




# Interoperable smart card data management in public mass transit

Filip Covic<sup>1</sup> · Stefan Voß<sup>1</sup> 

Accepted: 21 September 2019 / Published online: 9 October 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

Due to a lack of shared practices of deployment, installation and application, the first commercial smart ticketing projects were built on proprietary specifications limiting their scope of integration and compatibility between them. As a result, a move towards global standards and specifications can be observed in current research as well as in practical applications. Therefore, interoperability in public mass transit has become a central aspect of e-ticketing. In this paper, we develop a standardised process on how to handle the emerging smart card data in an interoperable environment. The goal is to present a unified approach where data mining tools and model applications can be tested and implemented in every region embedded in the integrated network. The Interoperable Smart Card Data Chain (ISCDC), which is presented in this paper, provides a continuous procedure for standardised data handling and management. Using insights from expert interviews with German public transit entities, we deduce best practices on how to implement the ISCDC effectively.

**Keywords** Smart card data · Intelligent Ticketing System · Interoperability · Standardisation in public transport · e-ticketing

## 1 Introduction

Since the 1990s smart cards have been commercially used in public mass transit to capture the benefits of an IT-supported ticketing system. Famous examples include the Octopus Card in Hong Kong launched in 1997 and the OysterCard in London launched in 2003. The main tasks of these cards may be seen in collecting fares/revenues and in processing up-to-date information on transactions for planning purposes by helping to understand usage profiles (Pelletier et al. 2011; CEN 2003). The surrounding smart card infrastructure facilitates and improves the ticketing process,

---

✉ Stefan Voß  
stefan.voss@uni-hamburg.de

<sup>1</sup> Institute of Information Systems, University of Hamburg, Von-Melle-Park 5, Hamburg 20146, Germany

i.e. the linkage between the operator and the passenger. A holistic ticketing system has to present innovative solutions to manage access rights to the operator's network, organise the sales of tickets through different distribution channels, enable a convenient top-up and make effective instruments for validation and inspection available (Baric et al. 2011). However, an initial lack of shared practices of deployment, installation and application can be observed (van Oort et al. 2015).

As a result, interoperability and standardisation in smart ticketing have been experiencing an increased interest by various academic, public and industry stakeholders. Recent smart ticketing implementations and projects are developing their concepts based on common specifications to ensure compatibility and upgradability. Due to the changing preconditions extending the number of involved players, the effects on availability, quantity and quality of arising smart card data must be examined (Robinson et al. 2014). These are vital to appraise the effects on data mining and applied models to evaluate them.

By analysing the various models in literature, this paper focusses on requirements for smart card data to enable a model application in an interoperable environment. For this reason, in this paper the interoperable smart card data chain (ISCDC) is developed which states the elements and processes of standardised smart card data collection, management and provision. The ISCDC is a conceptual framework based on currently available technical standards providing a possible approach to integrate the latter. As auxiliary tool, expert interviews with representatives from major German public transit entities were conducted to support the identification of implementation issues of the ISCDC and align its conceptualisation with practical insights. The German public transit industry is closely cooperating within the interoperable VDV-Kernapplikation which is the only association in Germany for managing and organising a common e-ticketing standard. For this reason, an expert interview with the director of the joint managing association VDV-Kernapplikation was used to get an overall view on the interoperable cooperation of public transit entities. Additionally, an expert interview with the project leader for smart ticketing from a major German public transit operator was conducted to get practical insights from the point of view of operators in the German smart ticketing landscape.

The paper is organised as follows. Section 2 gives a literature review focusing on relevant topics for this paper, smart card data for planning purposes in public mass transit and intelligent ticketing systems (ITS) interoperability. Though, subsequent sections are also moderately interleaved with additional references to extend this review. Section 3 presents the stepwise development of the ISCDC. Section 4 discusses the results of the ISCDC development aligning these with insights from expert interviews. Finally, Sect. 5 draws a conclusion and highlights further aspects to be considered in future research.

## 2 Literature review

Academic, public and industry institutions provide research material on ITS, smart card data and their interoperable implementation. Thereby, the emphasis of studies highly depends on the type of the institution conducting it. Generally, the outcome

of this review shows a high focus on data analysis in academic literature, on institutional implementation in studies by public institutions and on technical issues, i.e. standardisation, in industry sources. All types of literature try to incorporate practical inputs from ITS projects to back up their results and to draw conclusions from them. Nevertheless, a gap between theoretical data model development in academic literature and practical (interoperable) data utilisation guidelines by non-academic sources can be identified. For the context of this paper, we provide an overview of literature on smart card data and interoperability.

## 2.1 Smart card data

Pelletier et al. (2011) offer a comprehensive literature review of smart card data utilisation and modelling. They classify smart card data mining tools, model applications and general data utilisation on the *strategic* (long-term data for customer behaviour analysis and demand forecasting) *tactical* (data for schedule adjustments and analysis of trip patterns) and *operational level* (supply and demand indicators, smart card system operations). Based on these categories they evaluate the data types used, the performed analysis by each author and the resulting benefits of the respective data examination. A recent collection of works in this field is provided by Kurauchi and Schmöcker (2017).

Dealing with specific smart card data utilisation, the research can be grouped into data analysis as an instrument for planning purposes and marketing metrics. The main types for planning can be split into:

- a. Travel patterns and passenger behaviour (including the deduction of not directly accessible passenger information by scrutinising smart card data)
- b. Public transit performance indicators and efficiency

In reference to travel patterns and passenger behaviour, Kurauchi et al. (2014) develop a method to prove the hypothesis that passengers follow *hyperpaths* (complex route strategies perceived as attractive by passengers based on characteristic public transit indicators like train frequency or route length) by using a data set from London's Oystercard bus lines. They use a 2-Markov model and a 3-Markov model and consider route overlapping on various degrees to prove their hypothesis. Morency et al. (2006) measure the travel variability of passengers based on the activity of smart card utilisation, boarding numbers per day by fare types (measuring variability between weekdays), numbers of different stops/stations used by one passenger with increased time of smart card utilisation and boarding time (clustering of passengers to identify days when the passenger load is highest). Addressing a similar topic, Zhong et al. (2015) use multi-day smart card data to measure variability in the transportation network on the individual and aggregate level. They employ statistical analysis, correlation matrices and network clustering methods and, afterwards, apply these to a case study in Singapore. Bouman et al. (2012) model the passenger flow based on agent-based micro simulation to detect passengers' traveling patterns

and analyse changes in public transit policies and their influence on passenger travel behaviour, public transit activity and utilisation.

In addition, recent developments in data mining have spurred the application of related methods to identify general patterns in long-term commuting (Goulet-Langlois et al. 2016; Ma et al. 2017; Nishiuchi et al. 2018), to close gaps in collected smart card data (Kusakabe and Asakura 2014) and to better understand and distinguish transfer trips from unlinked trips (He and Trepanier 2015; Hong et al. 2016) including the surrounding activities of passengers (Nassir et al. 2015).

Other works on data analysis focus on related issues such as, e.g., data reliability (Qin et al. 2011), loyalty estimation based on a time model (Trepanier et al. 2012), transfer time (Jang 2010), behavioural change (Asakura et al. 2012), revenue management (Lovric et al. 2013; Wang et al. 2015) and short-term planning as well as modelling for OD-matrices (Chen et al. 2011; Tavassoli et al. 2018; Viggiano et al. 2017). Eventually, these collected, processed and analysed data can be used to make prediction of individual and group passenger behaviour based on different attributes like time and location using statistical evaluation methods as in (Zhao et al. 2018).

With reference to public transit performance indicators and efficiency, Trepanier and Morency (2010) present supply, demand, average mode occupancy and capacity indicators that are calculated and compared based on smart card data collected in Quebec. Additional references, field studies and implementation projects are, e.g., Moore and Giuliano (1998), Chapleau et al. (2008), Trepanier et al. (2009), and Kusakabe et al. (2010). Formal data management issues can be found, e.g., in the collection ITSO 1000-0-9 (2010).

## 2.2 Interoperability

Yoh et al. (2006) provide a literature review in the framework of a regional case study in California centering on interoperability issues. They evaluate the difficulties regional implementations are facing when setting up a joint ITS. They conclude that the technical framework (e.g. available standards) for operating interoperable smart ticketing is well established while the institutional framework (organisational, financial, acceptance issues) poses a significant risk for compatible adoption. In this context, Iseki et al. (2008) and Cheung (2006) try to show the financial impacts of interoperability in the framework of a cost-benefit-analysis from projects in the USA and the Netherlands, respectively. Other works on cost-benefit-analysis also try to move towards socio-economic issues; see, e.g., Welde (2012). However, the main drivers behind research on smart card (data) interoperability and standardisation are organisations established by public institutions and authorities. This can be observed by the comprehensive study on interoperable smart ticketing issues conducted by the Transit Cooperative Research Program (TCRP 2006) based in the USA which assesses the data, technical, institutional and security dimension of interoperable smart ticketing, overall. The UK Department for Transport (DfT) (2009) has a similar approach towards integrated e-ticketing by presenting the initial proprietary situation in the UK and afterwards developing economic and political measures to

promote interoperable and open smart ticketing adoption by operators and passengers nationwide. Similarly, van Oort et al. (2015) present the current state in the Netherlands and analyse sources of unreliabilities and bottlenecks in the common nation-wide smart card data system. They develop a software tool for monitoring the public transit network of all operators based on smart card data and describe potentials for improving the public transit service. Overall, it is well known that smart cards and smart card data can be beneficially used to detect travel patterns, improve ticketing, support long-term transport planning and observe and analyse network performance; see, e.g., Blythe (2004), Briand et al. (2017), Utsunomiya et al. (2006). In addition, reference to financial obligations may become a major issue (Buchholz-Stepputtis and Voß 1999).

### 3 Interoperable smart card data

The contribution of this paper is to devise a common conceptual framework based on currently available technical standards and implementation procedures that can be generally applied to collect, share and utilise smart card data interoperably between different public transit entities. It integrates best practices already formulated in literature and implemented in practice to be used as guidelines to design and operate interoperable smart ticketing systems. The necessity of the aforementioned model arises due to the following gaps identified in the literature review above: A notable practical application of academic smart card data models on continuous public transit planning cannot be observed. In addition, there is a lack of substantial discussion of practical interoperability issues and, particularly, the meaning and context of interoperable smart card data and their handling. Thus, a need to link technical and institutional practices of interoperable smart card data administration and exchange is identified. These concerns are major topics in the following, balancing practical implementation and theoretical analysis while providing a fundament for interoperable smart card data utilisation.

This section presents a procedure on how to handle and organise smart card data in a standardised way making data commonly available for the different types of model application and data mining tools. The ISCDC has a two-fold structure integrating the administration of smart card data with interoperable operations of an e-ticketing infrastructure. Thus, before presenting the ISCDC, Sects. 3.1 and 3.2 discuss the central aspects and issues of these two components separately. This supports the understanding of the underlying mechanisms and provides the necessary background for the integration of an interoperable smart card data utilisation. For this reason, a general overview of smart card data is given that helps to apprehend the requirements on data per se that are identified during the analysis of the research models. Afterwards, the term *interoperability* is defined to enable a general understanding in the context of public mass transit. Finally, these two aspects are integrated to formulate the ISCDC.

### 3.1 Smart card data

Smart cards enable a way of wide-ranging data analysis. The type of analysis is dependent on the various kinds of available and collected data sets. Smart card data can be comprehended as delivering information for an automated continuous travel survey where the data fulfil the basic requirements of being complete, linking transaction data to detailed operational data and connecting users to specific points of interest (Chapleau et al. 2008).

To fulfil the mentioned requirements of interoperable smart card data, Table 1 provides an overview of the most common data collected and processed in literature, field studies and implementation projects (compiled from Hong et al. 2016; Kusakabe et al. 2010; Chapleau et al. 2008; Trepanier et al. 2009; Moore and Giuliano 1998). Despite the potentials concerning data collection, some limitations have to be noted. For example, it is not or only rudimentarily possible to receive direct information about the passenger's socio-economic status, his trip purpose or the origin and destination of his overall journey. Still, methods have been developed to estimate these data using existing passenger statistics (Lee and Hickman 2014). In addition, if the card is linked to personal data like the name and address (depending on whether the user is registered or anonymous), conclusions can be drawn from these records. In some cases where the passengers are not required to check-out, the alighting stop is not registered which is essential for a thorough application of data models and, therefore, must be (quantitatively) estimated (Alsger et al. 2015; Tavasoli et al. 2018).

One of the main advantages of collecting passenger data with smart cards is the ability to analyse them on a disaggregate level in addition to a facilitated collection and evaluation on an aggregate basis (Bouman et al. 2012). The facilitated utilisation of aggregate data stems from the omission of standardised surveys and observations conducted manually by operators. On the one hand, this relates to data about travel time, making it possible to assess usage profiles for different points of time and average travel patterns as well as to provide a general public transit market representation (Ma et al. 2017). On the other hand, smart card data enable the understanding of passenger flows between different points of interest based on aggregate spatial data (White et al. 2010). Disaggregate data for the analysis of individual travel patterns bear the opportunity to recognise passenger groups and solve issues regarding the classification of trips as linked or return trips (Trepanier et al. 2007; Nassir et al. 2015). Finally, smart card data make an accurate progression of passenger flows per period feasible where it is possible to analyse data, e.g., by overall usage, seasonal usage or trip rate by user (Trepanier and Morency 2010).

### 3.2 Interoperability in public mass transit smart ticketing

The fundament of ITS-interoperability in public mass transit is depicted by universal *standards* which define uniform procedures and methods for generic applications. *Specifications* incorporate the standards into a specific industry

**Table 1** Smart card data and classification

Journey-related data	Operator-related data	User-related data
Data describing a journey of a passenger	Data concerning the operator's fleet and equipment	User- and card-specific data
Transaction-ID: Continuous numbering of an individual event or transaction linking all data		
Departure location (spatial)	Route number/direction	Card-ID/Security-key (card information)
Alighting location (spatial)	Vehicle number/Driver number	User data (personal information, stored centrally)
Time and date of network entry and exit (temporal)	Terminal number	Fare product issued as entitlement
Modes of transportation used	Stop/Vehicle location (GPS-based)	Validity of card and fare
		Transaction type (pre- or post-paid)

framework for a purpose-built user application. In this context, interoperability can be defined as follows:

Interoperability in the framework of ITS in public mass transit describes

- a set of institutionally applied standards and specifications by public transit entities
- to enable seamless travel for passengers using one compatible medium across different operators on different networks (cf. CEN 2003; NPRA 2005)
- with integrated data exchange and coordination between stakeholders (cf. TCRP 2006).

This definition includes three layers of interoperability which can be achieved consecutively (compiled from The EC Smartcards Study Consortium 2011; CEN 2003):

Layer 1: *Technical compatibility* (interoperability of medium)

Layer 2: *Acceptance* (interoperability of application)

Layer 3: *Validity* (interoperability of product)

In the first layer two pieces of equipment by different operators must be technically fitted to each other, e.g. by transmitting signals on a shared frequency. Secondly, the applications must be allowed to communicate with each other, e.g. by containing the appropriate security keys for communication. On the last layer, the product stored on a smart card has to be valid on the approached public transit mode by validating it with a terminal enabling the user to check-in or check-out interoperably. Finally, the infrastructure in which smart card data is generated, collected and processed creates specific roles for the various tasks and responsibilities of data handling in an interoperable environment. Figure 1 shows the roles and their tasks within an interoperable ITS infrastructure. It classifies them into the front-end which includes all card and user centric operations and the back-end which manages a data warehouse for all incoming data resulting from customer transactions and measuring equipment. It is a central database where all information is processed to be used for the analysis and administration tools applied by operators and agencies. In this context, an *application* is regarded as the platform for interoperable communication, validation, clearing and travel entitlement, used by different roles and stored on an appropriate medium.

Standards and specifications have been developed for a joint framework of definitions and assignments of roles (e.g. NPRA 2005; VDV-Kernapplikation 2010; Scholz 2012; Calypso 2010; KCEFM Kompetenzzentrum 2011). The roles depicted in Fig. 1 are interacting on different levels depending on the current task that must be processed. The figure illustrates the necessity for a standardised and interoperable data exchange to administer fare management, user data, transactions and fleet data coming from a multitude of public transit operators. Thus, the following description of the roles shown in Fig. 1 discusses the various tasks (depicted on the arrows) that must be performed to interact within an ITS to make it operational:



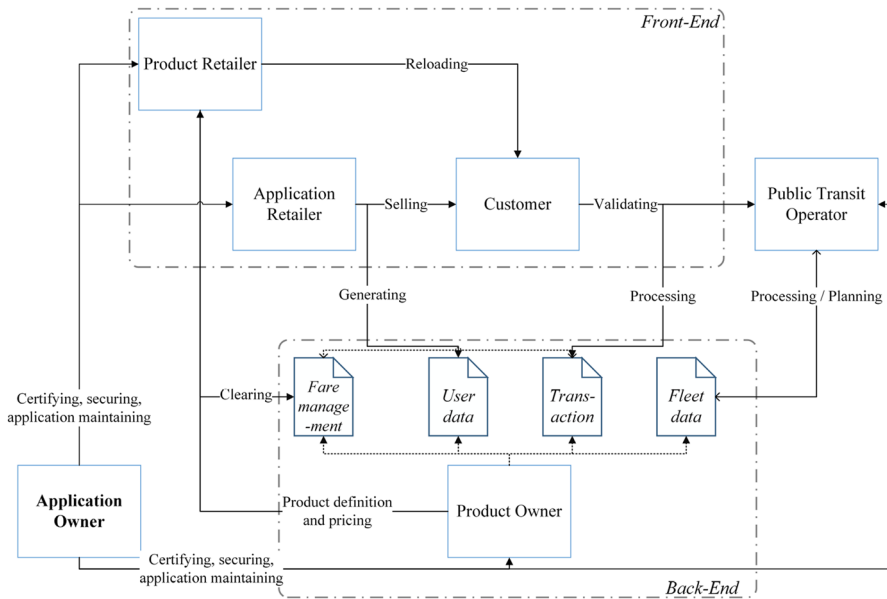
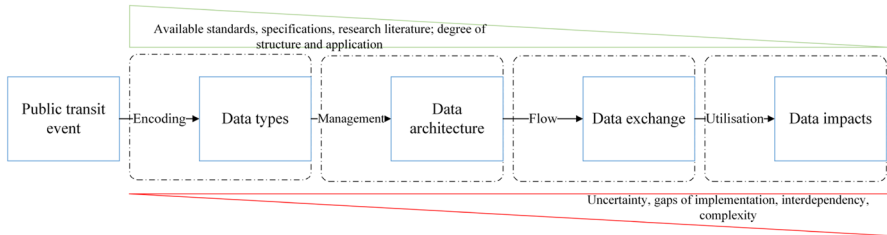


Fig. 1 Interoperable ITS role model

- *Customer*: contractual partner of the ticket entitlement, utilises the public transit service, responsible for payment and balancing his credit
- *Public Transit Operator*: provides the public transit service to the customer, main body for generating and transferring operational customer data
- *Product Owner*: operates the back-office with clearing and data management responsibilities, responsible for fare structure and ticketing system
- *Product Retailer*: contractual partner of the ticket entitlement, charges the customer's smart card and bears the risk of payment, provides customer service
- *Application Retailer*: issuing the application in form of a smart card or similar medium to the customer, establishes a customer account and generates initial customer data
- *Application Owner*: develops, enforces and maintains the application, certifies other roles, constitutes the supreme instance in terms of interoperability and issues evolving from it

The role model illustrates that an ITS and the surrounding infrastructure becomes interoperable by integrating an application owner who is identified as the interoperable element. Through the application, he coordinates and manages the interfaces between each stakeholder making him the premise of interoperable smart ticketing. Note that an entity can occupy one or more roles.



**Fig. 2** Interoperable Smart Card Data Chain

### 3.3 Interoperable smart card data chain

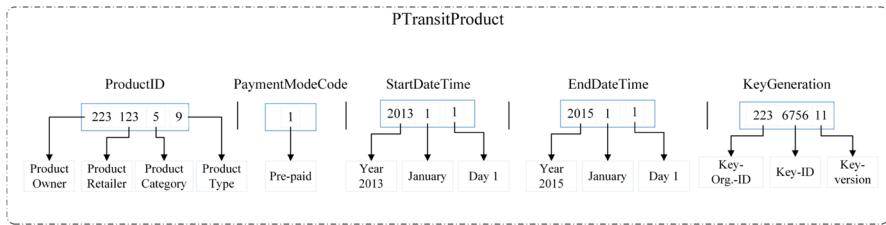
This section integrates the technical smart card data (Sect. 3.1) with interoperability (Sect. 3.2) to provide a holistic conceptual model for smart card data management and utilisation which is, to the best of our knowledge, neither present in academic literature nor in practical applications. For this reason, the Interoperable Smart Card Data Chain ISCDC is developed (see Fig. 2).

The ISCDC describes a procedure on how to standardise smart card data and manage them interoperably. Based on the requirements stated before, the ISCDC is a tool that supports the universal execution of data mining and model applications for planning and clearing purposes for all types of public transit entities, systems and modes. It ensures that the appropriate smart card data quality for applying the models is attained. The ISCDC consists of four steps once the data processing is initiated by a public transit event (e.g. a check-in at a terminal).

The subsequent paragraphs discuss each of the steps of the ISCDC separately. Particularly, EN (European norm) standards provide a basic source for setting the framework for interoperable public transit which is the fundament for smart card data sharing. The relevant and main ideas of these standards are incorporated in the descriptions for each step of the ISCDC. Additionally, this procedure helps to assess the amount of standards available for a certain ISCDC step in order to identify gaps that must be closed by further scientific research and technical literature. Finally, a loose example is applied in each step to offer a generic methodology on how to implement each step based on these standards. The same example is traced throughout all steps of the ISCDC.

#### 3.3.1 Data types and encoding

Interoperable smart card data is characterised by being a highly augmented data set as a result of enhanced and integrated data mining and sharing by a group of public transit operators (cf. van Oort et al. 2015). Therefore, the ability to be centrally stored and processed within a common infrastructure is the fundamental concept of them. The first step in organising smart card data is to provide a common procedure to encode the data generated at the front-end. Each data group (object) is comprised of codes and identifiers (IDs) which, in turn, are comprised of semantic and basic



**Fig. 3** Example of an encoding of a public transit product

data. The objects, with specific attributes assigned, depict one central instance of an ITS. The object *PTransitProduct* is used to illustrate exemplarily which attributes must be included to define a public transit product set up by the product owner to be issued as public transit entitlement to passengers.<sup>1</sup> It incorporates elements of the VDV-Kernapplikation (2010), NPRA (2005) and structures of ITS0 1000-0 (2010). For reasons of clarity, the code is not presented but only an exemplary interpretation of the elements that must be encoded to provide the needed information for smart card data encoding (see Fig. 3):

The first four elements in Fig. 3 identify the product by generating a unique ID. It is composed of an agency's ID defining the product and one selling it. The generic product category and the specific product type finalise the object's identifier or label. The next seven elements constitute the data set, i.e., the content of the object. In this case the payment code is one property of the product where its status can be assigned from a list of status enumerations. Moreover, the product is valid for 2 years after which it has to be revised by the product owner. The last three elements secure and authenticate the data structure by stating an organisation responsible for issuing the key and providing a key-ID with a continuous version number. After encoding the *product* object and its attributes it must be embedded within the application. For this purpose, the *product* has to be linked to the *entitlement* object, which constitutes the issued product to the customer.

Next, the encoded data must be stored and arranged within the application of the smart ticketing system. From the perspective of interoperability, it is advisable to form independent modules with common interfaces to allow an intra-smart card communication. Consequently, the data sets can be directly read and written by all participants in compliance with their designated access rights. This disaggregation of stored data enables the separate handling of public transit and payment data by transferring them to the relevant organisation directly (TCRP 2006). Table 2 illustrates the data storage and transfer inside a module structure.

The first module, as the central identification instance of the customer, can be retrieved by all licensed partners engaged in the interoperable agreement. The Holder-ID serves as a customer number identifying the passenger and the

<sup>1</sup> ISO 8824 provides the relevant syntax (Abstract Syntax Notation One) to represent the respective data elements.

**Table 2** Interoperable data storage and transfer within smart card application modules

Identification module	Transaction module	Public transit module	Third-party module
Independent application modules stored separately on the smart card			
Card-ID	Transaction-ID	Operator-ID	Partner-ID
Holder-ID	Mode of payment	Journey-related	Partner certification
	Stored value	Operation-related	Stored partner product
	Bank account data	User-related (e.g. entitlement)	Loyalty value/purse

Security (Secure Application Module): Key management, data access rights, card authentication

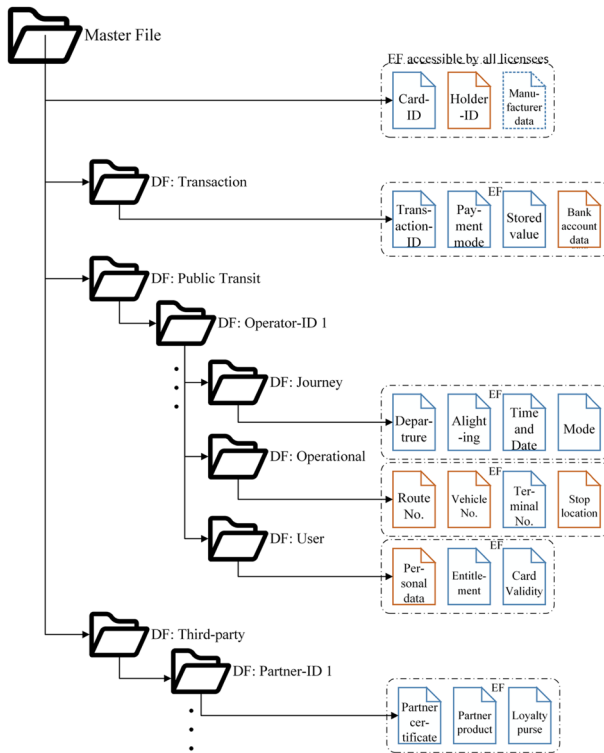
**Table 3** Basic back-end-centric data

Configuration data	Contractual data
Specification and standards (documentation)	Organisational agreement data (role definition)
Technical manuals of components and equipment	Certificates and licences
Application operating system (OS) data	Financial data for funding
Application programming interface (API) data	Clearing ratio
Blacklist (list of blocked users, see card validity)	Data ownership and location
Master and application keys	Access rights and responsibilities
	(Integrated) fare management data (list of offered fare products)

corresponding personal data while the Card-ID is used for an anonymous tracking of events respecting passengers' privacy. This module is always linked to the transaction currently performed. Therefore, the transaction module contains the Transaction-ID to associate a specific event with the relevant data stored, changed or deleted. Moreover, it holds all monetary data with information about the mode of payment. Interoperable travel data can be embedded in the public transit module where they are organised separately according to the operators gathering them. Another module can be established for third-party partners storing their products, loyalty programmes and entitlements on the card. All modules work on the basis of a secure access module (SAM) that authenticates every data access by terminals, manages the keys for encryption of the particular module and validates the access rights of the organisation writing and reading data on the smart card (cf. Calypso 2010; Blythe and Carr 2004).

The back-end becomes the central element of an interoperable ITS with increased complexity. Thus, it is essential to address a further data type classification in addition to the technical categories stated before. In an isolated smart ticketing system, customer and travel data as well as an optional purchase record are collected at the front-end and transferred to the back-office afterwards. Hence, *front-end-centric data*<sup>2</sup> are the main source of communication and information of the operator's service. By building an interoperable ITS with many organisations to coordinate, a new type of data becomes at least as important as front-end-centric data. As a result of the agreement concluded by agencies, operators and third-party partners, contractual data has to be made available to all participants to be aware of the scope of cooperation and the responsibilities held by them. Specifically, data about the clearing mechanism, rights, roles and funding need to be logged to allow an automatic transfer of transactions between partners. Additionally, configuration data for installing the joint application on smart cards, setting up interoperable components and

<sup>2</sup> Note that this does not imply that this data is purely stored at the front-end, i.e. on the smart card. Rather it distinguishes between data collected at the front-end that is transferred to the back-office and back-end-centric data that deals with data not emerging from the front-end operation. Thus, front-end-centric data is stored in the back-office, too, and not all front-end-centric data are stored on the smart card.



**Fig. 4** Front-end-centric file architecture according to ISO 7816-4

equipment and running the back-office are other necessary elements to consider. As these data sets affect the passengers only indirectly, not influencing their travel behaviour in the first place, they can be referred to as *back-end-centric data*.

Table 3 gives an overview of basic back-end-centric data types to be filed in the back-office. It is not possible to provide a complete list as they always depend on specific contracts or agreements confirmed by all parties involved in an interoperable ITS. Thus, Table 3 serves as a template with the option to adjust or add data types according to the organisations' contract (TCRP 2006).

### 3.3.2 File architecture and data management

The next step in the ISCDC is to organise the encoded form in a standardised file structure with corresponding commands to access them by agencies and operators. ISO 7816-4 fulfils this requirement for smart cards in general, which can be adopted for public mass transit. Hereafter, this file architecture is employed on the module structure proposed before. Two types of files exist to structure smart card data: *Dedicated Files (DF) (Directory)* which contain *Elementary Files (EF)* that incorporate the data stored. A special type of DF is the Master File which represents the root file

for all other directory files. Figure 4 applies the standardised file architecture on the front-end-centric interoperable module structure from Table 2.

The blue files indicate that data is stored at the front-end (i.e. smart card) as well as at the back-end. In contrast, data in the orange files are transferred immediately or linked to the back-office without storage on the smart card (for reasons of privacy and storage size the file contains only a link to the respective data in the back-office). The Transaction-ID and the DF *Journey* are stored temporarily on the smart card as a log of a specified number of last transactions to facilitate the reconstruction of customers' recent public transit events. The Master File integrates the identification module of Table 2 to be accessible by all registered partners. In addition, it contains specific card manufacturer data like application configuration and encryption data for operating the smart card properly. The other modules are depicted by DFs representing one module each. The Public Transit module/DF and Third-party module/DF contain sub-DFs for each organisation participating in the interoperable ITS. The presented file architecture<sup>3</sup> allows to define methods for data retrieval based on referencing according to file identifiers and names or stating the paths to a file.

After establishing a shared basis for a smart card data file structure, the data management by organisations engaged in interoperable smart ticketing (a topic hardly addressed in literature) becomes indispensable. The outline is built on a mutual agreement between all participants. It defines the previously mentioned organisations' data access rights and responsibilities which are organised in a rights and responsibility panel accessible by all members. It has to be decided which data is stored centrally in the back-office and which data is limited to a decentralised access by an agency which needs to transfer the relevant data upon request. From the perspective of back-end-centric data, the partners have to agree on how data ownership is managed and who is liable for fraudulent data and data misuse, especially regarding privacy issues (TCRP 2006).

When establishing the panel, the following characteristics of roles need to be considered: the public transit operator is focused on optimising his fleet by gaining access to the operational data he generates and handling the customer validation. The product owner is able to access a large share of data sets as he is the main body organising and managing the ITS while also profiting from the operational data gathered by smart cards to develop a tailored fare policy and pricing. On the other hand, both types of retailers are mainly focused on front-end-centric data being the organisational interfaces between the customer and the ITS operators. The customer is allowed to access almost all data generated by him to track his journey records and payment receipts. However, he is limited only to his personal data reducing substantially the data set he receives. The application owner is the technical instance, highly back-end-centric, overseeing the seamless technical integration and communication of all participants to make and maintain the system interoperable.

---

<sup>3</sup> The file architecture presented here corresponds to a generically applied structure for open specifications examined in this paper. The content of the DFs and EFs may vary depending on the relevant specification. The file structure in this case, is based on the findings and propositions from the previous sections.

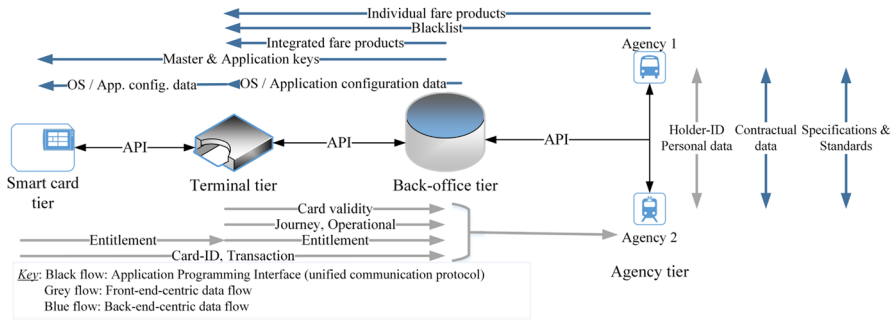


Fig. 5 Interoperable data flow

Note that the rights and responsibilities vary according to the organisations' distribution of tasks within their network. Especially, it is possible that work packages are shifted from the product owner to the product retailer depending on the organisational structure of the local public transit network.

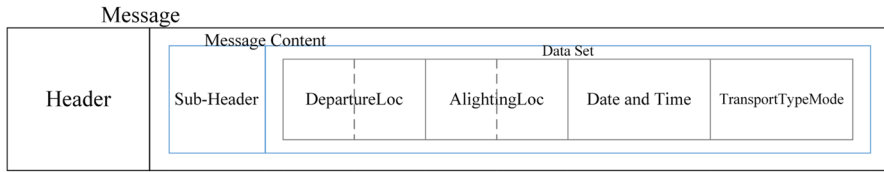
### 3.3.3 Data flow and exchange

For the purpose of data flow and exchange the relevant interfaces and the means of data transfer through them are analysed. First, the different tiers of data flow must be examined. The components of an ITS are used for classification. Four tiers can be identified where communication takes place.

1. *Smart Card Tier*: the card itself with the relevant front-end-centric data stored on it
2. *Terminal Tier*: instance connecting the front-end and back-end
3. *Back-office Tier*: data warehouse as central data communication instance
4. *Agency Tier*: handling of sensitive data that are not regarded as openly available through the back-office for all members of the ITS

Figure 5 (based on TCRP 2006) shows an idealised setup and the corresponding data flow with different stream types. Two types can be observed with front-end-centric data being transferred from the front-end to the back-office and back-end-centric data flowing vice versa. Furthermore, it is evident that data elements are not transferred through all tiers. Instead, they only appear within different stages depending on their properties and relevance. All contractual data and passengers' personal data are only processed inside the back-office to maintain the compulsory security and privacy standards. The technical basis is depicted by the black flows in Fig. 5. In an interoperable e-ticketing framework common application programming interfaces (API) are defined allowing the embedding of specific components' applications into the entire system. Note that in Fig. 5 the agencies only communicate via the back-office without establishing a separated API between them. Thus, the back-office is not only considered as central data warehouse but also as a standardised





**Fig. 6** Visualisation of message structure

tool for inter-agency communication. Decentralised data held by agencies can be transferred upon request through the back-office. The same is true for data that have to reach the front-end.

The streams of data flow show the places and directions of data movement in an ITS. Subsequently, types of data exchange can be addressed on this basis. Primarily, two main exchange purposes arise:

1. For *Clearing*: payment information and stored value on the smart card are collected and processed to assign the right revenue share to each agency. The practical execution defines claims and liabilities of each organisation over a specified period after which a net settlement is calculated and transferred to the benefiting organisation.
2. For *Operation*: information on fleet efficiency, ridership and public transit usage is traded according to the privacy policies to optimise the joint interoperable fare management and smart ticketing network.

The smart card data exchanges can be encoded analogous to the Abstract Syntax Notation One as described above for the data encoding. Figure 6 visualises exemplarily the interlaced message structure for an exchange of information for operational purposes between agencies. Hence, this sample focuses on transferring journey-related data from one agency to another (compiled from NPRA 2005; VDV-Kernapplikation 2010; EN 1545-1). Overall, messages are structured along three main layers (based on ITSO 1000-9).

Before the actual data is transferred, an initial header is defined to state the necessary information for identification of type, time, sender and receiver. The header specifies the exchange form (either as a specific file or message containing the relevant information), generates a continuous Message-ID and indicates the key required to enable the organisations to read the message. The key defines by whom the transferred data can be accessed in the back-office. It is possible to make them available for all organisations (encrypting with one globally shared key), restrict them to a certain pre-defined group (specific groups of organisations holding a shared key) or make them confidential where the data is only suitable for the addressed entity (one key shared between sender and receiver) (ITSO 1000-9 2010). Afterwards, the actual information is encoded and passed. For this example, the front-end-centric EFs in the DF *Journey* outlined in Fig. 4 are sent from one agency to another. It contains a sub-header and the relevant data set. The sub-header denotes the circumstances in which the data were generated by naming the Transaction-ID, participating operator

and Card-ID for this transaction. The second element is the specific data set about the journey. The departure and alighting locations (Location-ID) are composed of an identification number and the enumeration of points of interest. The same is valid for the mode of transport which is represented by an enumeration of transport means complying with the very same standard.

### 3.3.4 Data utilisation and impacts

The last step in Fig. 2 highlights the general core of smart card data and examines the identified benefits and obstacles in the framework of interoperability. The aim is to derive opportunities for an integrated data utilisation that can impact the integrated network design and coordination of participating organisations.

As mentioned above, the literature reviewed for this paper does not address an interoperable utilisation of smart card data in public mass transit. Therefore, an analysis of data impacts on interoperability is not existent. Thus, this paper initiates the discussion on these issues by providing recommendations for the first step of a three-step procedure to enable an interoperable smart card data utilisation. The last two steps have to be approached individually by the user on the basis of his specific outcome of suitable smart card data models:

1. Review of existing smart card data models outside interoperability. Subsequently, potential impacts and applications of these models on interoperable smart ticketing are deduced
2. Based on these findings the relevant data models are advanced and extended to integrate them into an interoperable ITS
3. A joint infrastructure is set up to facilitate the practical implementation of the proposed data models

Smart card data facilitate the coordination of an ITS by dovetailing the participating operators helping them to adapt to the needs of an interoperable smart ticketing system. Furthermore, if operators focus on similar evaluation tools the data quality can be enhanced significantly by combining and sharing resources (processing power of the information and communication (ICT) infrastructure), know how (application and proper handling of various data models by different organisations) and customer data. Synergies between participants can lead to the build-up of a large analytical tool set that improves the operational efficiency and customer convenience beyond the possibilities of an isolated operator.

Table 4 addresses the first step by exemplarily selecting publications from the three categories of possible data utilisation identified in the literature to illustrate the implications of standardised and interoperable smart card data. Primarily, a more detailed description of the particular focus is given.

Afterwards, potential impacts of data utilisation in an interoperable framework are deduced.

Considering clearing aspects, smart card data can aid in concurring on common clearing ratios by precisely analysing which operator bears the greatest passenger load

**Table 4** Implications for interoperable smart card data utilisation and impacts

Research publication	Research focus and smart card data application	Implications for interoperable smart card data utilisation and impacts
Leung (2000)	<p><i>Marketing metrics and tools:</i></p> <p>Conjoint analysis to determine customer preferences on smart card design and features</p> <p>Customer relationship management (CRM) tools to increase customer retention</p> <p>Identification of smart card features that boost smart card adoption and acceptance</p>	<p>Perform joint analysis to agree on common interoperable smart card design and features</p> <p>Perform joint analysis to agree on common interoperable fare structure design</p> <p>Share various data of operators and agencies to provide tailored advertising and products for passengers (augmented data pool for targeting the right customers)</p>
Trepanier et al. (2009)	<p><i>Public transit performance indicators and efficiency:</i></p> <p>Development of public transit public supply and demand indicators</p> <p>Evaluation of mode occupancy and capacity</p> <p>Data used: Card ID, Fare Type, Time, Route ID, Boarding and Alighting stop, Validation info, Operational info (e.g. vehicle number)</p>	<p>Identification of operators bearing the highest passenger capacity load with the possibility to offset load to other operators, thus, smoothing the capacity between them. This leads to a balanced public transit network and more convenient passenger journeys</p> <p>Matching different operators' supply with passengers' demand</p> <p>Provide implications for revenue distributions by analysing which operator(s) contribute(s) the most to passenger travel</p>
Kurauchi et al. (2014)	<p><i>Travel patterns and passenger behaviour:</i></p> <p>Examination if passengers follow hyperpaths when choosing their bus line in London's Oyster ITS</p> <p>Bus lines only are chosen due to different generation of data sets on different modes</p> <p>Data used: Card number, Route ID, boarding time</p>	<p>Extension to all public transit modes by harmonising the data sets generated on each mode</p> <p>ISDCDC ensures a complete data set by incorporating all data identified important in literature and ITS projects</p> <p>Easy implantation of standardised data filters to reduce sample according to study needs</p> <p>Identification of transfer trips becomes easier as the entire network is covered, thus, leaving no gaps when tracing passengers' travel patterns</p> <p>Easy transfer of study to other cities to verify results with the same data types and quality</p>

or which operator is able to target the most customers in terms of public transit and smart card activity. This supports the mitigation of operator concerns about a fair revenue distribution.

In contrast, interoperable data utilisation for operational purposes poses significant risks and constitutes organisational issues that can potentially hamper a wide adoption. Considering that these data represent sensitive information regarding public transit operations, it is questionable how keen operators will be to jointly share efficiency and passenger metrics with their competitors. Additionally, agencies or operators already incorporating data models to analyse operational efficiency are reluctant to transfer their know-how to other organisations being afraid of losing their competitive advantage. These worries are amplified by privacy issues concerning the sharing of personal customer data between operators to create movement profiles and expand tailored advertising from the perspective of customer relationship management (CRM).

Another issue is the content-wise compatibility of collected data. The technical and formal compatibility has already been established by the preceding discussion in this paper. However, it does not guarantee that certain data is equally important to different operators. For instance, train operators are, amongst others, interested to estimate the choice of train by passengers based on the minimisation of time between checking-in at a terminal and boarding the train (Kusakabe et al. 2010). This time interval has no relevance for bus operators as passengers check in after boarding the bus. Therefore, operators in an interoperable ITS face a highly increased data load in which it is critical to sort out relevant data for own data utilisation. This leads to an increased demand of analytical resources while it is also difficult to define relevant criteria for measuring operational efficiency and passenger travel patterns for the specific situation. Overall, the complexity of data handling increases considerably binding capital and other resources potentially impacting corporate performance negatively.

Finally, an interoperable ITS is composed of a sizeable number of different stakeholders engaged in a contractual agreement. The decision-making process is prone to delays and complex coordination procedures. Changes of schedules or fare types require an adjustment of the interoperable ITS depending on the degree of (fare) integration. As a result, short-term planning on the basis of supply and demand indicators is not feasible in a network of this kind. These issues can only be incorporated in a strategic evaluation of the network design when assessing customer behaviour and forecasting the demand. Schedule adjustment according to passenger travel patterns is only possible on a tactical level where the business processes between organisations are well defined and clearly structured, meaning where interoperable e-ticketing is mature and established over an extended period of time.

## 4 Discussion of ISCDC and expert input

### 4.1 Expert interview aim

Expert interviews were used to align the conceptual framework of the ISCDC with insights from representative industry stakeholders and to see whether the ISCDC

addresses some of the issues encountered in application. The expert interviews support the assessment of the practical significance of smart card data and their interoperable dimension. While the chosen experts are high-level representatives of the e-ticketing landscape in Germany, speaking for a multitude of public transit operators and organisations, the insights received from these interviews serve only as auxiliary input to refine the general ISCDC devised from available standards and specifications.

## 4.2 Expert panel and selection

Mr Steffen Bork, project leader e-ticketing, HVV (Hamburger Verkehrsverbund), and Mr Sjeff Janssen, director of the VDV-Kernapplikation (VDV-core application, VDV-KA) were chosen. The VDV-Kernapplikation is the sole industry association for implementing an interoperable smart ticketing system between the major public transit operators in Germany. VDV-KA is the application owner developing, maintaining and certifying the interoperable e-ticketing application used by more than 200 public transit organisations in Germany. Hence, Mr Janssen being the director of this association can provide a representative and overall view on efforts by all public entities in Germany collaborating in e-ticketing. In essence, this expert was selected because he directly relates to the conceptual research of this paper from a wide-ranging industry perspective.

As the HVV is one of the major public transit associations in Germany combining over 30 public transit operators (some of them also organised in the VDV-Kernapplikation), Mr Bork, the responsible project leader for this entire system seems appropriate to give a representative account on the integration efforts done by the all entities involved. He was selected because he is the chief coordinator of 30 public transit operators being able to give a holistic view of a multitude of key players involved in collecting, sharing and utilising smart card data and ticketing systems.

## 4.3 Expert interview methodology

Both interviews had been conducted during this research project in a semi-standardised problem-centric way. Mr Bork was interviewed in person while the interview with Mr Janssen was conducted via phone. We developed the questionnaire based on issues of interoperability identified in literature and standards during the development of the ISCDC. Expert interviews were preferred as the increased complexity of the examined topic demanded an in-depth approach by interacting with representatives from central organisations that are operating, advancing or maintaining a smart ticketing system.

Subsequently, key insights by the interviewees are stated supplemented by the relevant conceptual contribution by the ISCDC.

1. Understanding the scope of the term interoperability in public mass transit to pursue the idea of standardised smart card data:

*Interviews:* In this regard, the interviewees agree that it must be distinguished between technical and institutional interoperability. While the former is well established through standards and specifications the latter has not been widely adopted yet. According to Mr Janssen, this is a particular result of the generally slow technical adoption of e-ticketing in the beginning, obstructing a homogeneous framework of interoperability in Germany. A central aspect is the elimination of fragmentation to make data sharing and interoperable utilisation feasible. In consequence, a holistic view on smart card data utilisation and resulting impacts is not possible at the moment.

*ISCDC:* The ISCDC offers a connection between the technical part (Data Types and Encoding, Data Architecture and Management) and the more institutional part (Data Flow and Exchange, Data Utilisation and Impacts) by linking them through a formal structure covered by necessary technical standards. Essentially, it stipulates available and relevant standards, the degree of complexity and the necessity of interoperable interaction between public transit organisations for each element of the chain. Overall, it states a conceptual understanding of interoperable smart ticketing.

2. The different data types are encoded and managed according to their designated purpose:

*Interviews:* The interviewees are aware of the fact that smart cards bear the opportunity to gather an enhanced data set especially improving master file data and close existing data gaps: They give public transit organisations the possibility to identify the customers' background by collecting and storing personal data. On a basic level they can help to statistically evaluate passenger flows and shifts in ridership, therefore closing gaps in evaluation that evolved from manual surveys and the collection of aggregate data by automatic counting devices in vehicles. In particular, these results from the utilisation of checking (check-out) mechanisms allow the ascertainment of boarding and alighting locations as well as the mode and journey time. These data are identified as central by the VDV-KA. Due to privacy constraints, data evaluation at the HVV is restricted to anonymous and statistical evaluation only.

*ISCDC:* The ISCDC integrates validated technical standards to manage the data mentioned by the experts. By employing the proposed data types modularisation and encoding procedures it can be guaranteed that relevant data is stored as intended and that privacy considerations are obeyed. This is achieved in the ISCDC by designating data stored at the front-end (i.e. smart card) and back-end according to constraints of storage size and the given priority to privacy and security (e.g. no personal data on the smart card).

3. A seamless sharing and utilisation of data between cooperating agencies and operators is dependent on the acceptance of the technical framework and establishment of the institutional framework:

*Interviews:* In the HVV implementation the product retailers have been identified as central administrators of data emerging from the ITS. For this reason, they are in charge of the generated customers' personal data, transactional sales and blocking list data which are considered to be the most important ones. For these types of data an interoperable exchange exists to perform a monthly settlement while there is no operational data exchange and smart card data utilisation for a joint optimisation of the public transit network (Mr. Bork). The HVV implementation is prioritising CRM through an improved product offer and advice on the basis of an integrated smart ticketing platform while the aim is not to utilise smart card data for operational purposes.

Discussing the general situation in Germany, interoperable data exchange is still in its initial phase where public transit associations and agencies organise the utilisation of smart card data individually depending on their capabilities and requirements (Mr. Janssen). In these circumstances, the VDV-KA as application owner is able to provide basic support without interpreting this task as its core competence. It is stated that the interest in specific data depends on the entity involved: service operators are more interested in operational data than product owners to improve fleet efficiency.

*ISCDC:* The last two steps of the ISCDC formulate a flexible and generic framework that can be adapted to the current data exchange requirements of the involved entities. By being tailored to the standardised role-based-approach, each public transit entity can be integrated by assigning the respective tasks in the institutional framework according to its desired data exchange and utilisation. The separation of smart card data between front-end and back-end enables to identify the data flows and pools to aid the entities in setting up the physical infrastructure.

In summary, the two-fold structure accompanying the ISCDC can be stated as follows: The technical framework manages all equipment and interfaces to be implemented on the basis of common standards in the first two steps. The institutional framework deals with organisational aspects of data handling and acceptance barriers of agencies and customers (regarding sharing and privacy) formalised technically in the last two steps.

## 5 Conclusion and further research

One of the key issues of current smart ticketing applications is the lack of joint standards and methods on how to approach smart card data handling and processing. This paper proposes a four-step procedure to respect all necessary attributes of smart card data. Starting from the basic encoding the subsequent steps lead to the eventual goal of enabling a cross-agency universal data sharing and model application for planning purposes in public mass transit. It incorporates the current state of research literature and practical insight from the German application owner and a major association complying with the German interoperable framework. The results obtained in this paper help future research on smart card data utilisation to embrace an environment where its outcomes can be tested and implemented on a broad scale of applications. Hence, data quality and the quantity of useable data as input for

the respective research can be improved. The adjustment of the ISCDC with expert insights help to integrate the ISCDC practically into an agency's operating ITS. Thus, this paper provides an interface approach which can be adopted by theoretical research literature as well as by operating public transit organisations that aim to streamline and standardise their smart card data processing for planning and clearing purposes.

Further research needs to target the complementary analysis of the technical and institutional framework of the ISCDC. The equipment infrastructure as well as the business environment has to be evaluated to propose methods on integrating the ISCDC into an established ITS. Testing the ISCDC's suitability by fully implementing it into practical public transit planning is a further important factor to consider. In addition, research literature should initiate a discussion on the emerging opportunities and benefits of applying data mining tools, models and algorithms in an interoperable environment with an extended number of players involved. New approaches on smart card data analysis can be developed based on the enhanced availability of comparable data sets.

In conclusion, the ISCDC initiates the examination of impacts of interoperable data utilisation on theoretical model development and specific public transit planning.

**Acknowledgements** An earlier version of the paper had been handled by Prof. Matthew G. Karlaftis.

## References

- Alsger AA, Mesbah M, Ferreira L, Safi H (2015) Use of smart card fare data to estimate public transport origin–destination matrix. *Transp Res Rec* 2535:88–96. <https://doi.org/10.3141/2535-10>
- Asakura Y, Iryo T, Nakajima Y, Kusakabe T (2012) Estimation of behavioural change of railway passengers using smart card data. *Pub Transp* 4:1–16. <https://doi.org/10.1007/s12469-011-0050-0>
- Baric D, Gambetta R, Novacko L (2011) Calypso electronic ticketing technology: the contactless standard for a modern transport. *Mech Transp Commun* 3:77–83
- Blythe PT (2004) Improving public transport ticketing through smart cards. *Proc Inst Civ Eng* 157:47–54. <https://doi.org/10.1680/muen.157.1.47.36592>
- Blythe P, Carr J (2004) Challenges and opportunities for the future roll-out of transport smartcards in the UK. Transport Operations Research Group; Cassie Building, University of Newcastle, Leeds
- Bouman P, Lovric M, Li T, van der Hurk E, Kroon L, Vervest P (2012) Recognizing demand patterns from smart card data for agent-based micro-simulation of public transport. In: Proceedings of the 7th Workshop on Agents in Traffic and Transportation, Valencia
- Briand A-S, Côme E, Trépanier M, Oukhellou L (2017) Analyzing year-to-year changes in public transport passenger behaviour using smart card data. *Transp Res Part C* 79:274–289. <https://doi.org/10.1016/j.trc.2017.03.021>
- Buchholz-Stepputtis K, Voß S (1999) Finanzdienstleistungen, Zahlungsverkehr und Electronic Commerce im Internet mittels SmartCards. In: Hesse H, Rebe B (eds) *Vision und Verantwortung*. Olms, Hildesheim, pp 445–463
- Calypso (2010) Calypso handbook. Calypso Networks Association. <http://cna.adcet.com/index.php/downloads/category/5-free-downloads?download=19:calypso-handbook>
- CEN (2003) Facilitating smart card technology for electronic ticketing and seamless travel—Part 2: development of smart card based interoperable ticketing systems (CWA 14838-2). European Committee for Standardization. <https://joinup.ec.europa.eu/solution/facilitating-smart-card-technology-electronic-ticketing-and-seamless-travel-part-2-development-smart/about>



- Chapleau R, Trepanier M, Chu K (2008) The ultimate survey for transit planning: Complete information with smart card data and GIS. In: 8th international conference on survey methods in transport: harmonisation and data comparability, Annecy
- Chen X, Guo S, Yu L, Hellinga B (2011) Short-term forecasting of transit route OD matrix with smart card data. In: 14th international IEEE conference on intelligent transportation systems, Washington. <https://doi.org/10.1109/ITSC.2011.6082929>
- Cheung F (2006) Evaluation and cost-benefit analysis: public transport smartcard in the Netherlands. AVV Transport Research Centre Ministry of Transport, Public Works and Water Management
- Goulet-Langlois G, Koutsopoulos HN, Zhao J (2016) Inferring patterns in the multi-week activity sequences of public transport users. *Transp Res Part C* 64:1–16. <https://doi.org/10.1016/j.trc.2015.12.012>
- He L, Trepanier M (2015) Estimating the destination of unlinked trips in transit smart card fare data. *Transp Res Rec* 2535:97–104. <https://doi.org/10.3141/2535-11>
- Hong SP, Min YH, Park MJ, Kim KM, Oh SM (2016) Precise estimation of connections of metro passengers from smart card data with smart card transaction data. *Transportation* 43:749–769. <https://doi.org/10.1007/s11116-015-9617-y>
- Iseki H, Demisch A, Taylor B, Yoh A (2008) Evaluating the costs and benefits of transit smart cards. Institute of Transportation Studies, University of California, Berkeley. California PATH Research Report
- ITSO 1000-0 (2010) ITSO TS1000-0-9 Interoperable Public Transport Ticketing Using Contactless Smart Customer Media – Part 0-9: Concept and Context. v2.1.4
- Jang W (2010) Travel time and transfer analysis using transit smart card data. *Transp Res Rec* 2144:142–149. <https://doi.org/10.3141/2144-16>
- KCEFM Kompetenzzentrum (2011) Elektronisches Fahrgeldmanagement in NRW. Migration zur VDV Kernapplikation Rahmenlastenheft PV-System Funktionsbereich Berechtigungsart WEB
- Kurauchi F, Schmöcker J-D (eds) (2017) Public transport planning with smart card data. CRC, Boca Raton
- Kurauchi F, Schmöcker J-D, Shimamoto H, Hassan S (2014) Variability of commuters' bus line choice: an analysis of Oystercard data. *Pub Transp* 6:21–34. <https://doi.org/10.1007/s12469-013-0080-x>
- Kusakabe T, Asakura Y (2014) Behavioural data mining of transit smart card data: a data fusion approach. *Transp Res Part C* 46:179–191. <https://doi.org/10.1016/j.trc.2014.05.012>
- Kusakabe T, Iryo T, Asakura Y (2010) Estimation method for railway passengers' train choice behavior with smart card transaction data. *Transportation* 37:731–749. <https://doi.org/10.1007/s11116-010-9290-0>
- Lee S, Hickman M (2014) Trip purpose interference using automated fare collection data. *Pub Transp* 6:1–20. <https://doi.org/10.1007/s12469-013-0077-5>
- Leung V (2000) Improving the smart card acceptance by the public transport operation through conjoint analysis to determine the user preference. *Manag Lit Rev* 2:52–81
- Lovic M, Li T, Vervest P (2013) Sustainable revenue management: a smart card enabled agent-based modeling approach. *Decis Support Syst* 54:1587–1601. <https://doi.org/10.1016/j.dss.2012.05.061>
- Ma X, Lui C, Wen H, Wang Y, Wu YJ (2017) Understanding commuting patterns using transit smart card data. *J Transp Geogr* 58:135–145. <https://doi.org/10.1016/j.jtrangeo.2016.12.001>
- Moore J II, Giuliano G (1998) Functional evaluation of the Los Angeles smart card field operational test. *Transp Res Part C* 6:247–270. [https://doi.org/10.1016/S0968-090X\(98\)00018-7](https://doi.org/10.1016/S0968-090X(98)00018-7)
- Morency C, Trepanier M, Agard B (2006) Analysing the variability of transit users behaviour with smart card data. In: Proceedings of the 2006 IEEE intelligent transportation systems conference, Toronto
- Nassir N, Hickman M, Ma ZL (2015) Activity detection and transfer identification for public transit fare card data. *Transportation* 42:683–705. <https://doi.org/10.1007/s11116-015-9601-6>
- Nishiuchi H, Kobayashi Y, Todoroki T, Kawasaki T (2018) Impact analysis of reductions in tram services in rural areas in Japan using smart card data. *Pub Transp* 10:291–309. <https://doi.org/10.1007/s12469-018-0185-3>
- NPRA (2005) Specification for Interoperable Electronic Ticketing System, Handbook 206-3. Norwegian Public Road Administration (NPRA)
- Pelletier M-P, Trepanier M, Morency C (2011) Smart card data use in public transit: a literature review. *Transp Res Part C* 19:557–568. <https://doi.org/10.1016/j.trc.2010.12.003>
- Qin LJ, LV Y, Zhang L-N, Cai X (2011) Evaluation of the reliability of bus services based on GPS and smart card data. In: Proceedings of the 2011 IEEE international conference on quality and reliability, Bangkok. <https://doi.org/10.1109/ICQR.2011.6031695>

- Robinson S, Narayanan B, Toh N, Pereira F (2014) Methods for pre-processing smartcard data to improve data quality. *Transp Res Part C* 49:43–58. <https://doi.org/10.1016/j.trc.2014.10.006>
- Scholz G (2012) IT-Systeme für Verkehrsunternehmen. Dpunkt, Heidelberg
- Tavassoli A, Mesbah M, Hickman M (2018) Application of smart card data in validating a large-scale multi-modal transit assignment model. *Pub Transp* 10:1–21. <https://doi.org/10.1007/s12469-017-0171-1>
- TCRP (2006) Smartcard Interoperability Issues for the Transit Industry, TCRP—Transit Cooperative Research Program Report 115, Transportation Research Board of National Academies
- The EC Smartcards Study Consortium (2011) Study on Public Transport Smartcards—Final Report
- Trepanier M, Habib K, Morency C (2012) Are transit users loyal? Revelations from a hazard model based on smart card data. *Can J Civ Eng* 39:610–618. <https://doi.org/10.1139/2012-048>
- Trepanier M, Morency C (2010) Assessing transit loyalty with smart card data. In: 12th WCTR, Lisbon
- Trepanier M, Morency C, Agard B (2009) Calculation of transit performance measures using smartcard data. *J Pub Transp* 12:79–96. <https://doi.org/10.5038/2375-0901.12.1.5>
- Trepanier M, Tranchant N, Chapleau R (2007) Individual trip destination estimation in a transit smart card automated fare collection system. *J Intell Transp Syst* 11(1):1–14. <https://doi.org/10.1080/15472450601122256>
- UK Department for Transport (DfT) (2009) Developing a Strategy for Smart and Integrated Ticketing
- Utsunomiya M, Attanucci J, Wilson N (2006) Potential uses of transit smart card registration and transaction data to improve transit planning. *Transp Res Rec* 1971:119–126. <https://doi.org/10.3141/1971-16>
- van Oort N, Sparing D, Brands T, Goverde RMP (2015) Data driven improvements in public transport: the Dutch example. *Pub Transp* 7:369–389. <https://doi.org/10.1007/s12469-015-0114-7>
- VDV-Kernapplikation (2010) Hauptdokument mit Basisobjektmodell (BOM). KA-Technische Spezifikation v1.107 (KA BOM-SPEC)
- Viggiano C, Koutsopoulos HN, Wilson NHM, Attanucci J (2017) Journey-based characterization of multi-modal public transportation networks. *Pub Transp* 9:437–461. <https://doi.org/10.1007/s12469-016-0145-8>
- Wang ZJ, Li X-H, Chen F (2015) Impact evaluation of a mass transit fare change on demand and revenue utilizing smart card data. *Transp Res Part A* 77:213–224. <https://doi.org/10.1016/j.tra.2015.04.018>
- Welde M (2012) Are smart card ticketing systems profitable? Evidence from the city of Trondheim. *J Pub Transp* 15:133–148. <https://doi.org/10.5038/2375-0901.15.1.8>
- White P, Bagchi M, Bataille H (2010) The role of smart card data in public transport. In: 12th WCTR, Lisbon
- Yoh A, Iseki H, Taylor B, King D (2006) Institutional issues and arrangements in interoperable transit smart card systems: a review of the literature on California, United States, and international systems. Institute of Transportation Studies University of California, California PATH Research Report, Berkeley
- Zhao Z, Koutsopoulos HN, Zhao J (2018) Individual mobility prediction using transit smart card data. *Transp Res Part C* 89:19–34. <https://doi.org/10.1016/j.trc.2018.01.022>
- Zhong C, Manley E, Arisona SM, Batty M, Schmitt G (2015) Measuring variability of mobility patterns from multiday smart-card data. *J Comput Sci* 9:125–130. <https://doi.org/10.1016/j.jocs.2015.04.021>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.