partially supported
33	Security	77	4	7	6.40	0.831	Strongly supported
34	Telecom	77	2	7	5.73	1.242	Supported
35	Traffic	77	2	7	5.78	1.177	Supported
36	V2X (Vehicle-to-Everything)	77	2	7	5.48	1.294	Partially supported
37	Vehicle AI platform	77	3	7	5.92	1.085	Supported
38	Vehicle detection hardware	77	1	7	6.38	0.946	Strongly supported
39	Vehicle protocols	77	2	7	5.87	1.043	Supported
40	WI-FI	77	1	7	5.68	1.352	Supported

Table 7. Descriptive results.

Results	Not Supported	Partially Supported	Supported	Strongly Supported	Total
N	1	10	15	9	35

Table 8. One-sample T-test.

Hypotheses	Concepts	t (Test Value = 0)	df	Significance		Mean Difference	95% Confidence Interval of the Difference	
				One-Sided *p*	Two-Sided *p*		Lower	Upper
H7	Automakers	37.597	76	<0.001	<0.001	5.610	5.31	5.91
H8	Autonomy	48.922	76	<0.001	<0.001	5.961	5.72	6.20
H9	Car	40.354	76	<0.001	<0.001	5.662	5.38	5.94
H10	Car-pool	46.532	76	<0.001	<0.001	5.792	5.54	6.04
H11	Carsharing	33.655	76	<0.001	<0.001	5.442	5.12	5.76
H12	Client	36.682	76	<0.001	<0.001	5.494	5.20	5.79
H13	Client interface	62.746	76	<0.001	<0.001	6.065	5.87	6.26
H14	Customer perks	37.484	76	<0.001	<0.001	5.364	5.08	5.65
H15	Customer's feedback	50.687	76	<0.001	<0.001	6.000	5.76	6.24
H16	Data	55.513	76	<0.001	<0.001	6.234	6.01	6.46
H17	e-Mobility	58.569	76	<0.001	<0.001	6.026	5.82	6.23
H18	Fare	45.858	76	<0.001	<0.001	5.662	5.42	5.91
H19	Firms	39.722	76	<0.001	<0.001	5.299	5.03	5.56
H20	Fleet	40.402	76	<0.001	<0.001	5.519	5.25	5.79
H21	Fraud	30.717	76	<0.001	<0.001	5.299	4.96	5.64
H22	Insurance	47.686	76	<0.001	<0.001	6.130	5.87	6.39
H23	IT Support Infrastructure	78.892	76	<0.001	<0.001	6.429	6.27	6.59
H24	Mapping	41.669	76	<0.001	<0.001	5.896	5.61	6.18
H25	Metropolitan	38.314	76	<0.001	<0.001	5.468	5.18	5.75
H26	Mobility	70.470	76	<0.001	<0.001	6.234	6.06	6.41
H27	Norms	49.626	76	<0.001	<0.001	5.805	5.57	6.04
H28	Physical Infrastructure	48.439	76	<0.001	<0.001	5.831	5.59	6.07
H29	Privacy	39.504	76	<0.001	<0.001	5.779	5.49	6.07
H30	QR Code with sanitary test	25.615	76	<0.001	<0.001	4.597	4.24	4.95
H31	Remote Driving Service	29.726	76	<0.001	<0.001	5.247	4.90	5.60
H32	Researcher	35.676	76	<0.001	<0.001	5.182	4.89	5.47
H33	Ridesharing	33.082	76	<0.001	<0.001	5.143	4.83	5.45
H34	Security	67.585	76	<0.001	<0.001	6.403	6.21	6.59
H35	Telecom	40.458	76	<0.001	<0.001	5.727	5.45	6.01
H36	Traffic	43.094	76	<0.001	<0.001	5.779	5.51	6.05
H37	V2X (Vehicle-to-Everything)	37.170	76	<0.001	<0.001	5.481	5.19	5.77
H38	Vehicle AI platform	47.878	76	<0.001	<0.001	5.922	5.68	6.17
H39	Vehicle detection Hardware	59.121	76	<0.001	<0.001	6.377	6.16	6.59
H40	Vehicle protocols	49.379	76	<0.001	<0.001	5.870	5.63	6.11
H41	WI-FI	36.840	76	<0.001	<0.001	5.675	5.37	5.98

The high t-values observed in the one-sample T-test signify a significant deviation of the sample means from the hypothesized population mean, providing strong statistical evidence in favor of the hypotheses. The very low one-sided *p*-values, often below the conventional significance level of 0.05, further indicate the support for the acceptance of the concepts. Additionally, the positive mean differences signify a consensus among respondents in favor of the concepts being tested. Collectively, the statistical indicators

contributed to the robustness of the findings, reaffirming the support levels derived from the descriptive statistics and underscoring the reliability of the conclusions drawn from the survey data.

In short, the strong support for these hypotheses indicates a consensus among the survey participants regarding the critical factors that contribute to the success of SAV services. The positive mean differences, along with the narrow confidence intervals and very low *p*-values, reinforce the robustness of these findings. This alignment in opinions and the high level of support across concepts suggest a unified perspective among the respondents on the essential components and considerations that need to be considered by companies, cities, and governments when implementing SAV services in urban areas.

The initial analysis of descriptive statistics revealed valuable insights into the respondents' attitudes and perceptions regarding SAV services in urban areas. This examination provided a clear understanding of the tendency and variability associated with each concept. Subsequently, the one-sample T-test served as statistical validation, affirming and reinforcing the initial analysis by confirming the level of support for each hypothesis. This dual-method approach not only deepened our comprehension of the respondents' perspectives, but also enhanced the robustness of the findings. Together, these analyses offer a comprehensive evaluation of the relevance of various concepts, providing a solid foundation for understanding the factors that are crucial to the successful implementation of SAVs in urban areas.

5. Discussion

The reached artifact is mostly supported by scientific articles, and the consequent group of concepts is the result of several processes (SLR, topic modeling). Nevertheless, to come from science does not signify that it is science again, so our objective was the validation of the reference model to check whether it fit with real cases and the opinions of those who could theoretically implement and participate in an SAV service.

The model, in demonstration and evaluation, is a predictive and not a descriptive model because it is mostly supported by papers that anticipate a reality. Therefore, its confrontation with reality is, in a certain way, a litmus test to verify whether it makes sense or not.

These confrontations confirmed concepts and brought about some changes. In fact, updating the model with the incorporation of concepts that came from real cases in Phoenix and Beijing strengthened the model.

Considering the lessons from Phoenix, it is possible to have a shared autonomous vehicle service based on car intelligence and precise mapping long before full connections are available, though it is not possible to estimate when this will happen. In fact, while V2X is very important, it is not currently a reality, so accurate mapping is a workaround that allows for the existence of SAVs. Additionally, having an interface with the client and mobility provider is also important, as well as customer perks, targeting people with special needs with serious mobility constraints, and boosting client engagement.

Regarding Beijing, aside from the mapping already added because of Phoenix, the two new concepts not raised by the reference model but identified in Beijing were related to the increase in customer adherence to SAVs. The remote driving capacity aims to boost the customers' confidence, offering the capability that if something goes wrong, the car will not stop on the street, but will instead be driven to a safe place. The QR code can also be associated with the customer experience, which is very important in a COVID-19 environment when customers are still reluctant to access transportation recently used by other customers.

Next, the twenty-nine reference model concepts, with the addition of the four extra concepts found in Phoenix and the two in Beijing were evaluated by a group of decision makers with origins in industries that could participate in the digital ecosystem constituted by the identified concepts.

The survey findings suggested that nine of the concepts were strongly validated, and fifteen others were also validated, making a total of 24 concepts that the respondents agreed were crucial for the implementation of a shared autonomous vehicle service in an urban area. The not-validated concept was "QR Code with a sanitary test", which makes sense because the COVID-19 times are no longer, reinforcing, in a certain way, the validity of this study. Therefore, this concept will now be definitively excluded from the final model.

The 10 concepts that were partially supported need deeper diving, requiring further reflection.

Regarding the concepts "carsharing" and "ridesharing", it may seem that the respondents were skeptical about giving them more relevance, but since people often have a "car ownership" mentality, we are still slowly moving toward a shared mobility state of mind, so the concept of sharing such a personal item may take some time accept.

Following the Tesla strategy, a carsharing peer-2-peer concept would be useful to convince AV car owners to share these cars, lowering or even recovering these investments. In a sense, this would allow car owners to continue to use the car whenever they need the car, and receiving fees in return from a ride hailing service whenever the car is not needed. This would also facilitate SAV promoter companies, here called "firms", to enter the business, because the investment would not be as large through this rent-2-rent concept. Of course, this will not be easy to implement as the sense of ownership prevails, and it is still quite important for many people. On the other hand, this could be an important step toward SAVs because economically, it does make sense for either car owners or SAV platforms, so it will probably be a concept that is not adopted in the beginning, but over time.

Ridesharing might be perceived as something that threatens people because sharing a ride with strange people in a confined place without witnesses can be dangerous. This is likely to be a service segment for low-cost customers, and eventually to young people, with the relative weight related to the price decrease that a shared trip can provide.

Regarding the concept of "client", it is strange that the respondents did not find it highly relevant, but this service would not be achievable if there were no clients. It is possible that what was in mind was some risk aversion that can freeze SAV adoption, which is something to address with other concepts like customer perks or remote driving capacity to improve the customers' confidence.

When it comes to the concept of "customer perks", it seems essential to recognize the importance of customer satisfaction and the role it plays in the success of shared autonomous vehicle services. The partial support can be related to the lack of importance given to perks in comparison with the other concepts, which makes sense. These perks might not only be small gifts to attract customers, for instance, when Uber offered bottled water to its customers, but could also be child seats or wheelchair facilities, enlarging the service to a wider segment of customers.

The concept of "firms" may have been subject to misinterpretation, as it may have not been clear that we referred to the companies that implement this service as it was intended, but rather how it could benefit firms as clients. This concept should have instead been called "mobility provider" to avoid this type of doubt. To have SAVs in an urban area, someone must move forward and launch the service, so the existence of mobility providers is mandatory.

The concept of "fraud" had positive agreement amongst the respondents, but a considerable diversity in perspectives. This is an important concept to keep, as enhancing fraud prevention strategies is important for confidence among SAV service users.

Regarding "metropolitan", this concept may be valid in a further stage of implementation, as the initial stage focused more on an urban area, which may be the reason for some of the respondents' different perspectives. It seems clear that once implemented in more urbanized locations, the SAV trend will spread throughout metropolitan areas. Right now, the stakeholders' doubts make sense, because SAVs will not operate in all metropolitan areas, at least in an early stage.

The "remote driving service" concept, similar to the concept "firms", may have been misinterpreted, as this is a futuristic technology of the car driving itself to the nearest help point in the event of a malfunction during the ride. Since this is not a service that exists currently, it may be subject to some holdbacks in its utility and practicality. However, it is a concept to consider in terms of service security.

When it comes to the concept of "researcher", the respondents may have had a more practical perception and did not interpret this asset as fully necessary, especially when in comparison with the other concepts. Collaborative efforts between researchers and decision makers can ensure research activities, and aligned with the practical needs and expectations of stakeholders, can foster a mutually beneficial relationship, which is precisely the intention of the present research.

Finally, as previously addressed, the concept of V2X was understandably questioned, as this is a technology from the future being assessed in the present. This concept suggests seamless interaction between the vehicle grid and the surrounding infrastructure such as traffic lights, public services, and other cars. Although it was only partially supported by the respondents, it seems like a valid concept to keep in the model.

All in all, evaluating the reference model with Phoenix and Beijing and surveying the consequent 35 concepts, only one concept should be excluded—the QR code with the sanatory test—so the remaining 34 concepts should constitute the updated reference model, which is the main value proposition of the present research.

6. Updated Reference Model

The final set of concepts is represented in Table 9.

Table 9. Updated concepts.

automakers
car
autonomy
vehicle detection and positioning hardware
car-pool
carsharing
ridesharing
client
IT support infrastructure
V2X
data
e-mobility
fare
firms
fleet
fraud
physical infrastructure
insurance
metropolitan
norms
researcher
privacy
vehicle protocols

Table 9. *Cont.*

security
telecom
traffic
mobility
vehicle AI platforms
client interface
customer's feedback
mapping
customer perks
remote driving capability

Repeating the process followed by [3], that is, classifying all 34 concepts into ArchiMate layers and sublayers of the ArchiMate 3.1 Specification and then using the Archi—Open Source ArchiMate Modeling4, it was possible to model the obtained concepts as well as the relationships among them, according to the ArchiMate defined relationships.

The results are represented in Figure 4. These were not very different from the reference model updated by Phoenix in [3], with the addition of the remote driving capacity that was identified in Beijing, and not including the QR code used in Beijing, clearly rejected by the survey because the COVID-19 times are no longer.
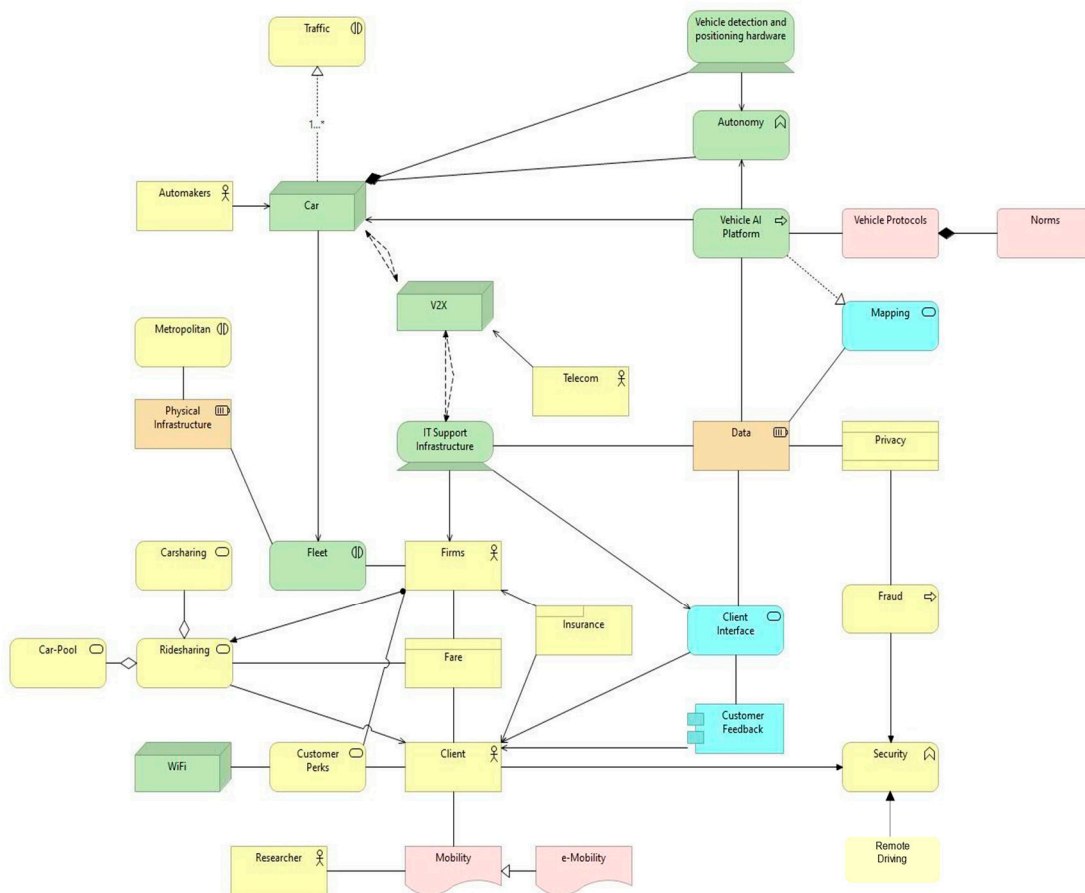
Figure 4. Updated reference model.

7. Conclusions

Overall, the main output of the referred match of the reference model, first with Phoenix and Beijing, and now with the survey, is encouraging.

In fact, the reference model is logical, and, to the best of our knowledge, its update with the realities in Phoenix and Beijing, now boosted by the present survey, represents a trustable instrument with which to make decisions in this mobility area.

Once the digital ecosystem behind SAV and its several concepts have been reached, it is crystal clear that SAVs will spread only if these concepts exist, and highlighting these conditions is itself a pre-condition for the decision-making process. The availability of an AV is not enough to have an SAV fleet operating in an urban area, and there is a list of assumptions that were here exhaustively mentioned that condition the existence of such a service, and those conditions now include the remote driving capability and do not definitively include the QR code to enter the car.

There are, however, some limitations related to this work.

An important constraint is that the reference model was built based on scientific works that were limited in time, which means that new papers published in the meantime can, in theory, transform the artifact here presented, which means repeating the work in this study with more real situations and consequently a newer survey.

Furthermore, the identified differences between Phoenix and Beijing prove that there is a need for local fine-tuning, that is, the reference model presented in this research will always need to consider the geography where the SAV service is deployed.

In novel research about multi-modal transportation networks, Ref. [16] precisely joined the adaptation to the innovation and to the collaboration, targeting the creation of a transportation ecosystem that is safer, more efficient, more accessible, and more sustainable for generations to come.

Finally, the survey only validated the concepts, and the respondents were not asked to add other concepts. The alternative would have been to risk a diversity of opinions that would have been impossible to incorporate into the model, which was already solid enough due to its scientific origin, either in papers or in processes.

As artifact mutability is related to evolution [4], that is, with a reliable and trustable foundation, this model is logically a never-ending creation, continuously updatable, which is also a lesson learned in the present research.

In this sense, further validations are expected to bring actualizations to the model, and future work is essential to keep the model actualized.

References

1. Martinez, L.M.; Viegas, J.M. Assessing the impacts of deploying a shared self-driving urban mobility system: An agent-based model applied to the city of Lisbon, Portugal. *Int. J. Transp. Sci. Technol.* **2017**, *6*, 13–27. [CrossRef]

2. Pereira, A.R.; Carneiro, P.S.; Da Silva, M.M. Evaluating a Reference Model for Shared Autonomous Vehicles using Phoenix. In Proceedings of the 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Tenerife, Canary Islands, Spain, 19–21 July 2023; pp. 1–10.
3. Sousa Carneiro, P.D.; Reis Pereira, A.; Silva, M.M.D. A Model for the Deployment of Shared Autonomous Vehicles in Urban Areas Based on the Research Literature. *Int. J. Transp. Dev. Integr.* **2023**, *7*, 199–213. [CrossRef]
4. Prat, N.; Comyn-Wattiau, I.; Akoka, J. Artifact Evaluation in Information Systems Design Science Research? A Holistic View. In *Proceedings—Pacific Asia Conference on Information Systems, PACIS*; AIS Electronic Library (AISeL): Atlanta, GA, USA, 2014.
5. Fettke, P.; Loos, P. Classification of reference models: A methodology and its application. *Inf. Syst. E-Bus. Manag.* **2003**, *1*, 35–53. [CrossRef]
6. Moody, D.L.; Soindre, G.; Brasethvik, T.; Solvberg, A. Evaluating the Quality of Information Models: Empirical Testing of a Conceptual Model Quality Framework. In Proceedings of the Conceptual Modeling ER 2002, Chicago, IL, USA, 13–16 October 2003. LNCS Volume 2503.
7. Ferber, R.; Sheatsley, P.; Turner, A.G.; Waksberg, J. *What Is a Survey?* EMBRAPA-DMQ: Brasilia, Brazil, 1980.
8. Mathers, N.J.; Fox, N.J.; Hunn, A. *Surveys and Questionnaires*; NHS Executive, Trent: Sheffield, UK, 1998; Volume 1998.
9. Martin, E. Survey questionnaire construction. *Surv. Methodol.* **2006**, *13*, 1–13.
10. Lefever, S.; Dal, M.; Matthíasdóttir, Á. Online data collection in academic research: Advantages and limitations. *Br. J. Educ. Technol.* **2007**, *38*, 574–582. [CrossRef]
11. Mills, J.D. SPSS: Textbooks: A Review for Teachers. *Stat. Educ. Res. J.* **2003**, *2*, 59–70. [CrossRef]
12. The Open Group. Archimate 3.1 Specification. 2020. Available online: https://pubs.opengroup.org/architecture/archimate31-doc (accessed on 26 August 2024).
13. Singh, A.S.; Masuku, M.B. Sampling techniques & determination of sample size in applied statistics research: An overview. *Int. J. Econ. Commer. Manag.* **2014**, *2*, 1–22.
14. Molnar, A. SMARTRIQS: A simple method allowing real-time respondent interaction in Qualtrics surveys. *J. Behav. Exp. Financ.* **2019**, *22*, 161–169. [CrossRef]
15. Francis, T.E. *The Best Practices of Executive Coaches with C-Level Executives*; The University of the Rockies: Denver, CO, USA, 2011.
16. Vemoori, V. Envisioning a Seamless Multi-Modal Transportation Network: A Framework for Connected Intelligence, Real-Time Data Exchange, and Adaptive Cybersecurity in Autonomous Vehicle Ecosystems. *Aust. J. Mach. Learn. Res. Appl.* **2024**, *4*, 98–131.