Tampere University

Lotte Karlsson

AI FOR DYNAMIC PASSENGER INFORMATION IN PUBLIC TRANSPORT

Master's Thesis
Faculty of Information Technology
and Communication Sciences
Examiners:
Tiia Järvenpää
Jyke Savia
November 2025

ABSTRACT

Lotte Karlsson : AI For Dynamic Passenger Information in Public Transport
Master's Thesis
Tampere University
Faculty of Information Technology and Communication Sciences
November 2025

Artificial intelligence methods are being increasingly considered for new application domains, including passenger information in public transport. This master's thesis examines this intersection of topics through a narrative review – to connect possible aligning problem and solution areas.

Modern passenger information in the field of public transport is firmly in a real-time and mobile information era. With the rise of accessible GPS technology in the 2000s, real-time vehicle location systems became more affordable and efficient for public transport providers. Gradually, this information became easier to access for individual passengers through consumer mobile devices. Following these trends in recent decades, advancements in artificial intelligence provide new opportunities for passenger information.

The review discovered a rising interest in personalization in passenger information research. Related to this, large volumes of data produced by passenger information systems are relatively underutilized. This data could be leveraged to create better passenger information systems. Personalization is thought to be one way to leverage this data – to better serve each user according to their passenger information needs. The mobile platform, while currently highly used in current public transport systems, is thought to be increasingly important as personalization platform.

Artificial intelligence methods are also thought to be helpful for personalization and otherwise enhancing passenger information systems and their related services. Artificial intelligence tools have capabilities to detect complex and abstract patterns from various types of data.

Overall, this review found that while some artificial intelligence technologies present beneficial potential in practical studies, real-world trial deployments are less frequent. Some factors hindering deployment are a lack of implementation frameworks, limited access to model training datasets, and a lack of frameworks for retaining explainable and ethical behavior. Current limitations are expected to evolve as research in the field progresses rapidly.

Keywords: artificial intelligence, passenger information, public transport, personalization

TIIVISTELMÄ

Lotte Karlsson : Tekoälyn hyödyntäminen dynaamisen matkustajainformaation edistämiseen julkisessa liikenteessä
Pro gradu -tutkielma
Tampereen yliopisto
Tietojenkäsittelytieteiden tutkinto-ohjelma
Marraskuu 2025

Tekoälyä harkitaan yhä useampiin uusiin sovellusalueisiin, mukaan lukien joukkoliikenteen matkustajainformaatio. Tämä pro gradu -tutkielma tarkastelee näiden aiheiden leikkauspistettä narratiivisen kirjallisuuskatsauksen avulla. Tavoitteena on yhdistää mahdollisia ongelmia ja ratkaisuja.

Moderni joukkoliikenteen matkustajainformaatio on vahvasti reaaliaikaisen tiedon ja mobiililaitteiden aikakaudessa. 2000-luvulla yleistyneen GPS-teknologian myötä ajoneuvojen paikantaminen muuttui joukkoliikenteen tarjoajille edullisemmaksi ja tehokkaammaksi. Vähitellen reaaliaikainen tieto saavutti myös matkustajat mobiililaitteiden kautta. Viime vuosikymmenten kehitystä seuraten tekoäly tarjoaa nyt uusia mahdollisuuksia.

Katsauksessa havaittiin kasvavaa kiinnostusta matkustajainformaation personointiin. Tähän liittyen matkustajainformaatiojärjestelmien tuottamia suuria tietomääriä hyödynnetään edelleen suhteellisen vähän. Nämä tietomäärät voitaisiin hyödyntää parempien matkustajainformaatiopalvelujen rakentamiseen. Personoinnin avulla voitaisiin tarjota matkustajainformaatiota suhteessa jokaisen käyttäjän omien tarpeitten mukaan. Mobiilialustat, jotka ovat jo nykyään laajalti käytössä moderneissa matkustajakäyttöliittymissä, nähdään tärkeänä personoinnin alustana.

Tekoälyä pidetään tärkeänä välineenä sekä matkustajainformaation personoinnin tukena että laajemmin muiden matkustajainformaatiojärjestelmien kehittämisessä. Sen tarjoamien työkalujen avulla voidaan tunnistaa monimutkaisia ja abstrakteja malleja eri tietolähteissä. Näiden tekoälyn kykyjen ajatellaan mahdollistavan suurten datamäärien hyödyntämisen.

Tämä katsaus havaitsi, että vaikka jotkut tekoälyteknologiat osoittavat potentiaalia rajoitetuissa käytännön tutkimuksissa – tosielämän laajempia prototyyppejä on vähemmän. Toteutuksen esteinä pidetään puutteellisia toteutuskehyksiä tekoälyn ja matkustajainformaation piirissä, rajallista pääsyä matkustaja-aineistoihin sekä puutteista selitettävyyttä ja eettistä toimintaa varmistavissa kehyksissä. Katsauksen esittämän nykytilan odotetaan muuttuvan, sillä alan tutkimus etenee nopeasti.

Avainsanat: tekoäly, matkustajainformaatio, julkinen liikenne, personalisaatio

USE OF AI IN THESIS

I have utilised AI tools in my thesis:

☐ No
☑ Yes

The AI tools utilised in my thesis and their purposes are described below:

Names and versions of AI tools: ChatGPT-4 and ChatGPT-3.5

Purpose of using AI tools: AI tools are used in this work to help structure chapters/sections and provide fast feedback/proofreading tools for phrasing and structuring of chapters. They are ***not*** used for text/content generation, text summary, sourcing or any other content-related part of the work.

Sections where AI tools were used: AI tools were used to support brainstorming of chapter structures in most chapters. AI is used to proofread most sections for grammar mistakes and to help against possible awkward phrasings. Possible corrections were still written and edited by the author.

I acknowledge that I am fully responsible for the entire content of my thesis, including the parts generated by AI, and accept accountability for any violations of ethical standards in publications.

PREFACE

This master's thesis was written as part of a project by Teleste Oyj to investigate the potential of new AI technologies for passenger information. The work has been supported by a grant from the Technology Industries of Finland Centennial Foundation (Teknologiateollisuuden 100-vuotissäätiö). A large thank you to both for providing a great opportunity to learn and contribute to an interesting project.

Particular thanks to Tiia Järvenpää, who welcomed me to co-operate with Teleste for my master's thesis – along with all my other friends at Teleste. Thank you to Jyke Savia, who along with Tiia acted as examiners to this work.

Thank you to my family, friends and girlfriend for supporting me along the way in my studies. Your encouragement and guidance have helped me reach the finish line.

Helsinki, 28 November 2025

Lotte Karlson

CONTENTS

1. INTRODUCTION 1
2. THEORETICAL FRAMEWORK 3
  2.1 Definitions.. 3
  2.2 Past Works 5
3. RESEARCH DESIGN 7
  3.1 Databases 7
  3.2 Inclusion and Exclusion Criteria 8
  3.3 Limitations and Challenges of the Study 8
4. RESULTS 9
  4.1 Dynamic Passenger Information 9
    4.1.1 Personalization 13
  4.2 Artificial Intelligence 21
    4.2.1 Computer Vision 25
    4.2.2 General Machine Learning 28
    4.2.3 Natural Language Processing 31
    4.2.4 Generative AI 34
    4.2.5 Agentic AI 38
    4.2.6 Edge AI 40
  4.3 Artificial Intelligence in Public Transport 41
    4.3.1 Personalization and AI in Passenger Information 42
    4.3.2 AI and Passenger Information System Enhancement 44
    4.3.3 AI for Enhancing Passenger Information Analysis 47
5. DISCUSSION 50
  5.1 Research Questions 50
  5.2 Limitations and future work 51
6. CONCLUSIONS 52
REFERENCES 54
APPENDIX A 62

FIGURES AND TABLES

Figure 1 AI, ML and DL relations. Adapted from Sarker (2021) 4

Table 1 Search entries. 8
Table 2 Objects of Personalization. Adapted from Vredenborg et al. (2025) 15
Table 3 Personalization Attributes. Adapted from Vredenburg et al. (2025) 16
Table 4 Evaluation of Personalization. Adapted from Vredenborg et al. (2025)........... 17
Table 5 Levels of Personalization and example table. Directly adapted from van Ardenne et al. (2025) 20
Table 6 Passenger Information Content Types. Adapted from Ait-Ali (2025)............... 62

LIST OF SYMBOLS AND ABBREVIATIONS

AD	Anomaly Detection
AFC	Automatic Fare Collection
AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
ASFF	Adaptively Spatial Feature Recognition
ASR	Automatic Speech Recognition
BiFPN	Bi-directional Feature Pyramid
CCTV	Closed-Circuit Television
CLIP	Contrastive Language Image Pre-training
CNN	Convolutional Neural Network
CV	Computer Vision
DAN	Deep Attention Network
DL	Deep Learning
DM	Diffusion Model
DNN	Deep Neural Network
DT	Digital Twin
FPN	Feature Pyramid Network
FCOS	Fully Convolutional One-Stage Object Detection
GAN	Generative Adversarial Network
GMM	Gaussian Mixture Model
GCN	Graph Convolutional Network
Gen-AI	Generative Artificial Intelligence
GNN	Graph Neural Network
GRU	Gated Recurrent Unit
HAR	Human Action Recognition
HMM	Hidden Markov Model
HRL	Hierarchical Reinforcement Learning
IDS	Intrusion Detection System
IoT	Internet of Things
ITS	Intelligent Transportation System
LLaVA	Large Language Vision Assistant
LLM	Large Language Model
LSTM	Long Short-Term Memory
MAS	Multi-Agent Systems
MATH	Mathematical problem-solving (AI benchmark)
ML	Machine Learning
MLP	Multi-Layer Perceptron
MMMU	Multimodal Understanding and reasoning (AI benchmark)
MVBench	Multi-Video Reasoning Benchmark (AI benchmark)
NAS-FPN	Neural Architecture Search (Feature Pyramid Network)
NLP	Natural Language Processing
PI	Passenger Information
PIS	Passenger Information System
PT	Public Transport
RAG	Retrieval Augment Generation
RBM	Restricted Boltzmann Machine
R-CNN	Region-based Convolutional Neural Network
RL	Reinforcement Learning
RNN	Recurrent Neural Network
RS	Recommender Systems
RTI	Real-Time Information

S-SL	Semi-Supervised Learning
SSL	Self Supervised Learning
SSD	Single-Shot Detector
SVR	Support Vector Regression
SVM	Support Vector Machines
Swin	(Transformer architecture for vision tasks)
TBML	Tree-Based Machine Learning
TTS	Text-To-Speech
TPU	Tensor Processing Unit
UI	User Interface
UX	User Experience
VAE	Variational Autoencoder
VCR	Visual Commonsense Reasoning (AI benchmark)
ViT	Vision Transformer
VLM	Vision-Language Model
WER	Word Error Rate
YOLO	You Only Look Once (object detection architecture)

1. INTRODUCTION

Public Transportation (*PT*) impacts a large amount of people in modern urban environments – enabling access to work, commerce and daily life. Reflecting the importance of PT, it represents a significant expenditure for governments. For example, the average for EU government spending on PT was 2.2% of their GDP. Transport use in the EU was observed to decrease in 2020 due to the COVID-19 pandemic but is recovering since 2022. (Eurostat, 2025)

The recovering interest in creating, maintaining, and using PT aligns with the recent rapid developments and interest in the domain of Artificial Intelligence (*AI*). With AI tools and technologies relying on mass amounts of data to work with, it becomes evident that there is great potential for the combination of PT data and AI. Easier analysis of mass amounts of data, rapidly generating dynamic content, and optimizing tasks and information are examples of AI themes which could potentially enhance current PT systems and create new ones.

Past studies and reviews have connected PT and AI, such as Jevinger et al., 2024, which presented a literary review of the domain. This study supplements the results of that study with a further focus on state-of-the-art AI technologies which the study did not examine. Jevinger et al. also used articles up to the year 2020, and followingly this review continues with articles from 2020 to 2025.

This review aims to conduct two combined literature reviews in the field of ***Artificial Intelligence*** (*AI*) and ***Passenger Information Systems*** (*PIS*), to gain understanding of the state of the art of both topics. The review then aims to use this as background to evaluate the potential to implement AI to public transport systems to enhance dynamic passenger information.

The review is conducted with the main perspectives of three research questions.

- ***RQ1:*** What concurrent needs and interests are found in research for passenger information systems in public transport?
- ***RQ2:*** What kinds of technologies does artificial intelligence present to answer the needs and interests described in RQ1?

The review continues with elaboration of the theoretical framework used to investigate in the following *Chapter 2* - introduces the research methodology, limitations and inclusion criteria in *Chapter 3*, presents results of the literature review in *Chapter 4*. *Chapter 5* continues with a discussion chapter, tackling the research questions and analysis of the results. *Chapter 6* is the summarising conclusion of the study.

2. THEORETICAL FRAMEWORK

This section grounds theoretical basis for the study with key definitions, including *artificial intelligence, machine learning, passenger information, dynamic passenger information* and *personalization*. Some important past works are highlighted - which play part in building general understanding for Passenger Information and Artificial Intelligence research.

2.1 Definitions

Artificial Intelligence can be difficult to define comprehensively – as definitions do not always agree what is included in the domain of AI and what is left outside of it. A proposal by the European Commission (2020) compiled definitions of AI to form a standardized definition. This definition could be summarised in the following way: "*AI systems are software and/or hardware which can tackle given goals digitally or physically, through perception and data interpretation, which is used to reason or process the best actions for achieving the given goal*". This definition is chosen for the purpose of this work.

AI is a massive research topic in which domains are often divided based on various goals that AI may be used for. Subdomains are then more specific areas that may be classified under a domain. Often it may be difficult to categorize some technologies into domains and subdomains – as the technologies commonly overlap. Notably – machine learning is the most common domain overlap in AI technologies. This following list is provided by the European Commission, while not exhaustive due to the discussed reasons – provides a helpful broad overview.

1. **Reasoning**: knowledge representation, automated reasoning, common sense reasoning.
2. **Planning**: planning and scheduling, searching, optimisation
3. **Learning**: machine learning
4. **Communication**: natural language processing
5. **Perception**: computer vision and audio processing
6. **Integration and Interaction**: robotics, automation, connected and automated vehicles.

7. **Services**: AI services
8. **Ethics & Philosophy**: AI ethics and philosophy research

(European Commission. JRC, 2020)

Machine Learning (*ML*) and its subset ***Deep Learning*** (*DL*) are particularly essential concepts. ML refers to AI techniques where pattern detecting, or predictive models are created through training them based on data. DL is then a subset of ML in which layered neural networks are used to similarly learn patterns from data. DL methods are generally the most relevant to this work, as many of the prevalent AI technologies are classified as DL methods. (Sarker, 2021) Further elaboration on these themes is provided in the results chapter regarding AI.

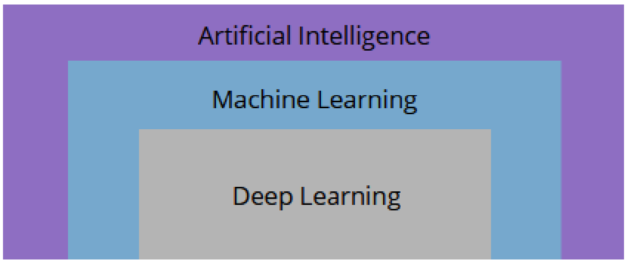

Figure 1 AI, ML and DL relations. Adapted from Sarker (2021)

Passenger Information (*PI*), *Traveler Information* or *Transit Information,* are terms that are commonly used in both academic and non-academic texts but seem to have limited definitions in common dictionary sources. For this work, this definition is used; "*Systems that provide users/passengers with relevant information regarding the status of a public transportation service*" (Ninčević Pašalić et al., 2023). For this work, the most common forms of urban public transport will be focused on – railway systems and buses. Railway systems in this work include train, tram and metro-based systems. ***Passenger Information Systems*** (*PIS*) are then systems which enable using and displaying PI data. They can act as singular systems but usually integrate into larger systems as individual subsystems. For example, a vehicle's GPS data may be transmitted to a central server, to be distributed to other PIS. Aligning with the PI definition, this work does not exclude systems more oriented to vehicle staff and operators.

To define ***Dynamic Passenger Information***, it should be separated from *static passenger information*. Static Passenger information refers to information, which is changed uncommonly, and is largely predetermined, such as planned schedules without any or limited real-time tracking. Dynamic passenger information has a strong connection to ***Real-Time Passenger Information*** (*RTI PI*), which is information that may change depending on circumstances. Examples of this can be user tracking for vehicle position or being notified for a delay. (European Commission, 2025) The distinction between dynamic passenger information and real-time information for the purposes of this work is the following:

1. Dynamic PI is not restricted to short-term reactivity or adaptations such as RTI PI. It may occur in the long-term.
2. Dynamic PI is not restricted to only information, but also the methods how it is delivered to users.

Personalization refers to specialized tools, environments, products which are modified for individuals based on their preferences, needs and characteristics. It has been widely used for commercial purposes for a long time – most notably including marketing. Outside of business connections, it may also be harnessed to improve user experiences (Fan & Poole, 2006) To improve service and better reach passengers – there also exists interest in studying personalization in the Public Transport context. AI also plays part in the joining of public transport and personalization - as the capability of AI technologies to adapt based on data has a strong overlap with studying user/passenger preferences and making adaptations. (Vredenborg et al., 2025) Personalization could for this reason be used to create increasingly dynamic passenger information. Vredenborg et al. (2025) provides a framework for examining personalization - where **personalization objects** are the *targets of personalization*, and **personalization attributes** *affect and determine the resulting alterations*.

2.2 Past Works

AI has been an active research topic for over 60 years, with intermittent breaks in funding and interest levels. In the last few years, AI interest has resurged due to favourable enablers such as advancements in ML and DL, available data sources for model training, hardware improvements and projects that made a strong public impression. (Jiang et al., 2022) The domains and subdomains of AI each hold vast amounts of research. This work is mainly interested in a broad scope and thus literature reviews are the most beneficial, alongside representative individual works. Some examples of broad reviews are Mienye & Swart (2024b) and Sarker (2021), which provide insight regarding DL. DL is thought and observed to be vital for the development of intelligent systems and advanced automation. Both works also introduce important types of DL architectures and techniques. While there are a huge number of different variants of DL techniques – many are extensions of others – such as ***Long Short-Term Memory*** (*LSTM*) being based around ***Recurrent Neural Networks*** *(RNN)* (Sarker, 2021). Mienye & Swart provide supplementary information to the work of Sarker, also including many application domains for DL. (Mienye & Swart, 2024b) Most AI topics also similarly were found to have useful literature reviews such as Large Language Models - a natural language processing technique strongly connected with DL. Minaee et al. (2025) provides historical insight into the topic – moving to modern taxonomy and popular techniques.

Literature reviews by Brakewood & Watkins (2019) and Ait-Ali et al. (2025) synthesize information about passenger information, real-time passenger information and their effects. These works underline the establishment of real-time information (*RTI*) practices in public transport – enabled by cheaper and more available vehicle location systems, data standardization efforts and mobile devices. According to Brakewood & Watkins, RTI systems were reported to decrease both perceived and actual waiting times for PT and increase satisfaction and feelings of security using PT along with other lesser benefits. The work provides valuable insight into the benefits gained from implementing RTI systems for PI.

One study similar to this review was found by Jevinger et al. (2023), which performed a mapping study to evaluate literature written about AI for improving public transport. Their work provides important background – and this work be seen as continuing similar themes with different focus points and timeframe. The work of Jevinger et al. used material from up to the year 2020, and this work will continue with a timeframe of 2020-2025. This study aims to focus more on AI technologies and their technical aspects - to understand why they could align with the needs of PT. Notably, an emergent trend covered by Vredenborg et al. (2025), van Ardenne et al. (2025)– personalization, combines the topics of AI and PI. These works heavily influence this work, as they were notable and broad literature reviews from 2020-2025 – which many others were not found by those criteria.

3. RESEARCH DESIGN

This work approaches the topics of AI and Public transport through *a narrative literature review.* Systematic literature reviews are often preferred for their reproducible methods including qualitative and quantitative synthesis and search technique, but for this work there were several other needs. Narrative literature reviews are aimed at summary, duplicate avoidance, and finding new study areas. More key characteristics are the ability to address multiple research questions or themes, and subjective study selections. (Ferrari, 2015) The narrative review method was selected for these reasons:

1. Better support for multiple small-scale reviews for Artificial Intelligence (*AI*), Public Transport (*PT*) and AI + PT.
2. This work is interested in connecting the needs and interest of PT with tools from AI, which is a summarising and exploratory goal.
3. This area of study is emerging in importance.

3.1 Databases

The study approached database selection by generally comparing the amounts of entries the chosen databases hold for the topics “Artificial Intelligence” and “Passenger Information” (or “Traveler Information, in case of overlapping topics”). *Table 1* shows the comparative amounts of entries in each of the chosen databases, *ACM Digital Library*, *Google Scholar* and *TUNI Andor*. ***Association for Computing Machinery Digital Library*** (*ACM Digital Library*) is a database containing peer-reviewed sources for computer science, information technology and related fields (Association for Computing Machinery, n.d). ***Google Scholar*** is a search engine which indexes a vast number of sources (Google, n.d). ***TUNI Andor*** is a library search service for Tampere University students, which provides access to sources which might otherwise be unavailable (Tampere University, n.d).

Table 1 Search entries.

Entries since 2020 in databases	“Passenger Information” AND “Traveler Information”	“Public Transport*”	“Artificial Intelligence”
Google Scholar	6200 + 4220	184 000	1 650 000
TUNI Andor	348 + 194	18 480	515 227
ACM	78 + 151	1504	176 180

Table 1 highlights how comparatively AI is a massive topic for research from 2020-2025 in relation to Public Transport – and the even more limited entries about PI.

3.2 Inclusion and Exclusion Criteria

This work is most interested in broad literature reviews from topics such as AI, PT and PI. Selections are based on relevance to these topics in their respective chapters. When literature reviews are not available – important and relevant individual studies can be included. As a narrative review this study avoids too many duplicates in selected works unless relevant or providing new information. Other criteria include English language, and the time frame of 2020-2025. Content before the target timeframe may only be used when highly useful and no more recent work was found.

3.3 Limitations and Challenges of the Study

To name limitations for the study methods, a narrative review is prone to researcher bias, as it is less replicable and comprehensive in comparison to a systematic literature review. A narrative review also poses limited scope for statistical analysis. Further challenges are presented by search-terms - since some of the wording for the PT field does not have well established standard terminology. The study was conducted while aware of these limitations and challenges.

4. RESULTS

The results are divided into three sections in which PI, AI and the combination of PI and AI are examined. Chapter 4.1 covers dynamic PI, with its subchapter dedicated to personalization. Chapter 4.2 cover artificial intelligence, with selections made for notable areas of technology which could act as possible or current enablers for PT and PI technology. This chapter attempts to avoid studies which are dedicated to combined PT, PI and AI. Chapter 4.3 is then the chapter which covers studies which combine the topics – to find out which kinds of applications are researched and considered in the future.

4.1 Dynamic Passenger Information

This chapter examines the findings related to the state-of-the-art of Dynamic PI research. This part will particularly draw on recent works on ***Real-Time Information*** (*RTI*) and PI by Ait-Ali & Peterson (2025) and Brakewood & Watkins (2019). Works by Vredenborg et al. (2025) and van Ardenne et al. (2025) highlight the theme of personalization and PI. These were identified as the most prominent review papers from 2019-2025, which covered general PI in PT without additional focus on themes like disruptions. Brakewood & Watkins is added despite being from 2019, as there was a notable gap between 2019 and 2025 for reviews.

Passenger Information (*PI*) is considered to be firmly in the era of dynamic passenger information. Dynamic passenger information is strongly built on top of accessible vehicular real-time information – which became common through availability of GPS during the 2000s. This decreased the price of **Real-Time-Information** (*RTI*) and made it more feasible to use for PT providers. Brakewood & Watkins (2019) also point out standardization efforts for transit schedules, and later the widespread availability of smartphones. RTI in public transport is characterized by live tracking of vehicles and their states – which are then presented to users through various platforms – mainly through passenger mobile devices and station or vehicle displays. (Brakewood & Watkins, 2019)

A literature review by Ait-Ali & Peterson (2025) acts as one of the few modern general literature reviews on passenger information *effects*. The review starts by defining ***Travel Situations*** as they heavily affect which kinds of PI the passenger requires, but also overall stress on the PIS:

1. *State of traffic*. This refers to the operational state of the desired transit line. The states can classify as normal, under disruption or cancelled.
2. *Purpose of trip*. Passenger motivations such as work or leisure.
3. *Travel phase*. Pre-journey – consisting of both at home and under trip to the station, at station, onboard and after the journey.
4. *Chosen travel mode.* Heavy rail and light rail, bus, metro.
5. *Weather conditions.* Conditions such as temperature, humidity, rain may play a part in PI needs and overall system use across passengers.
6. *Time of day.* Ait-Ali & Peterson list night-time and daytime, weekend or holidays as factors for PI.

(Ait-Ali & Peterson, 2025)

Information Content Types exist in multiple varieties. Ait-Ali & Peterson divide content to *static* and *real-time* information content. While *RTI* is generally favoured for its dynamic and up-to-date properties - *static* content is still relevant for printed bus routes and schedules. Static information content for bus routes is often found more on smaller stations, which may not have RTI compatible digital signage. Adding to the divide to static and real-time information - Ait-Ali & Peterson also distinguish between *active* and *passive* information. *Active* information is information that must be actively sought out, by using apps or reading signage. *Passive* information is notified to the user through audio announcements, push-notifications or SMS. Further information content type divides were identified: *qualitative vs. quantitative*, *descriptive vs. prescriptive* and *location-specific vs. transport situation-specific*. Ait-Ali & Peterson used basic passenger information content types provided by Swedish railway provider Trafikverket – found in Appendix A. (Ait-Ali & Peterson, 2025)

Communication Channels can be divided to stationary channels and mobile personal device channels. *Stationary channels* are often managed by stakeholders such as a public provider, or multiple private providers in markets which have multiple transit providers. Types of stationary channels commonly include *signage, speakers* and *screens* which appear at stations, stops, and inside the transit vehicles. *Mobile personal devices* allow for SMS and applications/web channels. Notably - the authors state mobile apps to offer the strongest potential for personalized RTI content. (Ait-Ali & Peterson, 2025)

Passenger User Groups were not covered thoroughly by Ait-Ali & Peterson, but the work suggested that there are different ways of compiling passenger groups to be examined or targeted by PI (Ait-Ali & Peterson, 2025). The work of Vredenborg et al. (2025) address factors which can be used in passenger group division and profiling in the form of Personalization Attributes. They are covered in Chapter 4.2.1 in Table 3.

Information Flow is related to the chain of stakeholders in which the information flows from start to endpoint. *Stakeholders* may include the driver, traffic leader, informants, operators and passengers. The organisation of the relationships between these stakeholders and the facilitation of their communication affects the delivered PI. Ait-Ali & Peterson also examine ***Information Quality,*** which is described as a measurable factor in overall transport quality – with pricing, service availability and journey time. The review did not provide a solid framework to examine quality but gave examples of delay information qualities: announcement speed, precision and update interval. Relevancy and accuracy are also stated to be factors in PI quality. One way to increase relevancy and accuracy is thought to be *personalization* – where personal needs would be catered. Ait-Ali & Peterson found works dating from the 2000s which already suggested that future PI directions would be characterized as dynamic, mobile, multimodal, personable and should stay relevant to changing times and population. (Ait-Ali & Peterson, 2025)

PI Effects studies that Ait-Ali & Peterson covered in their review, identified effect distinctions such as *Long-Term vs. Short-Term* and *Individual vs. Collective* effects. The main individual PI effects on passengers were *coping with uncertainty*, *reduced waiting time* and *path choice*. The same effects apply in both normal and disrupted conditions – with disruptions placing particular focus on pre-trip information to mitigate negative impacts. ***Evaluation*** of the effects are covered, with different approaches: *Analytical Modelling**,** Surveys* – which are particularly very popular*, Simulations* - which are very rarely used for PI effect evaluation and *Cost-Benefit Analysis.* ***Assessment*** of PI effects includes a focus on quantification. *Willingness-to-Pay* is one such approach which focuses on evaluation of PI services in monetary value, which was generally suggested this to be around 5-20% of fare pricing. *Accessibility* evaluation focuses on quantification of delays in minutes or percentage. *passenger satisfaction* and *social welfare* are also mentioned but not provided reasoning as how to quantify them. (Ait-Ali & Peterson, 2025)

Overall, while the work of Ait-Ali & Peterson, provide important research to this work, as one of the few recent PI literature reviews with a broad scope. While this is useful as background - the review has flaws. As examples – the work addresses some topics with only short paragraphs, does not provide a framework for PI effects and commonly relies on old literature – predating the RTI era. Another literature review focused on the effects of ***Real-Time Information*** (*RTI*) on passengers by Brakewood & Watkins

(2019). Although older, it adds to the understanding of PI effects. As PI in the last two decades has been heavily interlinked with RTI, findings in this research area are applicable to PI overall. RTI PI was found to have many positive effects to passengers in literature from 2004-2019 – some of which echo the work of Ait-Ali & Peterson:

1. *Decrease in wait times, and perceived wait times.* This is because users can more accurately predict the movement of their vehicle – accessing the information on their own mobile devices. This also reduced the feelings of uncertainty.
2. *Decrease in total travel time*. RTI allows users to make more adaptive choices about their journey, and this can result in lower total travel times. Studies, which relied on simulation modelling, were not unanimous on the magnitude of reduction – they generally consider RTI beneficial.
3. *Amounts of PT use.* Studies covering survey data contained mixed results. When asked whether respondents whether RTI would increase their PT use – the responses ranged from 30% to 70%. Contradictory – however, multiple USA based studies which used econometric analysis to follow ridership, suggested only an increase of <2% use in PT with newly introduced RTI services.
4. *Satisfaction on PT*. Surveys covering user satisfaction generally noted a positive reaction to new RTI systems, but not a significant increase in general service rating.
5. *Perceived security*. Featured surveys had unclear results. Some surveys had small positive results only for daytime, and some for only night-time. Brakewood & Watkins suggest that other context might play a part in the results and call for studies including factors such as crime-rates in the survey area.

(Brakewood & Watkins, 2019) As noted by Ait-Ali, many PI evaluations which Brakewood & Watkins featured rely on survey results. RTI seems to positively influence survey ratings for wait times, travel times and feelings of uncertainty. However, the results for PT use, satisfaction and security held unclear or contradictory results – indicating that surveys may not capture the full response contexts or the responses are not accurate.

Complementing these findings - a conference paper by Baldauf & Tomitsch (2020) analysed the implementation of Swiss train station PI and compared their findings with other then concurrent papers. RTI routing services in mobile apps are found to be a standard service due to the previously discussed reasons, but Baldauf & Tomitsch consider the *presentation* of that information not being entirely solved – particularly for multimodal transport. Novel display methods and enhanced real-time guidance were suggested as potential research opportunities. In a sense – those ideas could also increasingly *blend the phases of travel* in the context of ticketing system use and guidance. *Personalization* is suggested to be another direction – where services could be capable of learning from past journeys – and reacting to variables such as the weather or traffic. The last PI research direction suggested was increasing *trust in public transport*. This is because the increasing digitalization, automation and personalization of PI services will also require efforts into factoring user trust. (Baldauf & Tomitsch, 2020)

Sensing technologies are increasingly being considered and used in public transport. Passengers are often detected and counted at vehicle entrances using ***IoT*** (*Internet of Things)* sensing technology, and crowd monitoring by CCTV is often present inside vehicles and at stations. Darsena et al. (2023) consider potential for new applications and enhancing existing ones. Security, crowd analysis, crowd prediction and crowd management are some considered topics. Suggestion by Darsena et al. include ways to *display crowding data and vehicle occupancy to users*, possible reservation ticketing systems for crowded vehicles, *more informed route suggestions* and long-term crowd analysis. Hindering factors were related to security, privacy and *IoT computation restrictions*, particularly when utilizing *edge computing* for complex security and privacy algorithms. (Darsena et al., 2023)

4.1.1 Personalization

Personalization was identified as an emerging area of interest in the PI reviews. This research direction has recently been synthesized by works such as Vredenborg et al. (2025) and van Ardenne et al. (2025). Personalization, according to Vredenborg, is a rising solution to simplify and tailor PI to passengers - to answer increasing amounts of available PI. The increased amounts of information may make ***Passenger Information Systems*** (*PIS*) more complex and difficult as user may access any kind of desired route information with filtering. Personalization is also a modern trend in other fields such as e-commerce, streaming and gaming – where it has been shown to enhance user experience, conversion rates and have positive impact on sales. (Vredenborg et al., 2025) Vredenborg et al. approach PI Personalization through a framework with three different dimensions - ***Objects of Personalization***, ***Personalized Attributes*** and ***Evaluation of Personalization***. These categorizations were synthesised from PI studies and help to understand the area more concretely.

The ***Object of Personalization*** refers to the targeted elements which are to be altered. Four personalization objects are classified into *Recommendations*, *Interactions*, *Information* and *Vehicle Settings*. **Recommendations** are decision-assisting elements in passenger navigation delivered through PIS. Recommendations are often divided to either route optimization or classification. In *route optimization*, search-based algorithms such as Dijkstra's algorithm or ML techniques such as reinforcement learning can be used to find shortest paths from route graphs. *Route classification* is then the Recommendation technique of selecting the fastest route from a pre-generated set, using ML techniques such as *collaborative filtering*.

Interactions represent PI delivery techniques as personalization objects. In more concrete ways, this can refer to the *delivery channels* or *methods*, *User Interfaces* (*UI*) and *modalities*. For example, enabling push-notifications or alerts can shift the interaction methods from passive to active. Text-to-Speech, visual content, or text-based information can represent personalized interaction styles that enhance accessibility. UI changes such as larger elements and their visibilities and layout can also represent personalized interactions. Content delivery can also be changed in terms of wording, tone or style. (Ait-Ali & Peterson, 2025) Vredenborg et al. highlight also the possibility of *Large Language Models* (*LLM*) in content delivery personalization, as they could enable conversational guidance or act as a use modality.

Information is a vital part of PI and can also be a personalization object. This can include any kind of PI - such as journey duration, routing alternatives and disruption details. In *navigation*, RTI PI may be utilized to find nearest vehicles and personalize guidance to help the passengers adjust routes, hurry or decide to wait for the next vehicle. The topic of *micro-navigation* is also featured – referring to personalized suggestions on where to navigate onboard the vehicle. As *disruptions* are events, which increase user-reliance on pre-trip information, personalized disruption messages may help users adapt more efficiently. **Vehicle Settings** are the fourth personalization object brought up, which would mainly refer to comfort attributes such as seat positions or climate control – but this area has not been widely explored in PT contexts, despite being popular for private vehicles, and current PT vehicles do not accommodate this personalization object. (Vredenborg et al., 2025)

Objects of Personalization			
Recommendations	Interactions	Information	Vehicle Settings
Route Optimization. Route Classification.	Modalities. User Interfaces. Content Delivery and methods.	Journey Duration. Vehicles. Navigation/Stops. Disruptions. Facilities. Fares.	Physical Comfort Attributes. Climate.

Table 2 Objects of Personalization. Adapted from Vredenborg et al. (2025)

The ***Personalization Attributes*** explained by Vredenborg et al. are the contexts and attributes in personalization and effect the results. They are divided to *Context-related* attributes ("physical, social, temporal and spatial") and *User-related attributes* ("preferences, characteristics and ownership"). **Physical** contexts relate to the traits and elements of the physical space, climate and facilities. **Social** contexts refer to the presence and behaviour of other passengers. **Temporal** context attributes are related to time factors such as time, timing, weekday, events and holidays. **Spatial** contexts refer to the current location of the passenger, from various precision perspectives such as location from GPS perspective, railway network or inside the vehicle. **Preference** attributes refer to implicit and explicit preferences and common user patterns – such as modalities, language, route preference, time of activities. **Characteristic** attributes refer to demographics, knowledge and status for special needs. **Ownership** attributes refer to tickets and subscriptions, vehicles and items that the passengers possess

Personalization Attributes						
Physical (Context)	Social (Context)	Temporal (Context)	Spatial (Con-text)	Preferences (Attribute)	Characteristics (Attribute)	Ownership (Attribute)
Infrastructure. Climate. Facilities.	Presence of unknown passengers. Staff. Company.	Current time. Current date. Journey timing. Weekday. Time of day. Events. Holidays. Seasons.	GPS lo-cation. Location in coun-try. Location in city. Location in vehicle.	Communication Mode. Interchanges Routes Time of travel. Distances	Demographics. Knowledge level. Special needs status.	Tickets. Subscriptions. Luggage. Items. Vehicles.

Table 3 Personalization Attributes. Adapted from Vredenburg et al. (2025)

The third personalization dimension is ***Evaluation of Personalization***, which Vredenborg et al. introduce categories of "*Evaluation Type*, *Evaluation Metrics* and *Evaluation Layers*". **Evaluation Types** can be user studies, stakeholder studies, expert analysis, use simulations or data studies. **User Studies** extract data from either current or potential user-bases, either using interviews or surveys, for qualitative and quantitative data types respectively. Adding to user studies, observations can examine daily use of transport and user tests examine the completion of a limited task or set of tasks. Field tests are observations which implement a new system to daily routines. **Stakeholder Studies** study people who are related to PI and PIS but are not the passengers or other end-users. Stakeholder studies are commonly qualitative, and are gathered using either interviews, observations, workshops or supplementary feedback during a project. **Expert Analysis** is conducted by domain experts to benchmark a system, study it in field conditions in a participant study or as system engineer. **Simulations** can be utilized in personalization evaluation to study user behaviour such as route selection and user preferences. This is done through randomly generated distributions based on real data. The simulated data may implement simulated vehicle GPS data, schedule data, *Automatic Fare Collection* (*AFC*) data, road and station status data and other miscellaneous data. Simulations were the most uncommon type of study according to Vredenborg et al. These evaluation types produce evaluation metrics, which are the factors being studied. **Evaluation Metrics** listed are "*Usability*, *Usefulness*, *Accuracy*, *Technical Performance*, *Model/Algorithm Efficiency*". **Evaluation Layers** are the third identified personalization evaluation factor. It is the process of representing the personalization processes into layers – and seek to analyse them separately. The different layers were *Input collection*, *Data Interpretation*, *World Model* which refers to system model validation, *Adaptation Evaluation*, *Adaptation Presentation*.

Evaluation of Personalization		
Evaluation Types	Evaluation Metrics	Evaluation Layers
Interviews (User).	Requirements	Input Collection
Surveys (User).	Usability	Data Interpretation
Observation (User).	Usefulness	World Model
User tests (User).	Accuracy	Adaptation Evaluation
Field tests (User).	Performance	Adaptation Presentation
Workshops (Stakeholder).	Model/Algorithm Efficiency	
Interviews (Stakeholder).		
Observation (Stakeholder).		
Field Tests (Stakeholder).		
Field tests (Expert).		
System Engineering (Expert).		
Data Study (Expert).		
Simulation (Expert).		

Table 4 Evaluation of Personalization. Adapted from Vredenborg et al. (2025)

The work by Vredenborg et al. synthesize personalization research into a framework with personalization objects as targets, attributes which can affect personalization decisions, and categorizations for personalization evaluation. Vredenborg et al. determine that the personalization trend in academic research has been growing for a decade. Some current PT providers are deploying personalization efforts – but many new possibilities are still in research phases. Future research directions for less explored topics that are pointed out are as follows:

1. Offering explainable personalization and recommendations.
2. Research of personalized content presentation styles.

3. Human-centered research on how personalized content influences decisions and reception.
4. Efforts to enhance and explore privacy, security and user control with personalized systems.
5. Research on accessibility and personalization.
6. Transparency efforts in study reporting practices.

Another notable literature review on PI personalization by van Ardenne et al. (2025) adds to the framework created by Vredenborg et al. by proposing a personalization taxonomy which determines 5 levels for how personalized a system is. The levels are based on system functionality and autonomy. This personalization framework also plays a part in personalization explainability and the creation of systematic definitions for PI – which have been previously less defined. Definitions for higher levels of personalization may also support AI integration design for PIS. The proposed framework consists of *three* functionalities with differing levels of *user* and *system autonomy* - ***Situational Awareness***, ***Target Identification*** and ***Trigger of Provision***. In the lowest levels – the system fully relies on user control, and on higher levels the system is highly aware of relevant contexts, can interpret the users and their needs – and provide information proactively. PI systems may place high in one category, and low in others. (van Ardenne et al., 2025) The resulting framework may be found in *Table 5*.

Situational Awareness is a functionality category which is focused on the system's ability to access status information about the vehicle, the transport network and passengers. A system with *No Awareness* is incapable of providing RTI PI and relies on static PI. *Historically Aware* systems can label high-level transport network data such as crowding and deliver this to all users. In these lower levels of situational awareness, the system is not aware of multiple contexts. On higher level functionality categories - *Supply Aware* systems access RTI data from multiple vehicles in the system and can provide route recommendations. *Demand Aware* systems also integrate personalization attributes such as the location of the passenger, preferences, behaviour and the trip phase. The focus is on providing personalized guidance to the passengers - rather than only informing about transit options. *Ecosystem Awareness* is the final level of system awareness, where environmental and societal contexts and attributes added. Examples of such factors are weather conditions, road congestion in bus routes or crowding due to a public event. (van Ardenne et al., 2025)

Target Identification is a functionality category where the level of user understanding is measured. This is achieved through user profile creation and management. Systems with *No Target* do not identify or profile individual PI needs and provides identical information throughout the transport network – often in static form. Systems in the *Geo-temporal Constraints* category divide users according to their location and route timing – often represented by systems where users can input locations and access route recommendations. This is not yet long-term user profiling. *Stated User Profiles* is a category which extends this to user profiles in which users may input preferences, mark favourite routes and stations and access this information later. Systems with *Revealed User Preferences* move to automated data collection – where user actions are stored including smart card use, ticket system use, and travel patterns. They can be used to provide recommendations and assign profile groups to users. This process is stated to benefits from utilizing machine learning practices for processing. The final target identification level in van Ardenne et al. are systems using *Projected User Preferences*. In this level, each passenger has a unique profile, based on a system interpretation of what the passenger is like. The system can personalize new route suggestions based on that. It differs from the last stage in that it would have predictive capability rather than only classification based on history. (van Ardenne et al., 2025)

The final functionality category is the ***Trigger of Provision***, which is focused on factors which trigger information delivery, and their timing. The focus is also in how proactive the system is. ***User Trigger*** refers to a system in which the user actively searched for PI. ***Supply Trigger*** refers to when the system delivers PI when an event triggers information delivery, such as disruption alerts. In this level, the user still needs to engage with the system to view the update. ***System Trigger*** is then the final level – where the system delivers proactive PI delivery. For example, this can mean push notifications for predicted delays - without the user actively searching for the information. (van Ardenne et al., 2025)

The resulting framework then divides the overall personalization levels to 0-4, depending on the personalization levels of the three functionality categories. The review goes into more detail, but this work is more focused on using this framework to define and categorize PI/PIS personalization levels. For future research agenda - van Ardenne et al. suggest that data storage needed for higher levels of personalization would need to be addressed by research. The database solution would have to be capable of functioning with fragmented PIS services. Data quality is also suggested to play a part in successful level 3 and 4 personalization – and thus advanced data collection methods are called for. Ethical questions are also introduced, such as user acceptance, bias, inequality in technology adoption, overall accessibility and privacy. A notion for mitigation is suggested to give the user the agency to opt in or out of data collection and advanced PI personalization. In a counterpoint – this could diminish the effects of the personalization efforts. (van Ardenne et al., 2025)

Personalization level	Situational Awareness	Target Identification	Provision Trigger
0. No personalization (User)	No Awareness	No Target	User Trigger
1. Customization (User)	Historical Awareness	Geo-Temporal Constraints	User Trigger
2. Partial Personalization (User)	Supply Awareness	Stated User Profile	User Trigger
3. Conditional Personalization (System)	Demand Awareness	Revealed User Profile	Supply Trigger
4. Full Personalization (System)	Ecosystem Awareness	Projected User Profile	System Trigger

Table 5 Levels of Personalization and example table. Directly adapted from van Ardenne et al. (2025)

This chapter utilized recent PI literature reviews and examined how the basics of PI is formed. The effects of PI and RTI PI were examined – extending into personalization, which is emerging as an important part of modern PI research. Vredenborg et al. propose a framework for structuring and defining personalization in PI. The framework is supported by van Ardenne et al. and their personalization level -framework. The understanding of the state of PI research is ready to be complemented by AI research, and to answer ***RQ1*** and ***RQ2***. The following interests and needs were found related to ***RQ1***:

1. *Personalization.* Personalization is emerging as a major theme for academic PI research.
2. *Data storage solutions*. Fragmented subsystems and data storage for PIS is a challenge for new developing new systems. Particularly proposed personalization systems would require centralized data solutions to work with the systems.
3. *Data collection solutions*. Current transport mobile applications and smart cards do not collect sufficiently sophisticated data to support the proposed personalization ideas.

4. *Content presentation.* Works such as Baldauf & Tomitsch (2020) suggest that PI content presentation may have room for new innovations, in digital signage and mobile applications.
5. *Sensing Technology*. IoT solutions for crowd sensing could be used to build new services related to crowd analysis and management. These technologies could also align with micro-navigation – to guide passengers inside vehicles.
6. *Ethical and legal questions.* User privacy, control, accessibility and legal matters are of vital importance when considering tools which utilize personal data.

4.2 Artificial Intelligence

Since 2020, there have been substantial improvements in many fields of AI. The Stanford HAI AI Index Report by Nestor Maslej et al. (2025) provides a large report regarding AI, including a variety of different topics such as research trends, economic trends and technical benchmarks. As an example of benchmarking efforts - LLM performances in competition-level mathematics against the MATH-dataset show large performance increases within 5 years, exceeding a human baseline in 2024. Strong performances were also found in other benchmarks such as against MMMU (Multimodal Understanding and Reasoning) which tested AI systems against expert-level multidisciplinary questions. The top scoring model for this benchmark, o1 – scored 78,2% vs a human baseline of 82,6%. Advancements in *Machine Learning* (*ML*), and its subfield *Deep Learning* (*DL*) have been major enablers for these technologies. (Nestor Maslej et al., 2025) This chapter will introduce select important ML topics, starting from classical ML models and moving to modern DL techniques.

Machine Learning is a key subfield of AI, characterized by the creation of algorithms which attempt to represent a set of data. Trained models may make predictions and decisions based on the training data. ML can be applied to a wide range of suitable datatypes, and thus ML extends into other AI domains such as Computer Vision and Natural Language Processing. ML models are most often divided based on their training methods – supervised, unsupervised, semi-supervised and reinforcement learning are the common general approaches. In ***supervised learning*** the focus is on creating "predictive" models. The computer receives labelled input-output data to create a predictive model which allows it to make predictions about new but similar data. In ***unsupervised learning***, the focus is on "exploring" models. The computer receives unlabelled data, where clustering, association mining and dimensionality reduction are techniques to group similar datapoints together and to simplify data by removing outliers. Emerging patterns are possible to be detected by using these unsupervised learning methods. ***Semi-supervised learning*** (*S-SL*) is a variant where some data-points are labelled, and others not. In ***Reinforcement Learning*** (*RL*) is focused on the creation of agents which receive feedback on their performance – and use a trial-and-error approach. *Classical ML* techniques embody these approaches. Examples of unsupervised learning model types are linear regression and logistic regression, support vector machines, K-nearest neighbours, naïve bayes, decision trees and artificial neural networks. Unsupervised learning approach examples are principal component analysis and clustering techniques. *Ensemble models* are combinations of other individual models. As an example - *random forests* utilizes multiple decision trees, classification and averaging for regression tasks. (Maleki et al., 2020)

Artificial Neural Networks (*ANN*) are particularly vital machine learning algorithms which are formed of interconnected nodes that are organized in layers. In a simple ANN there are three layers. The **Input layer** receives the initial input data. This layer sends it to the subsequent **Hidden layers** for computations, where activation functions process data to extract features or patterns. Finally, in the **Output Layer**, the final classifications or other outputs are prepared. Simple ANN:s can have just one hidden layer, but ***Deep Neural Networks*** (*DNN*) can have any number of hidden layers. The number of hidden layers is a defining factor for whether an DNN is classified as shallow or deep learning. Different types of DNN exist, with different focus points such as learning abstract and complex patterns or hierarchies, and extending to enhance image recognition, and other similar machine learning tasks. (Mienye & Swart, 2024b) With many different types of DNN, Mienye & Swart consider ***Convolutional Neural Networks*** (*CNN*) and ***Recurrent Neural Networks*** (*RNN*) to have been foundationally important to advance in fields such as Computer Vision and Natural Language Processing respectively.

CNN is notable for computer vision tasks. It is defined by use of convolutional layers which are useful for capturing edges, textures and shapes. Typically, the CNN architecture consists of convolutional layers, pooling layers and fully connected layers at the end. The integration of pooling layers helps reduce overfitting to a specific dataset and simplify the computations while preserving the spatial information. The final fully connected layer culminate into a prediction. (Mienye & Swart, 2024b)

Following CNN use for vision and recognition tasks, **RNN** are useful for sequential data processing, notably including language tasks such as NLP and machine translation. RNN can capture context with temporal dependency from previous time steps and could theoretically learn dependencies from long sequence data. However, the base

RNN architecture faced issues of vanishing and "exploding" gradients, causing them to be unreliable for long sequences. In such cases, an error signal sent backwards in the network would disappear or cause other issues. This was a problem with RNN until solutions were formed, such as ***Long Short-Term Memory Network*** (*LSTM*) - in which a specialized memory cell addresses the vanishing gradient problem. This is done utilizing a gate system to control cell the state updates. This allowed more consistent dependencies for long data inputs. Variants of LSTM exist, such as *Gated Recurrent Units* (*GRU*), which simplify the LSTM gate system, and aim to increase computation efficiency. (Mienye et al., 2024a)

Autoencoders are a type of DNN commonly used for unsupervised learning tasks. They work by encoding an input vector to a simplified latent space form which captures essential features and then reconstructing it based on their simplified latent space form. This encoding process reduces the dimensionality of the data, and can be, for example, used by flagging reconstruction errors to find anomalies. Autoencoders and their large variety of subtypes can also be used for a wide variety of other tasks, including feature extraction, denoising data, data compression, data denoising and synthetic data generation. One notable variant for autoencoders is *Variational Autoencoder* (*VAE*), which is utilized in generative AI. (Berahmand et al., 2024)

Transformer, which was created by Vaswani et al. (2017) presents as a notable DL Architecture which have become very important for natural language processing, computer vision, audio processing and other miscellaneous tasks. Transformers utilize an attention mechanism to detect global dependencies between input and output and sought to improve the then popular RNN model issues with long inputs. The transformer architecture has many strong features, such as efficient parallelizable training, scaling performance and architectural flexibility to fit many purposes. Pre-trained transformer models can also often be fine-tuned, rather than training from scratch. (Lin et al., 2022) ***Graph Neural Networks*** (*GNN*) are a beneficial way to process graph data. They utilize a technique to update each node representation based on adjacent node features, which allows for the capturing of complex relationships. Node classification, prediction of links, classifying graphs and nodes are useful for naturally graph-structured data such as "social networks, chemical molecules and communication networks.". (Mienye & Swart, 2024b)

Generative Adversarial Network (*GAN*) is a DL architecture which is useful for *generative AI* task such as producing realistic synthetic data – including images, language, video and audio. In its most common variant, it is formed of a *Generator* and a *Discriminator*. The Generator is responsible for generating noisy fake data, and a Discriminator trained by using real images determines whether the output is classified as fake. Then the Generator proceeds to repeat generations until the Discriminator components cannot determine the outputs as fake. A model capable of synthetic image material is then formed. (Sengar et al., 2025)

You Only Look Once (*YOLO*) are a family of frameworks for real-time object detection. This type of technology is highly useful for many kinds of fields such as the medical field, agriculture, security systems, intelligent transportation systems, traffic management, autonomous vehicles. Previous approaches such as R-CNN, advanced the field of real-time object detection. YOLO improved these approaches by accomplishing detections with a single network pass, rather than the previous approaches which first had to detect object regions and later classify them. YOLO attempts to balance speed and accuracy due to the focus on real-time processing, and is not universally the best solution for non-real time accuracy based tasks. (Terven et al., 2023)

Deep learning is overall a field which enables a huge number of technological solutions across a wide range of domains. However, they suffer from fundamental challenges. One large general issue is their **explainability** – as DL models are very complex and intricate – and their results are often difficult to explain. Thus, their trustworthiness is under question, particularly for systems which utilize DL for decision making. *Scientific visualization methods* such as heatmaps and saliency maps are one way to address explainability. *Model distillation* attempts to "*approximate a complex model by fitting a simpler model using the training set*". These are some attempts at providing explainable outputs. **Model bias** is also another type of DL challenge – as the models are based on training datasets – which may skew the results based on things that are over- or underrepresented. **Technical issues** are also constraints for DL methods. The complex models may be very memory and computation intensive – which restricts their application possibilities in environments such as IoT edge devices. The creation and refinement of small and performant DL models is an enabler for new DL applications. Helpful techniques are thought to be pruning extra parameters, quantization techniques to reduce memory size and knowledge distillation to simpler models. **Security-related issues** are also a factor – as manipulating input data may lead to unintended behaviour. **Data-related issues** also limit DL models, as valid training data for something may not be accessible or is unlabelled. Models may also be **Overfit** for a certain dataset, so that the model does not generalize the information beyond it. (Mohd Noor & Ige, 2025)

Although DL technologies suffer from certain challenges – their overall viability and strengths are not erased. The preceding overview of ML and DL featured some foundational architectures which play part in AI subdomains such as computer vision, natural language processing and generative AI. The next subchapters will feature overviews of select notable areas of AI.

4.2.1 Computer Vision

Computer Vision (*CV*) is a subdomain of AI interested in enabling machine perception, interpretation and analysis of image and video content. As CV is a large research area – this subchapter focuses on core and otherwise notable topics. Many of the previously described DL architectures apply for CV, such as CNN-based architectures for most central CV tasks such as recognition, GAN which can generate new data, RNN-based architectures for sequential video data, YOLO for real-time object detection and Transformer for various purposes as a versatile architecture. CV is applicable to fields which want to analyse image or video content. The medical field, for example, has been a major contributor to CV. Feature detection, recognition, 3D-modeling and image segmentation are very useful for the medical field, for detecting illness, diagnosis and even supporting robotic surgery. Biology is another field, where microscopic imagery can be analysed. CV is also used in agriculture for monitoring tasks or part of machinery, and in autonomous vehicles. (Cernadas, 2024) Central CV tasks can be broadly divided to *classification, segmentation* and *object detection* (Mahadevkar et al., 2022). These tasks can be extended to more complex topics such as human action recognition.

Image Classification is a fundamental CV area which is focused on differentiating object classes from image properties. Very basic classification tasks can include differentiating cars from flowers. The most popular technique for classification is CNN, which uses convolution kernels for feature extraction. (Zhao et al., 2024) Maurício et al. (2023) synthesized in their literature review a notion that ***Vision Transformer*** (*ViT*) is gaining an edge over CNN in some areas. The authors discovered ViT better for images with noise, due to the self-attention mechanism making the image information accessible from high and low layers. ViT was also found to be less computationally intensive than CNN. CNN were found to generalize better when small datasets are available – but ViT were able to function overall with less images. Despite this performance difference - Maurício et al. found all reviewed literature to consider combining ViT with CNN to perform the best. (Maurício et al., 2023)

Image Segmentation is focused on partitioning images to meaningful regions. It provides a foundation for pattern recognition, image categorizations and image understanding. Accurate segmentation plays a strong success factor for these tasks but is also challenging due to the possibilities of working with complicated backgrounds, possible poor image qualities and inconsistent image radiance. Conventional methods for segmentation are *thresholding*, *region and edge-based segmentation*. *Thresholding based segmentation* partitions images by comparing their pixel sets with a threshold – creating a background and a foreground. *Region based segmentation* focuses on detecting interiors of possible image entities and find their edges this way. *Edge based segmentation* is then the opposite, where initial edges are detected, and occluded by contouring algorithms. It assumes that image entities transform at edges. Brar et al. (2025) consider these conventional methods have some advantages over DL methods, as they are more lightweight, explainable and do not overly rely on a dataset for success. However, DL has been a strong point of interest in this field, as its strengths correlate with the processing of pixel data, and finding complex patterns. CNN, RNN, GNN are all technologies which have been used for image segmentations, including their vast number of variants and technologies based on them. Brar et al. anticipate future image segmentation research to revolve around iterations of DL techniques. (Brar et al., 2025)

Object Detection is focused on locating and classifying object instances within images. It builds upon image classification and image segmentation but requires also the prediction of object positions. Amjoud & Amrjouch (2023) provides a survey on deep learning for object detection. Ordinary object detection models consist of the *backbone*, *neck* and *head* modules*.* The **Backbone** module is a *CNN or ViT* network providing foundational feature extraction. VGGNets, ResNet and EfficientNets are pre-trained classification network backbones. **The Neck** module is a network over the backbone, which combines feature maps to support finding objects in different scales. Neck network examples are FPN, NAS-FPN, ASFF, PAN and BiFPN. **The Head** detection module is responsible for either *dense prediction* or *sparse prediction*, in which the final predictions are provided. Dense (single-shot) prediction network examples are RetinaNet, YOLO, SSD, CornerNet and FCOS. Sparse (two-stage) prediction network examples are Faster R-CNN and Mask R-CNN. Modern single-shot detectors were found to rival two-stage detectors accuracy. Transformer-based detectors such as Swin-L and Swin V2 were also highlighted as emerging performers. The research field of object detection moves rapidly. Amjoud & Amrouch considered future directions: optimizing the balance of speed and accuracy, tiny object detection, 3D object detection, multi-modal object detection and few-shot learning. (Amjoud & Amrouch, 2023)

Classification, segmentation and object detection are basic problems within CV, but they can be extended to various purposes. Human action recognition, anomaly detection are two examples.

Anomaly Detection (*AD*) for CV is focused on patterns and events which deviate from normal. Anomalies are often divided to *point anomalies*, which are individual sample irregularities, *contextual anomalies* which are abnormal due to the context and *collective anomalies* which are anomalous when compared to the entire dataset. DL brings additional distinctions of *sensory anomalies*, which are irregular textures or edges, and *semantic anomalies* which are samples where the sample belongs to a differing class. DL advancements are very relevant in this field, with them mostly replacing more classical ML models for AD such as Kernel Density Estimation and One-Class Support Vector Machines. AD systems are deployed in domains such as medicine, industry, infrastructure, financial security and cybersecurity. Some notable AD models are CutPaste, CSI and Spatial CL.

Central AD issues arise from the nature of the problem – datasets contain largely normal data rather than anomalies. Even when labelled anomaly data is available, new anomalies pose an issue. Due to these reasons, unsupervised and semi-supervised learning are the only practical options, which are not as performant as supervised learning algorithms. Recently, **Self-Supervised Learning** (*SSL*) is an emerging training technique, which better utilizes unlabelled data. In SSL, the models are trained through unrelated proxy tasks which are used to teach a generalized representation. This has enabled stronger models even now – but it is also an important topic for AD research in the future. (Hojjati et al., 2024)

Human Action Recognition (*HAR*) is a field strongly correlated with CV is interested in recognizing and classifying human actions. Sensors are sometimes integrated to HAR systems or relied on entirely – but many solutions are interested in video or image data. Human-Computer Interactions (*HCI*), competitive sports and surveillance are some areas where HAR is used. For HCI, HAR can be used to implement new modalities, such as smart home system gestures. For surveillance, HAR can be used to create autonomous real-time surveillance systems, which may provide early warnings, or prevent false alarms. Both ML and DL are popular in HAR systems. For DL this includes CNN, GCN but also LSTM, as the data for HAR is most often sequential. Some limitations and challenges for the field include data pre-processing and collection, dataset modelling and configuration, limited open-access tools, and computational intensity. (Karim et al., 2024)

4.2.2 General Machine Learning

This subchapter features areas of research, which provide further examples of the different applications for AI. The featured topics are *automatic speech recognition, time series data processing, intrusion detection in cybersecurity*. The selected topics highlight the variety of data types which ML-based techniques can be applied to.

A work by Ahlawat et al. (2025) discusses **Automatic Speech Recognition** (*ASR*) from a DL perspective. The core idea of ASR is to convert acoustic signals into words and characters. Classic approaches utilize algorithms such as MFCC for feature extraction and HMM and GMM for classification. Recently classic approaches have been outperformed in benchmarks by DL based approaches. Measuring ASR performances is done using the Word Error Rate (*WER*). WER benchmarks compare the model transcriptions to human ones - and in this way an error rate is extracted. The most successful models in the English language that Ahlawat et al (2025) inspected could achieve a WER of 1.4% against a clean LibriSpeech dataset. Contrasting this, rare language benchmarks – particularly those without wide data resources available held 10-20% WER. Contributing factors to differences in those performances are related to a lack of model training data for rare languages. Benchmark datasets also vary in difficulty, which leads to differing scores even for the same language. Factors such as tone, emotion and accent play into word recognizability. Thus larger contexts affects the results. (Ahlawat et al., 2025) While classic ASR is already a mature technology deployed in real world-contexts – DL is pushing the boundaries further. Ahlawat et al. (2025) state that transformer and its audio variants such as conformers have been leading recent improvements – improving many WER scores below 5% against many benchmarks. Despite the improvements, Ahlawat et al. identified and collected several challenges from ASR research. These include processing of sound intensity variations, addressing limited datasets for rare languages, identification of different speakers in a conversation and mixing of similar sounding word and performance. (Ahlawat et al., 2025)

Time Series problems are related to prediction of future values based on historical and sequential observations. They are important in strategic planning, resource management and decision making. Concrete time series problems which Hall & Rasheed (2025) present in their survey include electricity consumption forecasting, environmental quality assessment, meteorological predictions, medical diagnostics, traffic flow prediction, and financial domains. Their survey considered **Tree-Based Machine Learning** (*TBML*) models for transportation, urban mobility, anomaly detection. DL models were favoured for environmental and meteorological predictions and health monitoring.

The most performant technologies were found to be RNN-based models such as LSTM for DL models, and XGBoost, LightGBM and CatBoost for TBML based models. Despite transformer-based models being used for NLP tasks, which have common ground with time series problems - Hall & Rasheed consider transformer to be underexplored in the space. The limited number of studies which used them showed promising performances – utilizing the transformer capability in long-range dependency modelling. Notably, Hall & Rasheed concluded that DL models have not replaced TBML models currently. TBML models outperformed DL in tabular, noisy, missing and categorical data – while DL was stronger in unstructured data with spatio-temporal patterns. (Hall & Rasheed, 2025)

A review by Mohamed (2025) covered AI and ML in the domain of ***cybersecurity***. In cybersecurity - AI and ML practices are found particularly useful for their capability in real-time pattern recognition, anomaly detection, classification, and responses. Cybersecurity systems produce vast amounts of data through traffic logs, behavioural data, and external threat intelligence feeds, which can be leveraged by using AI. Mohamed (2025) identified five key areas where AI and ML are used in cybersecurity – intrusion detection, malware classification, behavioural analysis, threat intelligence, and automated response. **Intrusion Detection Systems** (*IDS*) are one vital part of cybersecurity architectures, where they monitor network traffic for potential threats and provide alerts or take responses. ML-based IDS can be used to learn normal traffic behaviour and avoid the issues of previous rule-based IDS - which struggled with detecting previously unseen threats and unconventional attacks. IDS also provide **Malware Classification** and **Behavioural Analysis** capability, where user patterns are tracked and classified – also using broader context such as devices and application use. **Threat Intelligence** refers to understanding new attack behaviours and methodologies and is often provided through *threat intelligence feeds,* where information is collected from multiple security parties. One emerging way to enhance threat intelligence is through Natural Language Processing (*NLP*) tools, which can use topic modelling and clustering to prioritize the vast amounts of threat data. Finally, ML improves IDS by supporting **Automated Responses** where the real-time monitoring can extend to refining detection thresholds and deploying countermeasures such as blocking addresses or providing threat alerts. ML has overall significantly benefited cybersecurity, and this trend is expected to continue. (Mohamed, 2025).

This section will cover ***Recommender Systems*** (*RS*), which are largely considered a separate area, along with CV and NLP and focused on presenting relevant and personalized information. *RS* answer the issue of with increasing data volumes on the internet

– by helping make intelligent predictions about user needs. Correlating with increased internet content data volumes and user-based data, ML is very useful for the domain. RS are generally categorized into Collaborative Filtering, Content-Based RS and Hybrid Systems as covered in a survey by Gheewala et al. (2025). The systems differ in approach to recommendation, with *Collaborative Filtering* compiling user past decisions and comparing it to other users and *Content Based RS* focusing on extracting content features and comparing them to user behaviour. *Hybrid systems* are then systems which combine two or more approaches to RS, such as different ML or DL models. A fourth type, *Knowledge graph-based RS* represents users, information items and features as graphs. (Gheewala et al., 2025)

DL is currently strong point of interest in RS research as DL is thought to outperform classical approaches in RS with the following capabilities:

1. Extracting hidden item and user features.
2. Understanding nonlinear user-item connections.
3. Next item prediction in sequential RS.
4. Capability to integrate resources from external domains.

Two approaches, *Rating Prediction* and *Top-N-ranking* are used for DL implementations. Rating Prediction is focused on how users would rate an item based on historical user data and recommending the best rated items. Top-N-ranking is focused on providing top N number of items the user would find useful or interesting. Traditional DL architectures which can be used for building RS include RBM, DBN, autoencoders, GAN, MLP, CNN, RNN. Some newer approaches underlined by Gheewala et al. include DAN, LLM, GNN and Hybrid models. Deep Attention Networks (*DAN*) use an attention mechanism to differentiate item weights which can enhance accuracy. Large Language Models are also being integrated for RS purposes and are thought to help with nuance and context. Graph Neural Networks (*GNN*) are utilized by modelling users and items as connected nodes - where their connections and interactions can be analysed for enhanced accuracy. (Gheewala et al., 2025)

Some current key application domains for RS were extracted in a review by Garapati & Chakraborty (2025). They include streaming services, social network services, e-commerce services, healthcare and education. Particularly health care and education are rising domains for RS application. In healthcare, they are used in conjunction with patient data and can be used in personalizing treatments or diet suggestions. Educations is another domain in which RS has been increasingly relevant – mirroring a transition to an increasingly digitized learning landscape. Learning materials can be personalized using RS to improve learning outcomes. For future domain applications for RS, Garapati & Chakraborty point their potential for smart city usage. Examples that they list include public transport route optimization, energy and waste management, urban planning, event discovery, tourism and smart buildings. The authors underline that before smart city use, RS privacy and bias concerns have to be taken into account. (Garapati & Chakraborty, 2025) To summarise – RS is integrating DL to support better understanding of user-item relationships and extending it to recommendations. RS has also shown capability to enhance personalization in new domains such as education, this may mirror interests in passenger information personalization research by works such as Vredenborg et al. (2025).

4.2.3 Natural Language Processing

Natural Language Processing (*NLP*) is considered a core AI field, seeking to enable computer understanding and use of natural human language. NLP is divided to two sections: Language *understandin*g and *language generation*. Understanding and generation capabilities are useful for several purposes: "text classification, machine translation, speech recognition, automatic summarization, question-answering systems, language models and other fields" (Feng et al., 2025). While the NLP research field is vast - Large Language Models (*LLM*) dominate current interest, enabled by DL and large amounts of available language data. Past approaches, including using classical ML approaches such as SVM and decision trees, were replaced by DL models such as RNN variants during the early 2010s – and later with the rise of the greatly successful transformer architecture after 2017. The release of transformer addressed issues with long-distance dependencies were improved greatly, creating ***Large Language Models*** (*LLM*) which can be trained for nuanced and complex language generation. This success ignited massive new interest in NLP research. (Feng et al., 2025)

To further talk about LLM - Minaee et al. (2025) provide a survey to the concurrent LLM landscape. Minaee et al (2025) dissect 3 different dominant architectures which are mostly based on transformer. The original architecture featured both encoder and decoder, but they are not always used, depending on context.

1. Encoder-Only. Best applications in sentence classification, named entity recognition and extractive question answering.
2. Decoder-Only. Best applications in text generation. Example: GPT models.
3. Encoder-Decoder. Best applications in text generation based on input context.

(Minaee et al., 2025)

LLM, which are commonly pre-trained using from 10 billion to over 100 billion text parameters as learning data, can approach a wide set of NLP tasks. Different tasks range from summarisation, question answering, coding, tooling tasks such as planning and self-criticism to emerging capabilities such as instruction following and reasoning. They can be deployed as standalone tools or integrated as modules within larger systems. LLM are in their nature probabilistic systems which produce output text based on vast training data. They make estimates regarding words and phrases based on input text and its word arrangement. Thus, training data is a highly important topic for language model development. Most common data sources include web crawl data from datasets such as Common Crawl, Wikipedia and books. During the LLM creation process, data cleaning is applied to process the data. This includes removing noise, detecting and filtering anomalies, balancing data categories, preprocessing text to remove punctuation or irrelevant tokens, and resolving dataset contradictions. Duplicate data is also removed, to address bias and help against overfitting - as this can affect how prominent the importance of a topic. This data is then further processed by tokenization, positional encoding and continued to pre-training, fine-tuning, instruction tuning and finally alignment against unintended behaviour. This complex multi-step process can be followed to complete a LLM model. (Minaee et al., 2025) While the subprocesses are highly complex and interesting - and act as background for discussing LLMs, the scope of this work is most focused on the potential use cases, advantages and limitations of AI technologies.

As LLM generate and answer based on the training data including how it was pre-processed for training, it introduces *limitations* for LLM functionality. Minaee et al. cover these LLM limitations in their work, as follows:

1. LLMs do not have an inherent state or memory and rely on external systems for such.
2. LLMs tend to give different answers to same prompts, due to their probabilistic nature.
3. LLMs cannot inherently gain new data and rely on possible external systems for this functionality.
4. LLMs in general require strong GPU for both training and serving.
5. LLM hallucinations happen – either intrinsic which feature direct conflict with source material, or extrinsic, where unverified or speculative information is given as output.

(Minaee et al., 2025)

As many systems which could be built using LLM would require features such as memory or new data – LLM are integrated with other systems to help remedy their shortcomings. One example is ***Retrieval-Augmented Generation*** (*RAG*) where input prompts are extracted to retrieve information from a relevant source, whether this is from a search engine or a database, and then the answer for the query is added to the original LLM input to form the answer. Thus, utilizing RAG, new information can be given as supplements to an LLM system. Other LLM augments include external tool integration, such as with integrating functions or services with LLM, for example utilizing API:s or databases. LLM agents – an area of rising research interest, presents LLM systems which utilize this augmented tool access to act as an autonomous entity, which could make context-dependent decisions and use external tools. Some examples covered by Mianee et al. included an LLM-based agent that could utilize a weather API for location-based information, or an agent that could make online purchases. The survey highlighted future directions for LLM research, which include a push for smaller and more efficient language models, interest in alternatives for to transformer and attention-based architectures, multi-modal LLM, improving LLM issues such as hallucination, LLM agents, Multi-Agent Systems and addressing security and ethics regarding LLM systems. (Minaee et al., 2025)

Vision Language Models (*VLM*) are an emerging combination of computer vision with natural language processing – and an example of multi-modal LLM, which use their language capability to provide machine reasoning. The goal of VLM is to merge both contexts together to help complete a task or goal. VLM are currently considered for tasks such as visual question answering, robotics and autonomous driving. This is achieved through CV feature extraction, which interacts with LLM models. VLM models vary in architecture as the field is developing. Earlier models such as CLIP utilized a separate text encoder and a text decoder to embed text together with vision features in a shared space. Newer models such as LLaVA interact with the visual features using an LLM model directly. (Li et al., 2025) Two example benchmarks highlighted by the Stanford HAI AI Index Report (2025) include:

1. *VCR*, where "common sense" reasoning multiple-choice questions regarding images are tested. The models must provide reasoning behind answers in this benchmark. Since 2018 when the benchmark was formed, a human baseline was reached by a model in 2024.
2. *MVBench*, where reasoning based on videos are tested. The benchmark includes temporal reasoning based on previous events in the video. Benchmarked models reached 69% average accuracy in 2025.

One potential technological synergy of VML is considered to be IoT. In a survey on edge network VML by Sharshar et al. (2025) discussed the potential of lightweight VML for the medical field, environmental monitoring, aerial imaging, autonomous systems such as traffic detection, navigation systems, autonomous transportation and surveillance. While optimized VLM are becoming better and deployable to real scenarios, edge environments present challenges such as optimization, model size, energy efficiency, sensor compatibility, data synchronization, privacy, security and edge network challenges. (Sharshar et al., 2025)

4.2.4 Generative AI

Generative AI (*Gen-AI*) is a field of AI focused on the creation of synthetic data - including images, videos, audio, text, 3D models, code and data. Gen-AI overall has strong overlap with other AI and ML fields including NLP and CV, as many tasks involve producing or transforming images and text. It is based around DL techniques to analyse patterns and structures and then replicating content with similar traits and qualities. Gen-AI also encompasses *content translation*, in which content is transformed, such as text into images or the modification of existing images. Gen-AI has been a major contributor to the recent rise in popularity of AI, as Gen-AI tools have been brought to public availability and attention through tools such as ChatGPT, and various text-to-image tools – capable of creating realistic text and visual content. As an example of Gen-AI subdomains - *image translation* is focused on generating altered versions of an original image. This can be useful for fields such as the medical domain and satellite imagery, where Gen-AI can be used to highlight important sections of the image to ease expert analysis. (Sengar et al., 2025)

Many types of DL architectures are used for Gen-AI in processing different types of data. For example – ***Generative Adversarial Networks*** (*GAN*) are most used for image, video and 3D generation tasks. ***Transformer*** variants find use in sequential tasks with their *attention* mechanism. While originally designed for language tasks but have since been adapted as ***Vision Transformers*** (*ViT*) - which has in many cases outperformed past CNN-based methods. Both Transformer and ViT have applications for Gen-AI. Other notable architecture areas include ***Variational Autoencoders*** (*VAE*) and ***Diffusion Models*** (*DM*). VAE use an encoder and decoder to encoder inputs to a latent space and decode it back to the original shape – and enable the model to learn from this process. (Sengar et al., 2025) Diffusion Models generally function by mapping simple data distributions to more high-dimensional distributions and reversing the process to learn a representation of the mapping. (Zhang et al., 2024)

Sengar et al. (2025) reviewed Gen-AI advancements explored in academic research from various domains. Transformer and VAE inspired architectures were explored for most of these categories – either as proven tools for the area or new explorations.

1. **Image Translation** tasks included studies for medical MRI, satellite imagery, image facial expression editing and style transfer, text-to-image translation and image upscaling.
2. **Video Synthesis** explorations included generations based on different data types. Featured areas included text-to-video and video-to-video, style translation videos, video/image classification, audio-based facial expression video translation, and 3D human motion prediction.
3. **Natural Language Processing**, where generative language models are a key research area, as covered in Chapter 4.2.3. Gen-AI language models can tackle problems such as question answering and assistance tasks. They can also be integrated with other systems for more complex workflows.
4. **Knowledge Graph Generation**, which refers to representing relationships between entities as graphs.
5. **Interdisciplinary Applications**. This area includes a wide range of applications - mechanical fault detections by utilizing GAN synthetic data, generating synthetic traffic scenarios for training data, music generation, handwriting generation, coding assistance and digital twins.

(Sengar et al., 2025)

The work by Sengar et al. chose to cover these research topics – but overall, they present only a narrow fraction of the vast number of research topics for Gen-AI, which would be impossible to capture comprehensively. Their work lacked coverage of the topics of transportation, public transport and smart cities. Specific works for this theme can be found. For example - a literature review by Yan & Li., (2025) covers Gen-AI for ***Intelligent Transportation Systems*** (*ITS*). In their work, the authors focused on urban transportation related systems such as traffic planning systems. They present studies where Gen-AI techniques are being researched as part of *traffic perception, traffic prediction, traffic decision-making* and *traffic-simulation*:

1. In ***Traffic Perception***, Gen-AI could be used for covering data gaps in noisy sensor data – when obstructed by weather or lighting variation. GAN-based approaches were found to be explored for topics such as traffic estimation, traffic data mining – where vehicle movement, traffic images and traffic videos are

analysed. GAN, Transformer, VAE and diffusion techniques were researched to support Traffic Anomaly Detection systems. In general, Yan & Li found benefits for using Gen-AI in this task field. Particularly advantages were found for tasks such as addressing missing data, improving accuracy and spatiotemporal dependency handling. The results, however, included variance when noise and other dynamic changes were present. Some researched task areas, such as traffic data mining and traffic anomaly detection were stated to be improving.

2. In ***Traffic Prediction*** – similarly to perception tasks, Gen-AI could be used to fill data gaps or inaccuracies for prediction purposes. Also, it could improve behavioural modelling and handling of spatiotemporal dynamics. Applications were found for *micro-level* predictions including human and vehicle movement prediction techniques - and *macro-level* predictions regarding road segment and region-related prediction. The latter include traffic flow, travel time and region-specific traffic prediction as examples. Yan & Li find Gen-AI to present high-quality results for traffic prediction, but struggle with interpretability and efficiency at the same time.
3. In ***Traffic Simulation*** - where transportation systems are modelled using mathematics to help urban planning or maintenance. As Gen-AI could be utilized to help create realistic scenarios – particularly rare cases with missing data, it can be useful for this task. Some work-in-progress exploration included modelling driver behaviour, traffic scenario generation, human trajectory generation, traffic flow generation and generating anomalous data. Yan & Li consider current models to be challenged by computational efficiency and scalability – and that large real-time deployments are not yet reached.
4. ***Traffic Decision-Making*** processes consists of route planning, traffic signal control, and autonomous vehicle control. *Route planning* benefits from Gen-AI as part of optimizing efficient routes. Research was found for navigation based on regular map data, multi-level map data, and egocentric images. *Traffic signal control* is thought to be possible to improve. Research is interested in addressing the processing of complex interactions and missing data gap for both single-intersection and multi-intersection coordination. Yan & Li present works that attempted LLM integrations with traffic decision systems by providing sequential modelling assistance. Yan & Li determine that Gen-AI for decision making systems still has improvement areas for handling unpredictable dynamics, addressing sensor limitations and missing rare event training data. (Yan & Li, 2025)

Yan & Li suggest that generic open challenges for ITS Gen-AI are related to handling multi-modal data from different sensors and sources, handling of spatio-temporal correlations, rare data sparsity for model training, data poisoning vulnerabilities, model explainability, real-time functionality, and resource and computing speed performance. They suggest future research to focus to work on addressing these problems. (Yan & Li, 2025) Overall, ITS systems present to benefit from Gen-AI as covered in this section. Themes discussed – such as Gen-AI's ability to address data gaps, generate new training data, enhance reasoning and function as part of subsystems are potentially valuable for other domains as well. Current research provides a wide range of application domains for Gen-AI, but some possible future overall trends were also considered by Sengar et al. One expectation was that model architectures are expected to improve and change, in the same fashion that transformer was adapted to vision tasks as ViT. Sengar et al. expect that techniques will be increasingly combined as hybrids, to take advantage of multiple types of strengths.

Hybrid architectures also align with the theme of multimodal data – including different datatypes for simultaneous ML processing of text, audio, images and video. Synthetic data is also considered a direction, to use in model training – especially which lack data on a certain theme. This theme was also found in other works covered in other chapters, such as in NLP - where niche language datasets cause language models to underperform in rare languages. Synthetic data could also similarly extend to any kind of applicable niche or unexplored topic for model training purposes. Senger et al. also highlight interest in Gen-AI for the education domain, mirroring similar future expectations for RS in the work by Chakraborty et al. (2025). Gen-AI could be used to enhance personalization and educational results by creating customized learning material to fit individual learning needs. Another notable domain is media creation and editing, which Sengar et al. expect to advance in quality and variety of tools available for creation, enhancement and multimedia content creation. (Sengar et al., 2025)

Despite Gen-AI being used for many purposes and continuing to expand to new domains – it has also raised criticism with *technical, ethical and sustainability* themes. Current Gen-AI models present ***technical*** vulnerabilities in themes such as ***out-of-distribution robustness***, which represent underperformance in scenarios that do not have training data. Another prominent issue is ***shortcut learning***, which refers to training data forming brittle correlations in a trained model, which affect the outputs. This refers to cases where a model learns to falsely connect two traits together, such as connecting daylight to humans - when all dataset pictures had images of humans during the day. Manduchi et al. (2025) suggest Causal Generative Models as remedying these issues, which refers to the integration of causal reasoning, rather than only data correlation. (Manduchi et al., 2025)

Sustainability concerns for Gen-AI are focused on resource utilization and efficiency – as model training can be slow and requires vast amounts of energy and produces emissions. ***Model Quantization***, which refers to reducing model weights and activations while not losing accuracy, is poised to produce more efficient and smaller models. ***Ethical*** questions for Gen-AI include questions about misinformation, security, fairness, and transparency. Production of increasingly realistic synthetic content can facilitate more believable misinformation – both intentionally and unintentionally. Altering human images, videos and voice, present potential for fraud, misinformation, and propaganda – also indirectly, as internet content becomes generally more untrustworthy. Unintentional examples may be a question answering model's training data containing inaccuracies. Such question answering models also present a major ***security*** concern, as inaccuracies may arise from "data poisoning"– referring to a party intentionally including malignant content in ML training data. (Manduchi et al., 2025)

Regardless of proposed criticism and noted limitations, Gen-AI has become transformative technology, and continues to rapidly improve in many directions. Generative AI has potential for many new uses and domains through realistic text and visual content. Trained Gen-AI models may be used to build useful tools – integrated as components into larger systems.

4.2.5 Agentic AI

Agentic AI is an emerging paradigm which is focused on the creation and management of autonomous goal-oriented systems and takes advantage of dynamic decision-making possibilities of ML and DL methods. Agentic AI systems may consist of one or multiple collaborating agents. Agentic AI shares a strong connection with *Reinforced Learning* (*RL*) for trial-and-error adjustment, *DL*, *Gen-AI* and in particular – *LLM*. Agentic AI are most often built utilizing LLM - giving them capability for natural language understanding and reasoning. Across different approaches, Agentic AI systems commonly use the following core components: "perception and world model", "memory", "planning reasoning, "execution and activation", "reflection and evaluation" and "communication, orchestration and autonomy". Agentic AI are notably integrated with means to interact with tools often by API access - and have access to live data regarding the use context. (Bandi et al., 2025) Acharya et al. (2025) provides lists three common types of Agentic AI architectures:

1. ***Multi-Agent Systems*** (*MAS*). Characterized by multiple agents sharing tasks, or collaboration between them to solve goals.
2. ***Hierarchical Reinforcement Learning*** (*HRL*). Features a hierarchy of agents in which higher-level agents divide sub-tasks for lower-level agents to complete.
3. ***Goal-Oriented Modular Architectures***. Features a modular design – where modules focus on a specific task only, in collaboration.

(Acharya et al., 2025)

Agentic AI is thought to be able to manage complex real-world scenarios with information processing and strategic planning. Bandi et al. (2025) highlight several domains where Agentic AI is used or experimented on:

- In **Healthcare**, agentic AI may assist diagnosis, treatment plans and personalization of care.
- In **Transportation** and Intelligent Transport Systems Agentic AI may be used to control dynamic routing and signal control, coordination of agents and traffic simulations.
- For **Software** Agentic AI can be used for building real-time responsive tools for various purposes.
- For **Manufacturing**, Agentic AI may be used for efficiency, automation and robotics, decision making, safety and maintenance for various kinds of machinery.
- For **Finance, banking and insurance**, Agentic AI may create risk profiling, automate forecasting for loans, treasury management, fraud detection and personalized agents. (Bandi et al., 2025)

Agentic AI appears to be a promising new type of AI technology, but while individual studies have scored strong against benchmarks – many unsolved issues hinder real-world deployments. Acharya et al. (2025) point out general major improvement areas for Agentic AI - a need for more *robust data pipelines* and *training datasets*, *incorporating feedback systems* and *human oversight* to improve *agent performance* and *ethical behaviour*. Domain-specific issues also limit agent development and deployment – for example, in healthcare systems require heavy data preprocessing, and struggle with data heterogeneity. General issues which the Agentic AI field is encountering and researching include overfitting and challenges integration with existing systems. (Acharya et al., 2025) Bandi et al. further emphasises limitations for long-term strategic reasoning, large volume data processing issues, explainability, coordination between agents, security and task alignment.

Adding to the technical unsolved issues - criticism regarding Agentic AI evaluations has been proposed. Meimandi et al. (2025) argue that academic and industry studies often favour technical benchmarks over human-centric and security benchmarks. They note that productivity gains suggested by benchmarks have not yet consistently projected to real-world deployments – particularly for safety critical domains. Also – the controlled benchmark tests may have projected productivity increases for real-world deployments for some domains, but real-world tests have not yet supported the benchmarks. (Meimandi et al., 2025) Despite criticism, both Acharya et al. and Bandi et al. underline the potential of Agentic AI, and expect systems to develop further into modular, scalable and transparent systems – and underline that alignment with human values, accountability and security are enablers.

4.2.6 Edge AI

Edge AI is the combination of AI with ***edge computing*** - a topic closely related to the ***Internet of Things*** (*IoT*). In IoT, physical objects are given unique identifiers and embedded with sensors, turning them into “things” which connect to the internet and send data to servers operating in the cloud. The IoT *edge* refers to the devices in an IoT network, which are close to the sensors and things, rather than the cloud where information is most often processed. Because the cloud relies on powerful datacentre hardware, it has historically been more efficient to process the sensor data in the cloud rather than on memory-constrained edge devices. Edge computing is an emerging paradigm focused on utilizing edge computing resources. (Gunjal, 2024)

Edge AI is then the process of implementing AI technology on edge devices. Although Edge AI is not a single AI technology, it plays an important role in the deployment and operation of AI systems. Edge computing has several advantages over cloud computing. Examples are lower amounts of latency, reduced network congestion and security benefits such as local data anonymization. Particularly the advantage of lower latency shares a connection to previously covered themes such agentic and safety-critical AI systems, which may require instantaneous reactivity. While many of the AI technologies covered in previous chapters receive criticism for their memory size and processing requirements – they are being addressed through research and development. More powerful hardware and the continuing efforts to create smaller and more efficient models through techniques such as quantization are enabling stronger models to perform at edge devices and thus enabling Edge AI. One major hardware enabler to Edge

AI is represented by Google's **Tensor Processing Units** (*TPU*) in 2018 designed to support AI inference and training in IoT edge environments. (Gill et al., 2025)

Gill et al. (2025) provide an overview Edge AI applications. IoT is a mature research field and widely adopted in real-world deployments, and ML has been widely used in conjunction with it - but comparatively Edge AI is a newer field, being first formed around 2014. Gill et al. present popular application domains for Edge AI. In *Healthcare*, Edge AI can be used in conjunction with wearable IoT devices to enhance real-time patient data collection. For *Smart Parking*, Edge AI can be used to find free spaces. For *Smart Homes*, IoT can be used for many small smart devices such as lighting systems and smart refrigerators. For *Computer Vision*, Edge AI finds use for industrial applications such as device monitoring systems with which make decisions. For *Cyber Security*, Edge AI can be leveraged for monitoring and managing attacks on a network. For *Transportation*, Edge AI can be used for improving traffic management systems in intersections. (Gill et al., 2025)

Gill et al. combine open challenges for advancing Edge AI. The first is infrastructure optimization, which would focus on the creation of stronger *hybrid architectures* from edge, fog and cloud computing. The advantage for enhanced hybrid architectures would be the capability to assign computation tasks to where it is most beneficial depending on context. A second challenge is the security and privacy concerns required to address with data such as biometric, or personal data being handled through AI on edge devices. AI security, explainability and predictability were also concerns for most authors writing about AI covered in this chapter. Finally, Gill et al also call for increased real-world experimental deployments, to identify limitations and find improvement areas for Edge AI and its applications. (Gill et al., 2025)

4.3 Artificial Intelligence in Public Transport

Building on results from the separate overviews from both AI and PT/PI, this chapter examines research which directly integrates both topics. The goal is to find out which AI technologies have been researched and utilized in the domain of PI. The precedingly covered domain of AI provides a strong set of diverse technology. This includes real-time visual data processing, classifying and segmenting data and creation of synthetic data. The research domain of passenger information aims to improving current process efficiency and attract new users to public transport by better UX. As RTI practices have become standard for large, developed cities – they provide valuable data which could be utilized in new ways such as enhanced personalization. There is growing interest in exploring past level 2 personalization (partial personalization) to levels 3 (conditional personalization) and level 4 (full personalization) (van Ardenne et al., 2025). This chapter is divided to the exploration of personalization in PI, PIS enhancement, and PI analysis tools.

4.3.1 Personalization and AI in Passenger Information

To examine studies dealing with personalization - the established personalization object and attribute framework can be used. The personalization attributes and contexts which determine personalization results are physical context, social context, temporal context, spatial context, preferences, characteristics, and ownership. From those, attributes – temporal, spatial are the most likely for the user to produce as a byproduct of PT application and smart card use. While the other contexts can be partly inferred – that process is more complex and may risk infringement of user privacy. Personalization objects then contain four categories – *recommendations, interaction, information* and *vehicle settings*. (Vredenborg et al., 2025)

Recommendations for PI are most often used for route classification and recommendation. Route choice prediction plays a large part in enabling recommendations and personalization. Past route choice prediction and process understanding methods have been focused on mathematical models such as discrete choice models. While common route choice systems can use simpler methods - Path Size Logit (*PSL*) and Mixed PSL are considered state-of-the-art examples. They use a principle of comparing probabilities of route choice based on a provided utility function. Marra & Corman (2025) highlight that new ML approaches are being experimented in the space including their trial DL model. Their approach claimed to outperform PSL in two ways. It could infer a non-linear utility function and was able to include complex interactions between multiple variables including the additions of data like the weather. (Marra & Corman, 2025) The work of Marra & Corman highlighted interest in DL for route choice prediction but suggested that further work is required for this topic. Supporting this view - a review article by Tian et al. (2024) regarding route choice modelling for urban rail transit networks suggested that ML and DL approaches are not commonplace for PT route choice modelling, and are emerging in the field of research. (Tian et al., 2024) Other domains such as travel and tourism studies and reviews were also often found from search results for AI in route planning and recommendations - which may suggest common interest and potentially transferable ideas. (Shi et al., 2025)

For ***Interactions***, we divide subcategories of *Modalities, UI* and *Content Delivery Methods*. A practical study by Romero et al. (2020) examined personalized **UI and content delivery methods**. The study approached this topic through an adaptable transit mobile application interface for buses – built utilizing machine learning methods and tested for 18 months in user groups of non-adaptive and adaptive UI. The goal was to reduce required selections, and manual route filtering by the user. This aligns with a major personalization goal of reducing cognitive load of PI noted by Vredenborg et al. Romero et al. built a ML model using Random Forests to use the users previous travel data to predict new user routes, aligning with level 3 situational awareness from van Ardenne et al. (2025). The study determined that adaptive mobile UI for PI seemed to require less interactive effort to use when successfully implemented, and more if the implementation was not successful – although with acknowledgements that the research topic is difficult to assess and further iterative research would be required. (Romero et al., 2020) Notably – the study relied on ML practices which may have advanced since 2020 but suggests that AI could be utilized smart UI for PI purposes. While interaction types through mobile device UI may be personalized - Schlegel & Titov (2025) suggest that as there is a general challenge in implementing personalization in the public transport context. This tension arises because systems need to ensure privacy, but act in public contexts close to other people. This affects the development of alternative modalities such as audio-based modalities. Visual information on embedded displays such as digital signage is also subject to this challenge.

Information is often related to information about the status of vehicles and non-route related information. This can be updates about relevant nearby vehicles which may affect the journey. Also travel times, relevant notification and disruption guidance are distinct from route recommendations. (Vredenborg et al., 2025) A peer-reviewed journal article regarding urban rail transit disruption management did not include discussion about machine learning or AI in this area – only that it is thought to be future research direction. (Wang et al., 2025) Studies from 2020 onward on vehicle status were not found from the perspectives of personalization and artificial intelligence.

The last personalization object type - ***Vehicle Settings*** is more focused on for private vehicles such as cars. While modern urban transit vehicles may have adjustable settings such as climate control systems, they affect every passenger similarly. Physical comfort attributes changes would also require the implementation of physical systems to support them – a large effort and expense for PT operators attributed parties. However, the current physical spaces have potential to be leveraged better such as by micro-navigation systems – guiding passengers toward stops, vehicles or seating. (Vredenborg et al., 2025)

4.3.2 AI and Passenger Information System Enhancement

AI integration as part of PI and PI-adjacent systems is being developed in many directions. Current research was found to be interesting in AI for personalized UI and route recommendations. These are heavily mobile-facing platforms. Other factors play part in PIS - such as signage, announcements, onboard guidance, vehicle routing and security. These PIS factors are present at stations and onboard transit vehicles, and most often contain general information directed toward all passengers as a group.

For ***Security***, one area of focus for PT and AI research has been to utilize computer vision techniques to monitor items and humans inside the vehicle. PT monitoring systems are commonly used for live monitoring or recording – but do not offer capability to identify or provide warnings. AI technology could provide help for this purpose. As covered - Computer Vision can be used to segment images, recognize entities and anomalies in video footage, 3D understanding, and these are all highly useful for this task area. For example, a study by Liu et al. (2025) utilizes a YOLO-based DL algorithm for vehicle monitoring system data to form a database of unusual passenger behaviour and anomalous items onboard. Their system was trained with data for face mask usage and abnormal object presence. The system achieved an accuracy of 95% in real-time detection for limited training scenarios including mask detection. The paper considered this technology useful for the PI context in the future. As this was only an experiment – a fully developed systems aimed for production use would have to train models with more robust and varying datasets. This presents as a challenging task with also ethical considerations – as niche situations might create false negative and false positive classifications. (Liu et al., 2025)

A further exploration of security-applicable CV technologies is presented by Meurie & Lézoray (2025) in their review of on-board action recognition models for PT. This refers to detecting *basic behaviours* such as vehicle boarding or exiting, *normal behaviours* such as sitting or standing, *abnormal behaviours* such as falling, *violent actions* such as fighting or shouting and *damaging actions* such as vandalism. The most popular recognized CV methods included CNN-based techniques such as YOLO and Single-Short Detection (*SSD*). Meurie & Lézoray provided critical analysis to the studies, as many domains such as bus and metro had their own popular research topics such as passenger counting for bus, and train studies commonly considering both audio and video – which presents the issue of too divided research. The trained study models were also found to be prone to overfitting against their training datasets. Seemingly dataset access is a major blocker, as many PT datasets are not publicly accessible and this issue extends to limited reproducibility, generalization issues, limited transfer learning and inconsistent benchmark comparisons. (Meurie & Lézoray, 2025)

AI research on **Directing Passengers** seems to focus on user mobile devices for personal wayfinding and routing. Along with the prominence of mobile devices – ***Passenger Directions*** are given through different outputs – both passive and active. Static and RTI signage, and audio announcements may be present both at stations and onboard the transit vehicles. ***Digital Signage*** is then a form of PI which is delivered through digital displays on stations or onboard the vehicle. This type of PI delivery seems to be underrepresented in academic research after year 2020, especially when considering the integration of AI with these signage systems. This seems to be the case at least for the isolated topic of “PI and digital signage”. As digital signage extends both underlying and connected ITS capabilities into the physical space, the research of that field is vitally connected. A focus for specific signage related research seems to be strongly linked with accessibility research (Kong et al., 2024). An older paper from Parker & Tomitsch (2017) determined that successful interactive digital signage implementations place focus on three factors. *Positioning* such that it is found at eye-levels where lots of potential users pass through, facing toward the entrance. *The content* was noted to avoid duplicating smartphone information, as users already have access to it and are found likely to avoid signage which may not seem helpful. Rather the signage should focus on providing supplementary information which is not easily accessible on the smartphone – such as a map of the space. *The functionality* must be responsive and fast – and users should understand its purpose and function before interacting with it. (Parker & Tomitsch, 2017)

AI tools might present novel ways to improve accessibility, through for example new user modalities or language tools. For example – the Seoul Metro has installed AI translation systems in an information-kiosk style implementation. It uses voice recognition as a modality – such that people may talk to the information kiosk in their native language and get directions via an integrated LLM. It is one example of new and alternative ways to interact with digital signage and access PI. Novel interactive displays may help visitors without access to a mobile application, or users without an available mobile device. This may underline the message of Parker & Tomitsch in which digital signage should focus on information that may be hard to access for users. (Seoul Metropolitan Government, 2024)

As digital signage is also an investment to the physical spaces, it may be one example to address the gap between the physical and digital space. Ferri & Popp (2023) suggest that the state of physical versus digital navigation is currently somewhat skewed towards mobile digital devices. In their study, they found that modern urban PT users are likely to put emphasis on personal mobile devices both for wayfinding and to cover for uncertainties in cases such as distinguishing stops at a station, or the side of the road and navigational confirming of beliefs. (Ferri & Popp, 2023) Kong et al. note that new digital tools may be less accessible to demographics such as the ageing population, which are then relying more on signage. Yet such demographics may benefit from digital adjustable font, audio and contrast features. (Kong et al., 2024) To summarise, AI may have potential to bridge gaps between physical and digital PI, and to improve digital accessibility. Tools such as successfully implemented AI information kiosks may allow users to be less reliant on mobile devices, applications or even knowing a specific language. Alternative modalities could make it easier to approach by more users and provide accessibility benefits. The tools would likely include language models, speech conversion and a way to model the transit network reliably.

Audio Announcements also play part in passive PI. Audio announcements may come in forms of pre-recorded audio, or live audio. While pre-recorded scenarios for PI such as next stop announcements are most common – new scenarios such as unexpected disruptions may require live audio announcements. Most commonly the driver gives out such announcements, which may present issues for clarity. Sapkota & Shrestha (2020) studied Oslo metro announcements for accessibility – and found that international students and hard of hearing are likely to struggle understanding driver announcements due to reasons such as poor speaker quality combined with the use of only one language, unclear speech and no written message. The authors developed a speech-to-text system for the driver to remedy this problem, such that announcement could be translated and displayed on digital signage and in multiple languages. The translations were done using Google's translation API. (Sapkota & Shrestha, 2020) The work supplemented driver announcements and attempted to improve audio announcement accessibility. However, the original announcement may still struggle from effects such as unclear speech, one language support and non-uniformity with other announcements. Modern ***Text-To-Speech*** (*TTS*) technology may offer the possibility to also convert speech to a unified form with enforced clarity and delivering driver announcement in multiple languages. DL is often utilized in modern TTS to create synthetic natural sounding language, by either large dataset pattern extraction or linguistic and phonological pattern extraction and language modelling. (Alwaisi & Németh, 2024)

4.3.3 AI for Enhancing Passenger Information Analysis

Along with both PI and PIS technologies which are either directly or indirectly affecting the passengers. Works were found discussing AI for commuter pattern analysis, maintenance prediction and transit network digital twins. Rather than focusing on passengers or in-vehicle staff, this group of technology is aimed at the PT providers. The vast amounts of available PI could be both collected and utilized more efficiently using AI techniques – to improve maintenance planning, asset management and operational efficiency. The focus is then also more on long-term operation.

Commuter Patterns analysis is one way to approach this area. As established, PI systems generate vast amounts of data through mobile device use, ticketing systems and smart card usage. Xiong et al. suggest in their review of commuter pattern forecast techniques that interest is growing in utilizing the data for better station flows, event adaptations and road condition optimizations. However - commuting patterns may be challenging to analyse accurately as some of the previously mentioned forms of data are incomplete. The targets for commuter flow forecast are mode choices, routing choices, destinations, durations and flow of commuters. Xiaong et al. identified several papers with different techniques for analysis ranging from discrete choice models, traditional ML, DL and combined models. They also summarized strengths and weaknesses for these methods. DL-based and combined models were reported to answer the complex dependencies in the data, but had their individual issues – and thus the authors did not prefer one method entirely over the other for general use, but selection of the best options for individual cases. (Xiong et al., 2024) Notably the work did not highlight the potential of Gen-AI for this domain, which was stated in multiple other works to help address data gaps.

Extending commuter pattern analysis to ***Demand Prediction*** - a review by Torrepadula et al. (2024) covers studies which use ML for predicting passenger volumes across specific areas and routes. Demand prediction may be approached from a long-term perspective for overall service improvement, or short-term perspective - where real-time data can be to make dynamic adjustments to services. Torrepadula et al. note a growing number of studies in the area - but also specify that 70% of works they found in this area cover metro station data, 22% buses. 19% trains and 2% or tram systems, which indicates an uneven study distribution. As this is a time-series problem, ML and DL techniques which deal well with temporal relationships perform the best. SVM and SVR (*Support Vector Regression*) were employed from simpler ML methods, moving to Autoencoders, RNN-based solutions such as LSTM and using GNN in conjunction with RNN to capture commuter group spatial relations. Torrepadula et al. suggest that transformer is also being explored since 2023 for this area, with promising results. This follows a trend of transformer architecture being found beneficial for new application domains. Torrepadula et al. summarise that while AI techniques in demand prediction are being explored, the domain struggles with lack of reproducible studies. Long-term analysis is less studied than station activity in the short term. A lack of open datasets for this domain is also seemingly a challenge for developing new technology. (Torrepadula et al., 2024)

Da et al. (2025) discuss the emerging role of **Gen-AI for Transport planning** – particularly how it could help alleviate issues such as the discussed data gaps in commuter pattern analysis and demand prediction. As noted earlier - Gen-AI tools based on technology such as GAN and VAE can be used to create and process synthetic data. This can play part in predictive modelling and simulations. Da et al. highlight Gen-AI potential for different transport planning themes. *Data Processing* uses are related to filling missing or corrupted data points and performing data fusions. *Descriptive Analytics* uses are related to examining commuter and vehicle patterns. *Predictive tasks* include forecasting traffic flow, arrival estimates, route performances and demand prediction.

Continuing with *Generative tasks* include the creation and modification of accurate synthetic datasets, creating simulative scenarios such as speculative routes. *Simulation tasks* are then the modelling of these scenarios such as complex traffic flow simulations. While limited studies in these areas prove potential, Gen-AI for transport planning tasks is suggested to be far from reaching its full potential. Current open research problems are stated to be technical barriers for PT practitioners, standardized evaluation metrics, uncertainty quantification, bias mitigation, ethics practices and explainable AI. (Da et al., 2025)

Combining these ideas further covered in a review by Thompson et al. (2025) regarding ***Digital Twin*** (*DT*) technologies for railway systems and overall PT. DT is an emerging field of technology though to become more feasible with advancing AI technologies. Digital Twins are real-time digital replicas of a physical thing – in this case a transit system. Live vehicle locations are currently discoverable by modern PT systems, but DT systems push this idea further by the integration of more sensor elements to monitor the technical aspects of vehicles. Examples of this are vibration sensors to gather track data for a railway vehicle. It could be analysed to enable predictive maintenance for the vehicle or tracks. A theoretical fully realized DT system could then enable deep understanding regarding most aspects about a transit network and make predictions based on historical data. Major enablers for this technology are thought to be ML methods for processing PT data which would be collected using IoT sensors. Thompson et al.

suggest that DT systems have wide potential which is thought to become more feasible through ML advancements and research for the domain. There are still limiting factors. Firstly, the integration of different types of accurate real-time capable sensors in a transit network is laborious and expensive. The development of supporting digital systems contribute to this, including their required extensive security requirements. There is also a lack of frameworks for implementing DT systems. Thompson et al. note that while these problems persist, their review found practical implementation case studies which were found to have success in a limited area. As an example - deep learning technique was used to improve track maintenance efficiency and reducing operational downtime in a study by Sresakoolchai & Kaewunruen (2023), which highlight that DT potential can be realized in limited ways in the current day. What Thompson et al. suggest as enabling steps for DT are policy instruments for more pilot projects, standardization efforts across PT subsystems for data formats and integration protocols and the development of implementation frameworks. (Thompson et al., 2025) As with other AI tools, many other fields are driving DT development, such as healthcare, autonomous driving, communication systems, energy, environment, infrastructure and virtual scenarios. A literature review by Chiaro et al. (2025) also suggests a link between Gen-AI and DT research. (Chiaro et al., 2025) Developments in these areas and domains are likely to reflect into PT and PI in some ways. Works by Thompson et al. and Chiaro et al. both present DT technology as potentially transformative in the future but still seemingly far from being fully realized.

5. DISCUSSION

This discussion chapter will be dedicated to answering the research questions RQ1, RQ2, and the discussion of limitations and future work. The following chapter will provide a conclusion.

5.1 Research Questions

1. ***RQ1:*** What concurrent needs and interests are found in research for passenger information in public transport?
2. ***RQ2:*** What kinds of technologies does artificial intelligence present to answer the needs and interests described in RQ1?

RQ1: Findings from PI and PT literature point out that the real-time information is a standard part of high-quality PI service in modern urban settings. Recent PI and PT literature reviews highlighted personalization such as Vredenborg et al. (2025), and AI in Jevinger et al. (2024). Overall, some highlighted need and interests across research are *leveraging underutilized data, providing more accurate and personable information*, *optimizing supporting systems* and *exploring* new potential provided by emerging *AI tools*.

RQ2: AI represents a broad domain of technology, that can act as an enabler for addressing RQ1 points. Any domain characterized by large amounts of available data holds potential to synergize with AI – including PT and PI. *Computer vision* algorithms such as *CNN*, *YOLO* can be used to process images or video content and build features for security and passenger guidance. *Natural language processing* is currently focused on pre-trained language tools which may be used in passenger-facing interfaces through conversational agents, or integrated in reasoning components for other systems, such as *visual language models*. *Generative AI* has potential for transport simulation which may extend to further optimization of transit routes and fleet management. It also supports the creation of synthetic data, which can supplement datasets with rare real-world examples. *Agentic AI* may be used to optimize and automate operational processes within PT. Advances in *Edge computing* may present new ways to deploy passenger micro-navigation, security and maintenance systems in transport infrastructure.

5.2 Limitations and future work

This narrative review covered a range of topics for two themes – AI and PI. The review successfully found answers to RQ1, RQ2. Because of the wide scope, each topic could only be explored at a high level. The choice of a narrative review limits reproducibility of the review, and the selection of the literature was ultimately at the discretion of the author. In several cases, the available literature which contained broad literature reviews into the topics from 2020 onward was limited to one or two authors, leaving less opportunities to compare perspectives. These factors introduce limitations to the review and should be considered.

Regarding future work – this study was focused on personalization, as the theme contained the most relevant and high-quality literature reviews since 2020 for the domain of PI. Contrasting with AI literature – the findings in both fields seem to have synergy. For example, leveraging available passenger data better, and using it for beneficial AI systems does seem to have alignment. While there is research suggesting benefits for personalization in other domains – sufficient studies on its *effects* for the domain of PI does not seem to be available. There is also a lack of research on *different types of personalization* including personalization attributes. Recent works have begun constructing frameworks for examining personalization to address this. The integration of personalization tools would increase demand for *data storage* and *standardization* efforts for it to be used with architecturally fragmented PT services. Personalization systems would include increased processing of user data – which introduces privacy and security concerns as noted by several authors. Route data is already available by smart cards and transit applications, and it may not present as wide of a concern as introducing new personalization attributes. AI could have synergy in new systems utilizing this data. Overall, it is important to figure out whether the benefits of personalization in PI are larger than the difficulties in their implementation. Aside from personalization, multiple other systems such as *security systems*, *analysis tools* and *accessibility enhancements* are potential use for AI. Technologies such as *digital twin* technology are also emerging with great promise for PT maintenance and fleet management, but research is in the early stages. Further research for these systems is required for the PT context – especially including *real-life prototype* validation.

6. CONCLUSIONS

The large scope of this narrative review inspected passenger information, artificial intelligence and their intersection in separate chapters – to cover a wide space of information about the topics and finally answer the research questions. A focus on topics on a mid-to-high level enabled the discovery of broad themes such as personalization. Currently personalization is an area of interest for Passenger Information research, included in reviews by Vredenborg et al. (2025) and Ait-Ali & Peterson (2025). This approach seeks to leverage hitherto underutilized large user data volumes, which are created as a byproduct of passenger smart card and transit application use. The benefits of personalization are thought to be related to the ability to cater *information, recommendation and interactivity* methods to the individual needs of passengers, rather than servicing everyone in the same way. This would put even more emphasis on the mobile platform, as it is thought to be easiest to implement personalization services through. Some examples can be personalized route recommendations and personal user interfaces. Van Ardenne et al. (2025) present the idea that higher levels of personalization are reflected through increased proactive system behavior such as pop-ups in mobile devices for personalized notifications.

AI, machine learning and particularly deep learning are thought to be enablers for personalization and for improving and creating new PI and PT services. Deep Learning methods such as CNN, RNN and Transformer -based architectures can be used to detect complex and abstract patterns from input data based on training. These technologies synergize with personalization as it requires the creation of user profiles based on data such as spatio-temporal travel contexts and personal attributes. DL is researched for many PT systems such as automated security monitoring, automated seating guidance – and enhanced real-time analysis tools for passenger flows.

While AI tools are increasingly becoming more efficient, many technology applications are still emerging domains, such as AI empowered digital twin technologies for public transport network monitoring, management and maintenance prediction – which is greatly promising idea but suffers from a lack of developed frameworks and publicly available training data. These kinds of solutions also require the large-scale deployment and maintenance of IoT sensors and the creation of secure supporting technologies which can be large expenses. Integration of systems such as these also require standardization efforts between different PT systems and subsystems – which may be challenging particularly if the transit system contains multiple PT operators.

Security, explainability and ethical questions are also factors which hinder AI tool development. Frameworks addressing these questions are still evolving and most often insufficient – highlighted by many reviews in this work. Security of AI systems is a large factor as they can be vulnerable to new types of attacks such as training dataset poisoning. User personal data is also increasingly important to secure – especially for systems which implement personalization. IoT devices, AI and edge computing may also bring new security vulnerabilities, which must be considered. As AI tools are often making decisions, their explainability can pose an issue – which efforts such as *explainable AI* is attempting to bring attention to. If the system functionalities cannot be explained – their decisions may be untrustworthy, especially when the use-cases are safety-critical. Ethical issues must be also answered, such as user privacy. For example, personalization tools would require more user data – are users consenting to that? Legal considerations are also a part of this discussion.

Overall, some AI tools are becoming more viable while others are still emerging or suffer from critical issues or lack research. The AI domain is vast, with many different DL-based tools being the most popular and often the most performant – but ML is not always outdated depending on the context. When considering implementation of AI technology, there are often multiple options – and multiple variants in those, which may make tool selection difficult. Rapid developments in the general field also bring difficulty for developers – as the current best approach could be outdated quite quickly. Not to say that this is new for the field of software.

An important final note from Jevinger et al. (2024) was that PT providers and operators often have limited interest in writing scientific publication. Thus industry-driven innovations and practical insights may remain underreported. Increasing collaboration between academic, private and public parties could support more comprehensive research on passenger information and public transport systems.

REFERENCES

Acharya, D. B., Kuppan, K., & Divya, B. (2025). Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey. *IEEE Access*, *13*, 18912–18936. https://doi.org/10.1109/ACCESS.2025.3532853

Ahlawat, H., Aggarwal, N., & Gupta, D. (2025). Automatic Speech Recognition: A survey of deep learning techniques and approaches. *International Journal of Cognitive Computing in Engineering*, *6*, 201–237. https://doi.org/10.1016/j.ijcce.2024.12.007

Ait-Ali, A., & Peterson, A. (2025). Assessing the effects of traffic information to passengers: A literature review. *Transportation Research Procedia*, *82*, 24–42. https://doi.org/10.1016/j.trpro.2024.12.026

Alwaisi, S., & Németh, G. (2024). Advancements in Expressive Speech Synthesis: A Review. *Infocommunications Journal*, *16*(1), 35–46. https://doi.org/10.36244/ICJ.2024.1.5

Amjoud, A. B., & Amrouch, M. (2023). Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review. *IEEE Access*, *11*, 35479–35516. https://doi.org/10.1109/ACCESS.2023.3266093

Association for Computing Machinery. (n.d). ACM Digital Library. https://dl.acm.org/

Baldauf, M., & Tomitsch, M. (2020). *Pervasive displays for public transport: An overview of ubiquitous interactive passenger services*. 37–45. https://doi.org/10.1145/3393712.3395335

Bandi, A., Kongari, B., Naguru, R., Pasnoor, S., & Vilipala, S. V. (2025). The Rise of Agentic AI: A Review of Definitions, Frameworks, Architectures, Applications, Evaluation Metrics, and Challenges. *Future Internet*, *17*(9), 404-. https://doi.org/10.3390/fi17090404

Berahmand, K., Daneshfar, F., Salehi, E. S., Li, Y., & Xu, Y. (2024). Autoencoders and their applications in machine learning: A survey. *Artificial Intelligence Review*, *57*(2), 28. https://doi.org/10.1007/s10462-023-10662-6

Brakewood, C., & Watkins, K. (2019). A literature review of the passenger benefits of real-time transit information. *Transport Reviews*, *39*(3), 327–356. https://doi.org/10.1080/01441647.2018.1472147

Brar, K. K., Goyal, B., Dogra, A., Mustafa, M. A., Majumdar, R., Alkhayyat, A., & Kukreja, V. (2025). Image segmentation review: Theoretical background and recent advances. *Information Fusion*, *114*, 102608-. https://doi.org/10.1016/j.inffus.2024.102608

Cernadas, E. (2024). Applications of Computer Vision, 2nd Edition. *Electronics*, *13*(18), 3779. https://doi.org/10.3390/electronics13183779

Chiaro, D., Qi, P., Pescapè, A., & Piccialli, F. (2025). Generative AI-Empowered Digital Twin: A Comprehensive Survey With Taxonomy. *IEEE Transactions on Industrial Informatics*, *21*(6), 4287–4295. https://doi.org/10.1109/TII.2025.3540473

Da, L., Chen, T., Li, Z., Bachiraju, S., Yao, H., Li, L., Dong, Y., Hu, X., Tu, Z., Wang, D., Zhao, Y., Zhou, B., Pendyala, R., Stabler, B., Yang, Y., Zhou, X., & Wei, H. (2025). *Generative AI in Transportation Planning: A Survey* (No. arXiv:2503.07158). arXiv. https://doi.org/10.48550/arXiv.2503.07158

Darsena, D., Gelli, G., Iudice, I., & Verde, F. (2023). Sensing Technologies for Crowd Management, Adaptation, and Information Dissemination in Public Transportation Systems: A Review. *IEEE Sensors Journal*, *23*(1), 68–87. https://doi.org/10.1109/JSEN.2022.3223297

European Commission. (2025, June 28). *Traveller Information—European Commission*. https://transport.ec.europa.eu/transport-themes/smart-mobility/road/application-areas/traveller-information_en

European Commission. JRC. (2020). *AI watch: Defining Artificial Intelligence : towards an operational definition and taxonomy of artificial intelligence.* Publications Office. https://data.europa.eu/doi/10.2760/382730

Eurostat. (2025). *Key figures on European transport: 2024 edition*. *2024*. https://op.europa.eu/publication/manifestation_identifier/PUB_KS0124021ENC

Fan, H., & Poole, M. S. (2006). What Is Personalization? Perspectives on the Design and Implementation of Personalization in Information Systems. *Journal of Organizational Computing and Electronic Commerce*, *16*(3–4), 179–202. https://doi.org/10.1080/10919392.2006.9681199

Feng, C., Li, Y., Chen, Z., & Guo, L. (2025). The Evolution and Breakthrough of Natural Language Processing: The Revolution from Rules to Deep Learning. *Proceedings of the 2024 5th International Conference on Computer Science and Management Technology*, 307–311. https://doi.org/10.1145/3708036.3708089

Ferrari, R. (2015). Writing narrative style literature reviews. *Medical Writing*, *24*(4), 230–235. https://doi.org/10.1179/2047480615Z.000000000329

Ferri, A., & Popp, M. (2023). Mind the gap: Navigating the space between digital and physical wayfinding in public transit. *European Journal of Transport and Infrastructure Research*, *23*(1), 63–84. https://doi.org/10.59490/ejtir.2023.23.1.6261

Garapati, R., & Chakraborty, M. (2025). Recommender systems in the digital age: A comprehensive review of methods, challenges, and applications. *Knowledge and Information Systems*, *67*(8), 6367–6411. https://doi.org/10.1007/s10115-025-02453-y

Gheewala, S., Xu, S., & Yeom, S. (2025). In-depth survey: Deep learning in recommender systems—exploring prediction and ranking models, datasets, feature analysis, and emerging trends. *Neural Computing and Applications*, *37*(17), 10875–10947. https://doi.org/10.1007/s00521-024-10866-z

Gill, S. S., Golec, M., Hu, J., Xu, M., Du, J., Wu, H., Walia, G. K., Murugesan, S. S., Ali, B., Kumar, M., Ye, K., Verma, P., Kumar, S., Cuadrado, F., & Uhlig, S. (2025). Edge AI: A Taxonomy, Systematic Review and Future Directions. *Cluster Computing*, *28*(1), 18-. https://doi.org/10.1007/s10586-024-04686-y

Google. (n.d). Google Scholar. https://scholar.google.com/

Gunjal, P. R. (2024). *Internet of Things: Theory to Practice* (First edition.). CRC Press.

Hall, T., & Rasheed, K. (2025). A Survey of Machine Learning Methods for Time Series Prediction. *Applied Sciences*, *15*(11), 5957. https://doi.org/10.3390/app15115957

Hojjati, H., Ho, T. K. K., & Armanfard, N. (2024). Self-Supervised Anomaly Detection in Computer Vision and Beyond: A Survey and Outlook. *Neural Networks*, *172*, 106106. https://doi.org/10.1016/j.neunet.2024.106106

Jevinger, Å., Zhao, C., Persson, J. A., & Davidsson, P. (2024). Artificial intelligence for improving public transport: A mapping study. *Public Transport*, *16*(1), 99–158. https://doi.org/10.1007/s12469-023-00334-7

Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). Quo vadis artificial intelligence? *Discover Artificial Intelligence*, *2*(1), 4. https://doi.org/10.1007/s44163-022-00022-8

Karim, M., Khalid, S., Aleryani, A., Khan, J., Ullah, I., & Ali, Z. (2024). Human Action Recognition Systems: A Review of the Trends and State-of-the-Art. *IEEE Access*, *12*, 36372–36390. https://doi.org/10.1109/ACCESS.2024.3373199

Kong, Q., Li, N., & Li, P. (2024). Literature Review on Age-Friendly Design of Urban Rail Transit way-finding Signage in the Context of Digital Economy. *2024 IEEE 7th International Conference on Electronic Information and Communication Technology (ICEICT)*, 1193–1196. https://doi.org/10.1109/ICEICT61637.2024.10670808

Li, Z., Wu, X., Du, H., Liu, F., Nghiem, H., & Shi, G. (2025). *A Survey of State of the Art Large Vision Language Models: Alignment, Benchmark, Evaluations and Challenges* (No. arXiv:2501.02189). arXiv. https://doi.org/10.48550/arXiv.2501.02189

Lin, T., Wang, Y., Liu, X., & Qiu, X. (2022). A survey of transformers. *AI Open*, *3*, 111–132. https://doi.org/10.1016/j.aiopen.2022.10.001

Liu, S., Bi, Y., Li, Q., Ren, Y., Ji, H., & Wang, L. (2025). A deep learning based detection algorithm for anomalous behavior and anomalous item on buses. *Scientific Reports*, *15*(1), 2163. https://doi.org/10.1038/s41598-025-85962-8

Mahadevkar, S. V., Khemani, B., Patil, S., Kotecha, K., Vora, D., Abraham, A., & Gabralla, L. A. (2022). A Review on Machine Learning Styles in Computer Vision—Techniques and Future Directions. *IEEE Access*, *10*, 1–1. https://doi.org/10.1109/ACCESS.2022.3209825

Maleki, F., Ovens, K., Najafian, K., ghani, B., Reinhold, C., & ghani, R. (2020). Overview of Machine Learning Part 1: Fundamentals and Classic Approaches. *Neuroimaging Clinics of North America*, *30*(4), e17-. https://doi.org/10.1016/j.nic.2020.08.007

Manduchi, L., Meister, C., Pandey, K., Bamler, R., Cotterell, R., Däubener, S., Fellenz, S., Fischer, A., Gärtner, T., Kirchler, M., Kloft, M., Li, Y., Lippert, C., Melo, G. de, Nalisnick, E., Ommer, B., Ranganath, R., Rudolph, M., Ullrich, K., … Fortuin, V. (2025). *On the Challenges and Opportunities in Generative AI* (No. arXiv:2403.00025). arXiv. https://doi.org/10.48550/arXiv.2403.00025

Marra, A. D., & Corman, F. (2025). Modelling route choice in public transport with deep learning. *Transportation (Dordrecht)*. https://doi.org/10.1007/s11116-025-10597-7

Maurício, J., Domingues, I., & Bernardino, J. (2023). Comparing Vision Transformers and Convolutional Neural Networks for Image Classification: A Literature Review. *Applied Sciences*, *13*(9), 5521-. https://doi.org/10.3390/app13095521

Meimandi, K. J., Aránguiz-Dias, G., Kim, G. R., Saadeddin, L., & Kochenderfer, M. J. (2025). *The Measurement Imbalance in Agentic AI Evaluation Undermines Industry Productivity Claims* (No. arXiv:2506.02064). arXiv. https://doi.org/10.48550/arXiv.2506.02064

Meurie, C., & Lézoray, O. (2025). A comprehensive review of on-board action recognition models in public transportation systems. *Expert Systems with Applications*, *290*, 128311. https://doi.org/10.1016/j.eswa.2025.128311

Mienye, I. D., & Swart, T. G. (2024b). A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications. *Information*, *15*(12), 755. https://doi.org/10.3390/info15120755

Mienye, I. D., Swart, T. G., & Obaido, G. (2024a). Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications. *Information*, *15*(9), 517. https://doi.org/10.3390/info15090517

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2025). Large Language Models: A Survey. *arXiv.Org*. https://doi.org/10.48550/arXiv.2402.06196

Mohamed, N. (2025). Artificial intelligence and machine learning in cybersecurity: A deep dive into state-of-the-art techniques and future paradigms. *Knowledge and Information Systems*, *67*(8), 6969–7055. https://doi.org/10.1007/s10115-025-02429-y

Mohd Noor, M. H., & Ige, A. O. (2025). A survey on state-of-the-art deep learning applications and challenges. *Engineering Applications of Artificial Intelligence*, *159*, 111225. https://doi.org/10.1016/j.engappai.2025.111225

Nestor Maslej, , Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki, Emily Capstick, Anka Reuel, Erik, Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Russell Wald, Tobi Walsh, Armin Hamrah, & Lapo Santarlasci, Julia Betts Lotufo, Alexandra Rome, Andrew Shi, Sukrut Oak. (2025). The AI Index 2025 Annual Report. *Institute for Human-Centered AI, Stanford University*. https://hai.stanford.edu/ai-index/2025-ai-index-report

Ninčević Pašalić, I., Ćukušić, M., Golem, S., & Jašić, T. (2023). Information Systems and Technologies for Green Public Transportation: In Z. Yilmaz, S. Golem, & D. Costescu (Eds.), *Practice, Progress, and Proficiency in Sustainability* (pp. 68–97). IGI Global. https://doi.org/10.4018/978-1-6684-5996-6.ch004

Parker, C., & Tomitsch, M. (2017). *Bridging the interaction gulf: Understanding the factors that drive public interactive display usage*. 482–486. https://doi.org/10.1145/3152771.3156162

Romero, O. J., Haig, A., Kirabo, L., Yang, Q., Zimmerman, J., Tomasic, A., & Steinfeld, A. (2020). A Long-Term Evaluation of Adaptive Interface Design for Mobile Transit Information. *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, 1–11. https://doi.org/10.1145/3379503.3403536

Sapkota, A., & Shrestha, R. (2020). Accessible Driver Announcement in Public Transportation: A Solution based on Speech-to-Text and Display. *2020 IEEE 8th International Conference on Smart City and Informatization (iSCI)*, 13–19. https://doi.org/10.1109/iSCI50694.2020.00010

Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, *2*(6), 420. https://doi.org/10.1007/s42979-021-00815-1

Schlegel, T., & Titov, W. (2025). A Sustainable Development Process for Visually Interactive Companions in Ubiquitous Passenger Information Systems. *Sustainability*, *17*(17), 7699-. https://doi.org/10.3390/su17177699

Sengar, S. S., Hasan, A. B., Kumar, S., & Carroll, F. (2025). Generative artificial intelligence: A systematic review and applications. *Multimedia Tools and Applications*, *84*(21), 23661–23700. https://doi.org/10.1007/s11042-024-20016-1

Seoul Metropolitan Government. (2024, March 18). *Seoul Metro Expands AI Translation Services Available in 13 Languages to 11 Stations -*. https://english.seoul.go.kr/seoul-metro-expands-ai-translation-services-available-in-13-languages-to-11-stations/

Sharshar, A., Khan, L. U., Ullah, W., & Guizani, M. (2025). Vision-Language Models for Edge Networks: A Comprehensive Survey. *IEEE Internet of Things Journal*, *12*(16), 32701–32724. https://doi.org/10.1109/JIOT.2025.3579032

Shi, R., Wei, D., Wang, L., & Li, C. (2025). Deep Learning for Predictive Travel Planning: Challenges and Opportunities. *Proceedings of the 2025 International Conference on Big Data, Communication Technology and Computer Applications*, 66–74. https://doi.org/10.1145/3727505.3727516

Sresakoolchai, J., & Kaewunruen, S. (2023). Railway infrastructure maintenance efficiency improvement using deep reinforcement learning integrated with digital twin based on track geometry and component defects. *Scientific Reports (Nature Publisher Group)*, *13*(1), 2439. https://doi.org/10.1038/s41598-023-29526-8

Tampere University. (n.d). Andor. https://andor.tuni.fi

Terven, J., Córdova-Esparza, D.-M., & Romero-González, J.-A. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, *5*(4), 1680–1716. https://doi.org/10.3390/make5040083

Thompson, E. A., Lu, P., Alimo, P. K., Atuobi, H. B., Akoto, E. T., & Abbew, C. K. (2025). Revolutionizing railway systems: A systematic review of digital twin technologies. *High-Speed Railway*. https://doi.org/10.1016/j.hspr.2025.05.005

Tian, Y., Zhu, W., & Song, F. (2024). Route choice modelling for an urban rail transit network: Past, recent progress and future prospects. *European Transport Research Review*, *16*(1), 52–15. https://doi.org/10.1186/s12544-024-00677-7

Torrepadula, F. R. di, Napolitano, E. V., Di Martino, S., & Mazzocca, N. (2024). Machine Learning for public transportation demand prediction: A Systematic Literature Review. *Engineering Applications of Artificial Intelligence*, *137*, 109166. https://doi.org/10.1016/j.engappai.2024.109166

van Ardenne, M. T., Cebecauer, M., Cats, O., & Ma, Z. (2025). Personalised passenger information systems in public transport: A review and a 5-level personalisation taxonomy. *Transport Reviews*, *0*(0), 1–32. https://doi.org/10.1080/01441647.2025.2537203

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. ukasz, & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, *30*. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Vredenborg, M., Van Kasteren, A., & Masthoff, J. (2025). Personalization in Public Transport Passenger Information Systems: A Systematic Review and Framework. *ACM Computing Surveys*, *57*(9), 1–38. https://doi.org/10.1145/3721478

Wang, Y., Li, Y., Chen, X., & Zhou, Y. (2025). Urban rail transit disruption management: A decade of research on network resilience, internal, and external strategies. *Intelligent Transportation Infrastructure*, *4*, liaf014. https://doi.org/10.1093/iti/liaf014

Xiong, J., Xu, L., Wei, Z., Wu, P., Li, Q., & Pei, M. (2024). Identifying, Analyzing, and forecasting commuting patterns in urban public Transportation: A review. *Expert Systems with Applications*, *249*, 123646. https://doi.org/10.1016/j.eswa.2024.123646

Yan, H., & Li, Y. (2025). Generative AI for Intelligent Transportation Systems: Road Transportation Perspective. *ACM Comput. Surv.*, *57*(12), 315:1-315:45. https://doi.org/10.1145/3719290

Zhang, C., Zhang, C., Zhang, M., Kweon, I. S., & Kim, J. (2024). *Text-to-image Diffusion Models in Generative AI: A Survey* (No. arXiv:2303.07909). arXiv. https://doi.org/10.48550/arXiv.2303.07909

Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. *The Artificial Intelligence Review*, *57*(4), 99-. https://doi.org/10.1007/s10462-024-10721-6

APPENDIX A

Table 6 Passenger Information Content Types. Adapted from Ait-Ali (2025)

Information Content	RTI	Prescrip-tive	Quantita-tive	Active
Train Number.				
Start and End Stations.				
Intermediate Stations.				
Operator Name.				
Arrival and Departure Times			X	
Arrival and Departure Platforms		X		
Disrupted Train Arrival Estimates.	X		X	X
Disruption Cause.	X			X
Disruption Summary and Fore-cast.	X		X	X
Booking and Miscellaneous Infor-mation.	X	X		X
Instructions During Disruption.	X	X		X

This table is adapted from Ait-Ali & Peterson (2025), and describers basic information content for PI, and their notable classifications.