THE IMPLEMENTATION OF DATA SCIENCE AND ARTIFICIAL INTELLIGENCE IN BUILT ENVIRONMENT

van Holst KAC, Kret J, van Oort TPW, Pfandler PB and Tatár M

Abstract

The age of Artificial Intelligence and Data Science has brought with it a rapid onset of technological, academic, and professional revolutions. The reality of these changes are met with skepticism, fear, and by some, incredible joy. This paper is an exploratory study into the influence of Artificial Intelligence and Data Science on the prospective job market for Built Environment professionals, and their attitudes towards these changes. A policy paper regarding curriculum recommendations is an extension of this report.

1 Introduction

The International Telecommunication Union (ITU) researched approximately 116 academic papers on smart and sustainable cities and proposed the following definition:

> ”A smart sustainable city is an innovative city that uses information and communication technologies (ICTs) and other means to improve quality of life, efficiency of urban operation and services, and competitiveness, while ensuring that it meets the needs of present and future generations with respect to economic, social and environmental aspects”. International Telecommunication Union (2014)

1.1 Problem statement

Domain managers at the data science and artificial intelligence programme were given feedback to make the programme more of a core business in Breda University of Applied Sciences. Therefore, we were given the assignment by domain consultants to research how our field can impact the other domains (e.g. Built Environment, Hotel Management and Facility Management), more specifically the awareness, attitude, knowledge and perception of data science and AI.

The challenge at hand is to assess the influence of data science and artifical intelligence on the prospective job landscape across the diverse programs of Breda University of Applied Science; also known as domains. Additionally, we need to gauge the existing levels of awareness and knowledge within each program. This information is essential to formulate strategic recommendations for curriculum adjustments across programs, ensuring they align with the evolving demands of the future job market shaped by Data and AI. The focus of this paper is the Built Environment domain.

Due to the pressure of making better and faster decisions a lot of organisations are adopting a data-driven culture, typically based on the use of analytics and business intelligence, that play a significant change in decision-making process. (Harvard Business Review, 2012) Managing and building cities require a lot of decision-making, the development of smart cities make way for a more data-driven society and increased efficiency and effectiveness in urban planning and resource utilisation, enhancing the overall quality of life for their residents.

1.2 Research objective

The aim of this paper is to assess the demand and job market for developing smart cities and how that demand shows, as well as investigating the current level of coverage of this topic in the curriculum of BSc Built Environment at Breda University of Applied Sciences, hereinafter referred to as BUas.

2 Literature review

2.1 The fourth revolution

Klaus Schwab, founder of World Economic Forum, divided the last 250 years into four revolutions. The first one being the invention of mechanical steam power, the second revolution being the invention of electrical power, the third is associated with the rise of Information Communication Technologies (ICT) and computers, which ended late 20th century and now we find ourselves in the fourth revolution: a world of machine intelligence, the smart city, digital health care and massive proliferation of information devices. (Batty, 2018)

2.2 The context of sustainable smart cities

The ITU performed a keyword analysis in which they studied the frequency of keywords in the academic papers. These keywords were grouped in different categories and given a percentage of occurrence. Table 1 shows these categories and occurrences and the importance of ICT, communication, intelligence, information in the discussion of smart cities:

2.3 Smart Cities and Technologies

The increasing urbanisation of the world’s population is a significant issue that demands attention. In the 1950s, only 30 percent of the global population lived in cities. However, by 2014, this urbanisation rate had surged to 54 percent, and according to United Nations forecasts, it is expected to rise even further, reaching 66 percent by 2050. In the current state of most urbanized areas, this could mean premature deaths to millions of people caused by air pollution and several other environmental factors. Wang (2015)

Category	% Occurrence
Quality of life and lifestyle	6%
Infrastructure and services	17%
ICT, communication, intelligence, information	26%
People, citizens, society	12%
Environment and sustainability	17%
Governance, management and administration	10%
Economy and Finance	8%
Mobility	4%
Total	100%

Table 1. Logical groupings of keywords and occurrence. International Telecommunication Union (2014)

Governments worldwide are actively exploring the adoption of smart city concepts as a means to achieve sustainability goals and elevate the quality of life for their urban populations. This entails the integration of big data applications to support smart city initiatives. Smart cities leverage a variety of technologies to enhance the efficiency of essential services such as healthcare, transportation, energy, education, and water management, ultimately aiming to enhance citizen comfort. Alahi et al. (2015)

In 2016, a research study focused on identifying the necessary technologies for building smart cities. However, given the rapid technological advancements that have occurred since then, there is now a noticeable gap in research. This gap arises from the significant changes and developments in the field of smart cities over the years. As a result, the findings of the 2016 study may no longer adequately address the current complexities and technological requirements associated with modern smart city development. Therefore, there is a clear need for updated research efforts to comprehensively address the evolving technological needs of smart cities in today's context. Mohanty (2016)

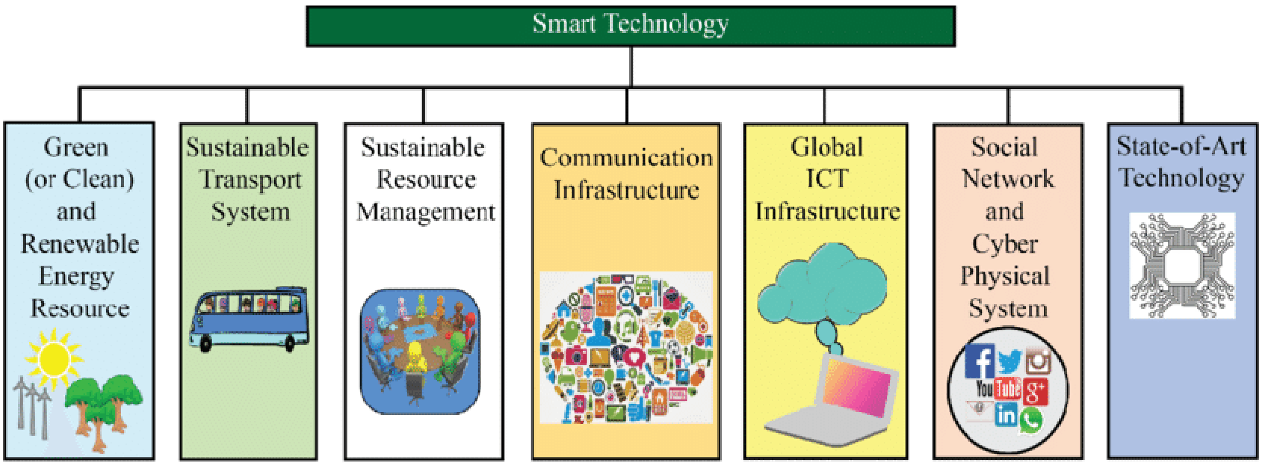

Figure 1. Technologies required for a smart and sustainable city. Mohanty (2016)

2.4 AI in Smart Cities

As smart cities develop, they give way for an entirely new category of data-driven decision making. From Big Data to Mega Data, Internet of Things (IoT) enabled data collection devices for smart cities will exponentially increase the amount of information we have about all aspects of a city. In order to make the most of all this new information, we must employ Machine Learning (ML) algorithms to analyze the vast amounts of data which a human simply could not comprehend.

One example of such an application is Predictive Maintenance. ML models connected to smart city sensors are able to instantaneously assess patterns and trends that may be overlooked by a human eye. This allows them to constantly update and monitor the 'health' of a city, and make smart suggestions on what needs to be optimized before a problem even occurs. (Alahi et al., 2023)) These suggestions may include reducing energy usage in a particular section of an office building, halting water supply to a region of a city because of a pipe-leak, controlling traffic lights that currently have no usage to enable smoother flow of infrastructure, and many more.

Currently, only 30 percent of smart city applications integrate AI to streamline analysis and decision-making. However, AI powered smart city initiatives are predicted to rapidly increase by 2025. (Alahi et al., 2023) para. 4) Besides improved resource management and increased sustainability, AI can also help in making cities a safer place. By analyzing CCTV footage, microphones and other sensors, AI can detect crime and alert authorities before it even happens. The rise of autonomous automobiles also enables a reduction of driving-related accidents, traffic related air pollution, and creates more opportunities for surveillance and analysis. (Cugurullo, 2020) Resource management is one thing, but monitoring and analyzing civilian behavior raises major privacy and data security concerns. These will need to be seriously assessed and considered when making the move to a more monitored and optimized future.

The applications of AI for a better future are endless. The bigger challenge will be arming future professionals with the necessary skills and tools to properly integrate, and manage the rapidly evolving technologies of tomorrow.

2.5 Cognitive aspects

The Knowledge-Attitude-Behaviour (K-A-B) model (Zhang and Dafoe, 2019) explains how people learn, feel, and act in a certain order. For example, when people learn about the good and bad sides of artificial intelligence (AI), they start to have opinions about AI. These opinions can then affect what they do, like using AI or supporting responsible AI. This theory helps us understand how knowledge, opinions, and actions are connected. In the context of this study, the research focuses on behaviours related to healthy eating and physical activity among middle-school-aged students. However, we can draw a parallel to the field of artificial intelligence (AI). In the same way that knowledge about health influences attitudes and behaviours in the health context, individuals who gain knowledge about AI may develop specific attitudes toward AI technologies, which can influence their behaviours in the realm of AI adoption and usage.

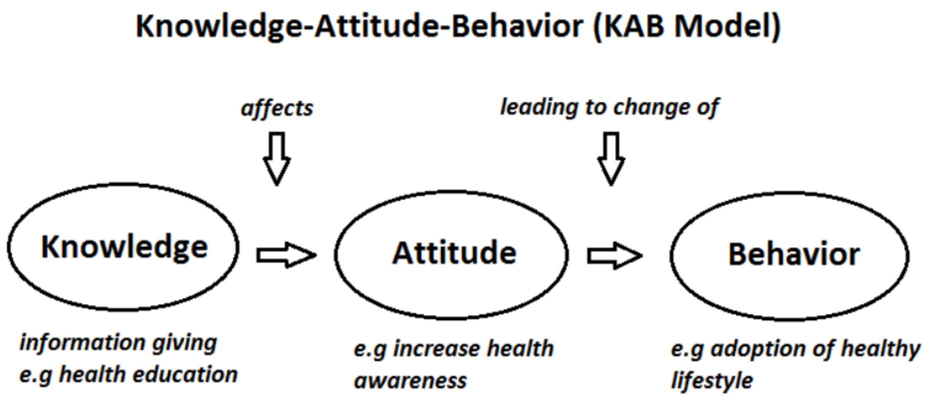

Figure 2. KAB Model flowchart

1. Awareness: knowing something; knowing that something exists and is important.(OED, 2023b)
2. Knowledge: the information, understanding and skills that you gain through education or experience.(OED, 2023c)
3. Perception: an idea, a belief or an image you have as a result of how you see or understand something.(OED, 2023d)
4. Acceptance: the act of accepting a gift, an invitation, an offer, etc. Commonly used is also acceptability the degree to which something is agreed or approved of by most people in a society.(OED, 2023a)

To develop recommendations for shaping attitudes and behaviours toward artificial intelligence (AI), we have designed a survey that assesses participants' AI knowledge, attitudes, and behaviours. This survey gathered data on AI-related knowledge levels, attitudes, and current behaviours. Using the survey results and the K-A-B model, we have formulated recommendations for stakeholders, focusing on addressing knowledge gaps. These recommendations will be communicated to policymakers, educators.

2.6 Research gap

We have identified technologies required to make a city a smart and sustainable city. However, the recent developments of data science and AI in combination with these technologies is yet to be researched. With this study we intend to revise the current state of the industry and approach the technologies from a data science and artificial intelligence perspective, in order to create curriculum recommendations in the form of a policy paper.

3 Research questions and hypotheses

3.1 Research questions

RQ: How does and will AI affect future jobs and how do we prepare students for this?

SQ1: What is the status regarding awareness, knowledge, perception and acceptance of staff and students towards this?

SQ2: How do students and staff compare on these metrics?

SQ3: What are the potential observed predictors for the outcome variables of research question 1, 2 or 3?

3.2 Hypotheses

1. RQ1 Null Hypothesis: There is no difference in psychometric observations.
2. RQ1 Alternative Hypothesis: Staff, specifically lecturers are more aware of AI tools employed in the industry today, new developments and their potential to reshape a city, but students are more accepting of this transition.
3. RQ2 Null Hypothesis: Student and staff have similar measurements across all categories, points lie no more than one unit apart.
4. RQ2 Alternative Hypothesis: Staff will show an average awareness score of 4, but will be hesitant and skeptical. Students will show awareness levels of 3, but the rest of the categories will have an average score of 4.
5. RQ2 Null Hypothesis: Student and staff have similar measurements across all categories, points lie no more than one unit apart.
6. RQ3 Alternative Hypothesis: Age is likely the largest contributing factor, and we will observe substantially more people under the age of 30 who embrace the new technologies.

4 Methodology

4.1 Design

Since the combination of data science and AI in Smart City applications is yet to be thoroughly researched, this paper aims to establish a starting point for further research on the attitudes, and consequences of such a transition. Due to a brief research period, small sample size, and lack of academic history, this study is exploratory by nature.

This research uses a combination of quantitative and qualitative research methods, also known as mixed methods. The quantitative part of this research will collect data on a larger scale and generate initial insights. This will be followed by a smaller-scale in-depth collection of information from our sample, when we identify small groups or people with different and interesting experiences.

Due to the time constraints, we chose to do a cross-sectional study type, based on our findings we can recommend to use this study as a starting point for a longitudinal continuation if deemed necessary.

4.2 Adhering to FAIR principles

The FAIR principles are guidelines to make your data Findable, Accessible, Interoperable, and Reusable, for more effective data sharing and analysis.

Findable: In the data collection process, all data qualitative and quantitative are assigned a Response ID or interviewee ID, for unambiguous referencing. In the survey metadata such as (duration, time, date, location, progress) is automatically recorded, during the interview this will be done manually or through the voice recorder. To whom it may concern, the data will be stored and shared with those involved on the GitHub repository.

Accessible: The data collected will be made available to the those involved, according to a Data Management Plan which can be found in the appendix. Since it is shared on a GitHub repository the authorisation is taken care of automatically by the educational institute (the administrators of the GitHub repositories). Only our team and lecturers have access to the data. The metadata will be included in the repository separately, as well as in the original data set.

Interoperable: The data will be stored in standardised data formats (Excel, CSV, Word, mp3) and will be using standardised vocabularies (likert scales) throughout the entire survey for consistency. Quantitative and qualitative data will be linked if possible using hyperlinks and persistent identifiers.

Reusable: the data will be interpreted and described in an unbiased manner in the data analysis part of this research. Re-usability protocols are written in the data management plan. This Methodology section gives a clear overview of how, why and when the data is recorded and processed. This is done using domain standards, under the supervision of M. Buckens and B. Heijligers.

4.3 Quantitative research

In the research team, we chose to send out a survey to our BUas students and lecturers to get information on the industry and curriculum. This method enabled us to gain insight on general, and domain specific categories. Naturally, there is overlap between the 300+ questions that were included in the survey, but the final table for analysis contains about 50 questions. For reproducibility, we include the survey in the appendix and provide all information required in the methodology section of this paper.

4.4 Qualitative research

For the qualitative part of the research, we chose to perform in-depth semi-structured interviews with professionals and students from the field. This enables the participants to talk about a topic in their own way, while still obtaining answers on specific questions. The interviews also gave us the flexibility of setting up in-person sessions with one person, rather than finding time in overlapping schedules to do a focus group. Questions asked were designed to measure the psychometric variables mentioned previously, but the explanation behind them as well. Knowing the why to the psychometric variables is crucial to provide valuable insights to the problem statement.

4.5 Sample and population

This study focuses on the population of professionals working in the built environment industry, more specifically urban design, urban planning and mobility. Because we studied students and lecturers, the age window is 18-67. We chose this window for the convenience of not requiring parental consent, and people relevant in the current workforce. The unit of analysis is on an individual level, because we have not gathered data on a higher level, like organisations.

In order to ensure the validity of the research and to be able to generalise the findings, the sample should represent the current industry. The chosen sampling method for this research is a non-probability way of sampling, convenience sampling. Convenience sampling is often used by students, because it is on basis of convenience or ease of access. This is especially true in the case of this study, as we have immediate access to a large population of professionals and students. Depending on the proportions of the results, boot-strapping and creating pseudo-responses could assist in a more accurate picture of the industry as a whole.

4.6 Procedure of data collection

For the quantitative part of the research we used Qualtrics to conduct the survey, because it is provided to us by the educational institute and already adheres to our policies with the license we have. Qualtrics is a cloud-based platform, that allows users to distribute and create surveys easily. This survey was be sent out digitally and through the use of posters to our participants, and the data was stored online in Qualtrics and later exported for analysis.

Demographic data is be numerical or categorical (depending on the demographic), while the data about the awareness, expectations and the use of data science and AI technologies will mostly be measured on a five-point-likert scale (agree/disagree).

The instrument for data collection in the interviews was the Microsoft Teams voice recorder and transcription service. The text document was used as an indexing file to record metadata and to provide information in general of the opinions to be explored in the interview. The voice recorder will be used to record all that is said in the interview, so that it can be replayed in the analysis process.

The data was be stored locally and in our GitHub repository according to the data management plan, which is based of the current BUas data management protocol. More information on this can be found in the appendix.

4.7 Procedure of data analysis

Upon receiving the final data set of the quantitative results, an in-depth cleaning was performed. There were two data sets available for analysis. A text, and a numerical one, which included text questions but the responses had been converted to numerical values. Most of the cleaning and analysis was completed on the numerical data set. Responses were filtered to the Built Environment domain, where metadata containing columns were dropped, and only the response-id was kept for manipulation purposes. Due to the large amount of questions, a system was implemented in the survey, where half of the respondents answered questions asked by the lecturers, and half of the respondents answered the questions posed by our team. This meant we had two data sets covering general attitudes, with one being more focused on built environment applications. However, both sets of questions roughly measured general attitudes. We will call questions by lecturers general, and the team's questions domain specific. Columns focusing on other group's research was dropped, and so were incomplete responses. This left us with approximately 22 responses. The analysis focused on the domain specific questions, which yielded 11 responses, but measured parallel attitudes to the general questions. Irrelevant questions which revealed little insight were manually sorted and dropped accordingly.

A major hurdle encountered during analysis was the inconsistent conversion of Likert-scale responses to numerical values. In order to do this, the text data-frame was approximated to the current dimensions of the numeric data-frame, responses were encoded correctly, and replaced in the numeric data-frame according to response-id. The final data-set contained 11 responses and 55 columns, 50 of which were relevant to measuring our variables of interest.

After cleaning, a descriptive analysis, a T-test and (multiple) linear regression was used to analyse the survey data. These analyses were selected for the insights they reveal in terms of statistical significance. Along with a correlational analysis, the linear regression will reveal if predictions can be made from the data points. It is safe to assume that there is some correlation between age, experience, role, and the attitude descriptors. However, correlation may not equal causation. The T-Test will further support the linear regression and p-values can be cross-referenced for further understanding.

For the qualitative data, we performed a thematic analysis. This method uses transcribed interviews, and codes them according to recurring topics. These topics were divided into themes, and correlations between 'codes' were be visualised and explored. This method was chosen for efficiency and collaboration. The insights from the thematic analysis support the decision to focus the quantitative analysis on the group's domain specific questions. Despite the wide variety of participants in the interviews, attitudes and knowledge are similar throughout the population, with expected outliers.

4.8 Ethical considerations

At the start of the survey and interview we will inform the participants about the topic of the research and their ability to withdraw from the research or not answer questions they do not want to answer. This process will be done verbally in the case of in-person interviews, or through a consent form in the case of a survey.

The participants are not required to fill in any data that makes them identifiable and the questions asked are not sensitive of nature. They will also be informed about the storage and accessibility of the data. And that the data collected will be recorded as honest as possible and not manipulated.

5 FINDINGS

The problem statement ultimately assigns two objectives to be explored. The first objective is measuring the attitudes and knowledge of the population in regard to AI within the industry; the second objective is finding the requirements to successfully prepare future professionals for a changing job market which heavily utilizes AI. The following results of the mixed-methods analysis should be interpreted with these objective in mind. These findings are from the individual perspective of Tatár M.

5.1 Quantitative Results

Responses of the Qualtrics survey were cleaned and sorted accordingly. Metadata was erased, except ”ResponseId”, to allow for identification of individual responses for data manipulation. The only columns that remained for analysis were the general psychometric questions posed by our lecturers, and the domain specific questions which were created by the team. Remaining responses were grouped by student or educator, and the means for all questions were calculated. The calculations were then grouped, and the average response for each psychometric variable was computed.

Visualizing the levels of awareness, acceptance, perception, and knowledge reveals that the alternative hypothesis for RQ2 is invalid when looking at awareness, but valid for acceptance. Educators show higher levels of acceptance towards the presence and possibilities of data science in built environment. However, students are more aware of this transition. See appendix 1.

Looking at a correlation table comparing the relationships between psychometric variables and the age of the participant reveals some interesting insights.

	Age Midpoint	**Acceptance**	**Awareness**	**Knowledge**	**Perception**
Age Midpoint	1.00000000	-0.01382334	-0.3383279	-0.80579879	-0.2449238
Acceptance	-0.01382334	1.00000000	0.2939898	-0.09137722	0.1362465
Awareness	-0.33832795	0.29398977	1.0000000	0.59784028	0.4266352
Knowledge	-0.80579879	-0.09137722	0.5978403	1.00000000	0.4127429
Perception	-0.24492379	0.13624649	0.4266352	0.41274286	1.0000000

Table 2. Correlation matrix of the study variables

According to table 2, most of the variable pairings have little to no correlation, but there are two which are worth mentioning. Awareness and knowledge have a correlation coefficient of 0.597, which supports the notion that awareness naturally leads to more knowledge about the subject, as one cannot exist without the other. We can also see that age and knowledge have a -0.805 correlation coefficient, which indicates an inverse relationship between age and knowledge. This means the older the participant is, the more likely they are to be unaware of certain technologies and advancements.

Linear regression summaries suggest age and knowledge to have a strong statistical significance, with this comparison being the only observation with a p-value lower than 0.05, thus supporting the alternative hypothesis. See the table ”Knowledge”.

Knowledge

Call:

```
lm(formula = knowledge ˜ age_midpoint + experience_midpoint, data = domain_data)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(¿—t—)
(Intercept)	3.8774	0.2319	16.717	1.4e-05 ***
age_midpoint	-0.0322	0.0093	-3.460	0.018 *
experience_midpoint	0.0512	0.0282	1.817	0.129

Furthermore, perception and age also show a high p-value, at 0.92, with experience also having a significantly larger p-value than in other variables, at 0.61. This proves the null hypothesis correct for RQ3. There is no statistical significance between perception and age, or experience in the field. In all other analyses, the p-values also support the null hypotheses. No prediction of attitudes can be made based off of these variables. See the following table.

Perception

Call:

```
lm(formula = perception ~ age_midpoint + experience_midpoint, data = domain_data)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(¿—t—)
(Intercept)	3.6780	0.5976	6.155	0.00165 **
age_midpoint	0.0025	0.0240	0.105	0.9201
experience_midpoint	-0.0390	0.0726	-0.538	0.6138

The t-test results comparing student and staff performances on the four variables are as follows.

Table 3. T-test Results for Awareness

Metric	**Awareness**
Estimate	-0.09274376
Estimate for Group 1	3.066667
Estimate for Group 2	3.15941
Statistic	-0.2518011
P-value	0.8088461
Degrees of Freedom	6.607349
95% CI Lower Bound	-0.9742885
95% CI Upper Bound	0.788801
Method	Welch Two Sample t-test
Alternative Hypothesis	Two-sided

Table 4. T-test Results for Knowledge

Metric	**Knowledge**
Estimate	-0.4460317
Estimate for Group 1	2.833333
Estimate for Group 2	3.279365
Statistic	-1.829503
P-value	0.1298712
Degrees of Freedom	4.753408
95% CI Lower Bound	-1.082644
95% CI Upper Bound	0.1905807
Method	Welch Two Sample t-test
Alternative Hypothesis	Two-sided

The highest p-value which can be observed is for awareness at 0.8, and the lowest for knowledge, at 0.13. These results indicate all null-hypotheses to be correct. A significant pattern cannot be observed between the attitudes, or knowledge of staff and students within the population.

5.2 Qualitative Results

Coincidentally, we interviewed 11 participants, which is the same number as the relevant responses for the quantitative analysis. We began the analysis by identifying 9 re-occurring themes. These were: comparison between students and staff, positive, negative feelings towards AI, challenges, opportunities, cognitive aspects (knowledge, awareness, acceptance, perception), specific technologies, necessary skills, and curriculum. Interview transcripts were reviewed and the relevant sections were transferred to a collective document.

The findings revealed similar insights to our quantitative analysis, without much additional information. The results were expected to demonstrate a clear divide between students and staff. Most respondents believed that awareness among the population was low, and there were mixed feelings towards AI throughout the sample. For example, a research manager who took part in creating the BE curriculum had negative feelings towards AI mostly because of his fear towards over-dependence. This can be seen in interview 4.2 at the timestamp 4:25 of the "Thematic Analysis" document. Ethical considerations contribute to fears relating to the misuse and sharing to data by large corporations. This can be seen in interview 5.2 at timestamp 0:32. A student showed negative feelings because of their perception of AI in interview 5.1 at timestamp 0:06.

Most of the negative feelings towards AI came from ignorance, or the overestimation of AI's capabilities. The ethical concerns are still very valid however. Even if well-founded and strict legislature is introduced, true intentions and actions of large corporations can have enormous consequences, and wrongdoing only has to happen once for there to be irreparable damage. Establishing limits for the jurisdiction of data scientists and built environment professionals is also mentioned as an integral component of implementing a balanced curriculum. Furthermore, there is an outcry for transparency regarding data ownership, access, and sharing between large corporations. Consolidation of knowledge and data is regarded as a necessary step towards an industry which utilizes AI to its fullest potential without making its professionals dependent on its assistance. However, this is the only step which is virtually impossible to address due to the greed of leading technological entities.

Attitudes towards AI were positive overall, and negative feelings could be easily subsided after basic education of what AI is, how it works, and its realistic potentials. The lack of knowledge and homogeneity in themes suggested that there is no statistical significance to be found in the qualitative data. Rather, it should be considered as the backbone for curriculum recommendations due to the deep insights. Resistance towards implementation of AI into the curriculum was only voiced through the over-reliance on AI, and professional being unable to work without it. Therefore, it is paramount to ensure students' competence before allowing them to use AI assistance.

5.3 Discussion

Even though the research generated some interesting findings, the statistical insignificance of the analyses can be attributed to the expedited nature of this study, and the lack of participants. Even though 580+ responses were recorded, we were only left with 22 respondents, half of which answered very general questions posed by our mentors. It is doubtful however, if more responses would yield significantly different results.

An unexpected problem arose while gathering qualitative data, as we found that a remarkably large portion of the Built Environment students and lecturers are away on excursions. This meant that most of the people we wanted to interview were unavailable, wouldn't reply, or would cancel their interviews. Some of us resorted to a brute-force method of finding interviewees, as we tried to find people on campus who had time, or could schedule an interview within a few days. There may be further insights to be uncovered if a larger sample size was interviewed. A focus group could also be created in order to gain deeper understanding of where students' interests lie when it comes to their education in terms of AI.

Further research can be conducted for specific needs, but these findings should be a solid foundation for the integration of AI into the curriculum. A general level of understanding can be gained of the attitudes and knowledge among the population of the university from the sample we gained.

From an individual perspective, I can confidently hypothesize that attitude levels (awareness, knowledge, perception, and acceptance) are similar throughout the other studies conducted by our domain.

6 Conclusion

Let us revisit the research questions directly.

RQ: How does and will AI affect future jobs and how do we prepare students for this?

Based on our literature review, the topic of smart cities encompasses close to all applications of AI within the Built Environment industry. It was chosen as the topic of focus precisely for this reason. It has applications in all three specializations of Built Environment, and gives an excellent overview of the competencies needed in order to prepare students for the evolving job market. AI and data science will not replace jobs in their entirety, but rather significantly reduce the complexity of the tasks within those jobs. It will be an enormous aid and versatile tool only to those that know how to maximise its benefits.

SQ1: What is the status regarding awareness, knowledge, perception and acceptance of staff and students towards this?

Observations show attitude levels to be low. The majority are unaware of current applications, and have misinformed expectations. People are fearful for the right reasons, but privacy and ethical concerns can only be resolved by governments. The benefits far outweigh the risks, which, with education, caution, and the correct approach can easily be mitigated. The null hypothesis is accepted.

SQ2: How do students and staff compare on these metrics?

Staff outscore students in acceptance, and students in all other metrics. However, this result can be refuted through the categorization of survey results. The null hypothesis is accepted.

SQ3: What are the potential observed predictors for the outcome variables of research question 1, 2 or 3?

According to the correlational analysis, and linear regression, Age and knowledge have a relatively strong correlation, thus rejecting the null hypothesis, and accepting the alternative. However, all other variables show age to have virtually no statistical significance. If anything, the analysis suggests that older participants are more willing to embrace the new technologies even though they know little about them. This should encourage the younger generations that they have the older ones' faith in implementing these changes correctly. The null hypothesis is accepted.

Preparing future professionals for an evolving job market is incredibly important, and can be done by carefully implementing AI into curriculums. However, it should be approached cautiously and methodically as to avoid indirectly harming their academic and professional development.

A Appendix 1

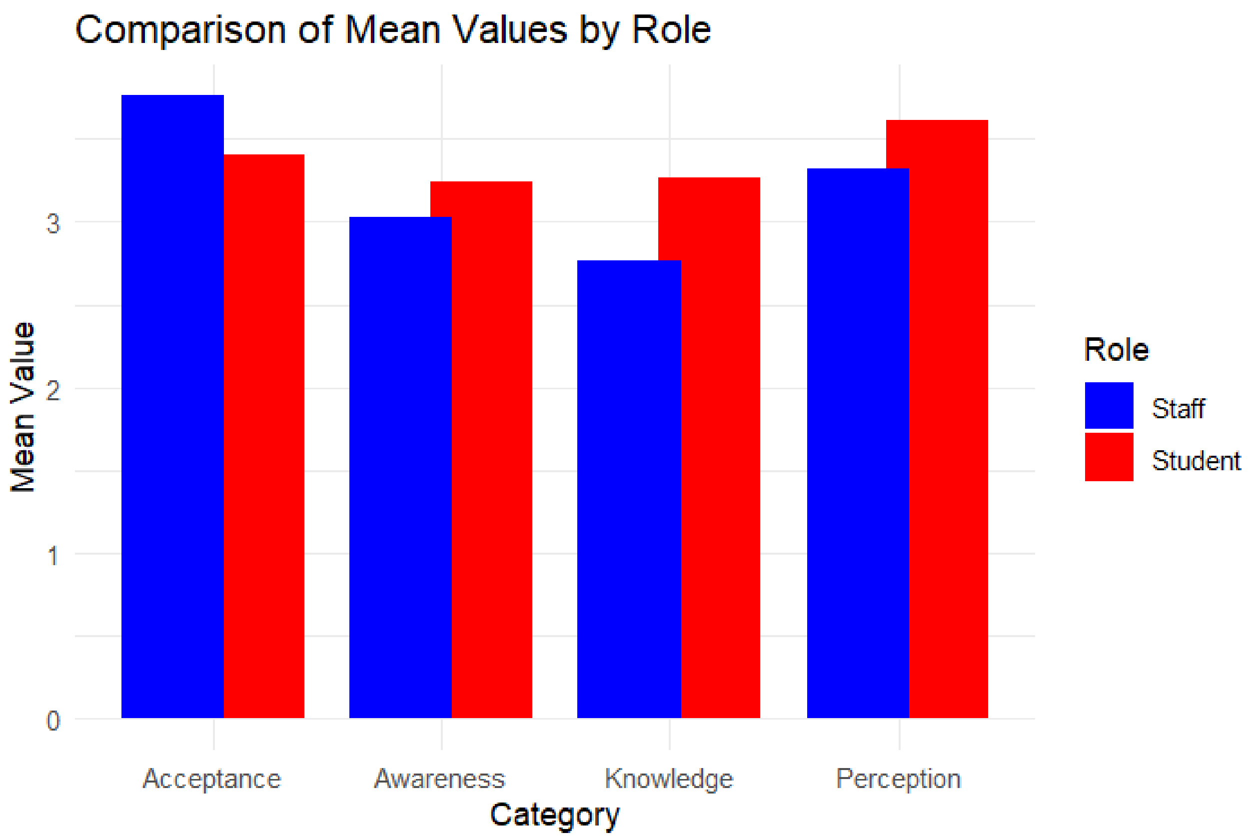

Figure 3. Means of Staff vs. Students

B APPENDIX 2

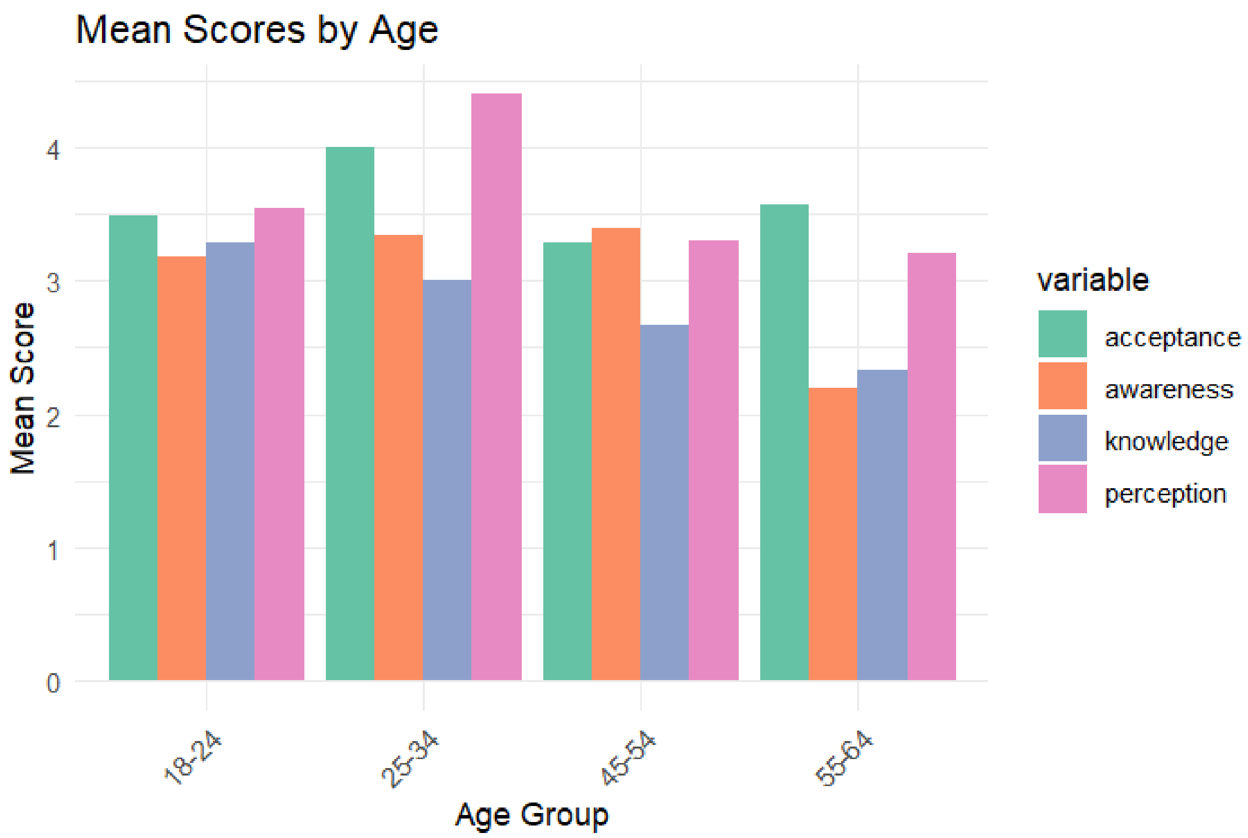

Figure 4. Means by Age

C APPENDIX 3

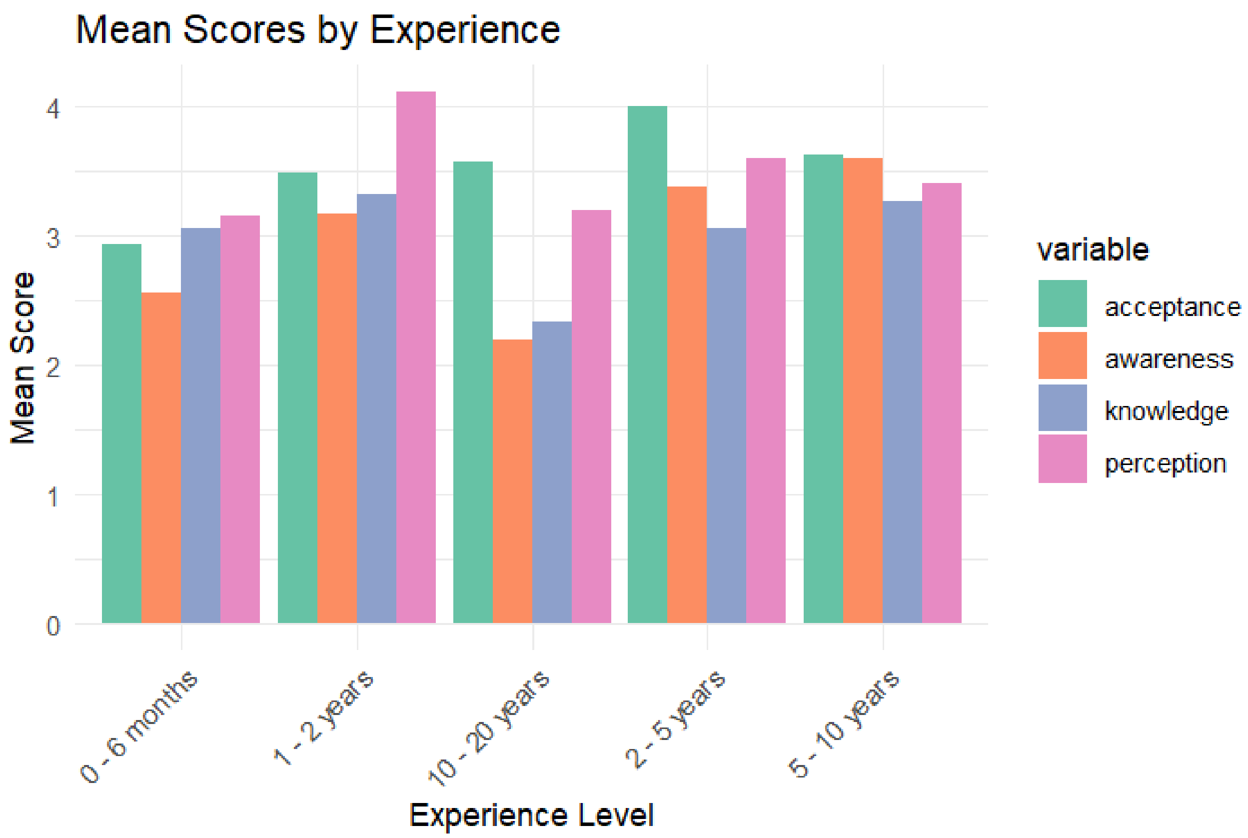

Figure 5. Means by Experience

D APPENDIX 4

Table not directly mentioned in paper.

Table 5. T-test Results for Perception

Metric	Perception
Estimate	-0.1071429
Estimate for Group 1	3.5
Estimate for Group 2	3.607143
Statistic	-0.3208969
P-value	0.7631676
Degrees of Freedom	4.335721
95% CI Lower Bound	-1.006524
95% CI Upper Bound	0.7922383
Method	Welch Two Sample t-test
Alternative Hypothesis	Two-sided

E APPENDIX 5

Table not directly mentioned in paper.

Table 6. T-test Results for Acceptance

Metric	Value
Estimate	0.4116
Estimate for Group 1	3.7857
Estimate for Group 2	3.3742
Statistic	1.4538
P-value	0.1917
Degrees of Freedom	6.6270
95% CI Lower Bound	-0.2656
95% CI Upper Bound	1.0887
Method	Welch Two Sample t-test
Alternative Hypothesis	Two-sided

F APPENDIX 6

Due to privacy reasons, access to the Data Management Protocol can only be granted through application for reuse, or reproduction through the Breda University of Applied Sciences.

G APPENDIX 7

Data Storage Protocol

van Holst KAC, Kret J, van Oort TPW, Pfandler PB and Tatár M.

File Naming & folder structure

This file is a guideline to be followed by the team during the project to ensure it is available for reuse. When you collect data of any sort from interviews make sure to use the following naming convention, guidelines have been provided by BUas.

A project acronym
Content description
File type information
Date : (make a choice how to write the date: for example DD-MM-YYYY or YYYY-MM-DD)
Creator name or initials
Version number
Use - or _ to separate elements in a file name

For example:

07022020_interview_BUaslibrary_audio.wav
07022020_interview_BUaslibrary_transcript.txt

Source:https://buas.libguides.com/rdm/storage

Version control

For version control we use github. Instead of naming the file after its version we have decided to make all changes in github which allows us to work more structurally and we can always go back to an earlier version of a file.

Readme files

In each file of significance a Readme.txt will be present explaining what is in the file and a short description of what is in the file.

H APPENDIX 8

[Survey Link](https://buas.eu.qualtrics.com/jfe/form/SV$_d6gVrdDpWIqjQRE$)

References

(2023a). Acceptance. 25/09/2023.

(2023b). Awareness. 25/09/2023.

(2023c). Knowledge. 25/09/2023.

(2023d). Perception. 25/09/2023.

Alahi, M. E. E., Sukkuea, A., Tina, F. W., Nag, A., Kurdthongmee, W., Suwannarat, K., and Mukhopadhyay, S. C. (2015). Integration of iot-enabled technologies and artificial intelligence (ai) for smart city scenario: Recent advancements and future trends. *Journal of Cyber Physical Systems*, 1(1):25–48. Link to article.

Alahi, M. E. E., Sukkuea, A., Tina, F. W., Nag, A., Kurdthongmee, W., Suwannarat, K., and Mukhopadhyay, S. C. (2023). Integration of iot-enabled technologies and artificial intelligence (ai) for smart city scenario: Recent advancements and future trends. *Sensors*, 23(11):5206. Link to report.

Batty, M. (2018). *Inventing Future Cities*. MIT Press, Cambridge.

Cugurullo, F. (2020). Urban artificial intelligence: From automation to autonomy in the smart city.

Harvard Business Review (2012). The evolution of decision making: How leading organizations are adopting a data-driven culture.

International Telecommunication Union (2014). Smart sustainable cities: An analysis of definitions. Technical report.

Mohanty, S. P. (2016). Everything you wanted to know about smart cities. *ResearchGate*. Link to report.

Wang, J. (2015). A literature survey on smart cities. *Sensors*, 15(9):21003–21024. Link to PDF.

Zhang, B. and Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. *Available at SSRN 3312874*.