Intermodal travel super app with agent system and data space: RealLab Hamburg implementation

Chris Schlueter Langdon1,2*, Johannes Eckert3

1. Deutsche Telekom, Germany
2. Peter Drucker School of Management, Claremont Graduate University, United States
3. Hamburger Hochbahn, Germany

*christoph.schlueter-langdon@telekom.de

Abstract
Climate change requires innovation to reduce the CO₂ emissions caused by mobility. We will describe the linking of two innovations, an agent system together with a data space, to enable a novel “super app”, which supports modal shift, i.e., the shift of trips using a personal car to intermodal travel involving the integration of public transport with micromobility, such as electric scooters and shuttles. Such mobility chains have been inhibited by the absence of corresponding data chains or data sharing due to a lack of data sovereignty. Our solution is an app for planning end-to-end intermodal trips “powered” by an agent system, which in turn sits on top of a data space to provide the necessary data. The data space utilizes technology that complies with the International Data Spaces (IDS) DIN Spec 27070 and Gaia-X objectives. Below we will report on a first case study application at RealLab Hamburg.

Introduction
How can we fix mobility? How can we reimagine it and reshape how we get from point A to point B to be more sustainable and better, of course, because very few of us would change our behavior voluntarily otherwise. Numerous studies confirm that mobility is broken, particularly in urban areas where it keeps taking longer and has become increasingly deadly for vulnerable road users, while CO₂ emissions continue to increase and exacerbate global warming. One possible solution in cities is intermodal travel, aided by the micromobility boom and its integration with public transport in particular: “[Micromobility] can perfectly complement buses and trains for the remaining kilometers to the destination. This makes public transport more attractive and can reduce car journeys” (Achim Berg, President of Bitkom, the Germany’s largest digital association, Bitkom 2019). However, this modal shift requires a change in information and business systems (Schlueter Langdon et al. 2021). For example, it is easy to see that end users would benefit from a single app or “super app” integrating different providers into a seamless mobility chain (Free Now 2022). But how can we do it? Peter Drucker, the founder of modern management, was known for giving timeless advice, such as to avoid reinventing the wheel. Instead of starting from scratch, this paper draws and expands on lessons learned from past mobility challenges, such as with air travel.

Lessons from U.S. airlines: Matching, profiler, and calculator
In 1978, U.S. President Jimmy Carter signed the Airline Deregulation Act into law (Statue-92-pg1705, link) that deregulated the U.S. airline business, allowing the entry of low-cost competitors, in turn...
 unhinging business models of existing providers. So how did they respond and survive? The legacy carriers changed their business models radically, shifting from revenue per route – still today’s dominant model of public transport (!) – to revenue per seat. What appeared to be a minor tweak in a financial spreadsheet required innovation and investment in new processes – and the systems and software to automate it all (see Schlueter Langdon 2021). This transformation of their business to revenue per seat gave birth to three types of systems (see Figure 2, Schlueter Langdon 2021): (1) airline reservation systems (ARS), (2) a loyalty system for frequent flyer programs (FFP), and (3) yield or revenue management systems (YM; for a literature review, see McGill & Van Ryzin 1999; for state of YM, see Carrier & Fiig 2018). Why these three systems? At the core of the shift to revenue per seat was the insight that, for the sake of profitability, they needed to sell every single seat at the highest price possible, essentially treating seats as perishable goods and customers as NOT created equal. From an analytics perspective, the challenge was doable: matching demand with supply, matching customers with seats. It required coupling inventory management with variable pricing: Start by using seat inventory to determine supply, using customer profiles to predict demand, and then using different price points to clear the market. Finally, learn from the results and adjust inventory next time around, for example, by using a larger plane or adding another flight. Back then, the problem wasn’t so much with the analytics but with the data – or more precisely, the lack of it. Where could you find the seat inventory for a destination? How can you keep track of different routes to the same destination? Airlines created reservation systems in order to manage seat inventory. Where can customer profiles be found to predict demand and establish price points? To be able to sell each seat at the highest price, you need insights into a customer’s willingness-to-pay (WTP). Predicting WTP, in turn, requires data on the type of travel (for business or leisure), budget (personal income or travel policy), sensitivity to travel time (daytime departure or red eye), travel duration (non-stop or stopover), convenience (economy class or business), and the context of the decision (traveling alone or with family) – and all of the above not at some aggregate, average level, but for each and every potential traveler. To collect this data, airlines invented frequent traveler programs so they could create traveler profiles. Finally, the matching of supply (using the seat inventory from the reservation system) with demand (using the traveler profile data from a loyalty system) is automated by means of a yield or revenue management system. Robert Crandall, former Chairman and CEO of American Airlines, gave yield management its name, calling it the single most important technical development in transportation management since deregulation (see Smith et al. 1992). Following the example set by airlines, our agent system concept features Matching as YM, Profiler as FFP, and Calculator as ARS.

Agent system

The agent system or “agent” is a market making system which matches supply with demand. The agent metaphor has become popular in mainstream computing and business, largely due to its suitability for the study of distributed systems. We follow Holland, an artificial intelligence scholar and genetic algorithms pioneer, in our conceptualization of a system who borrowed the term "agents" from economics “to refer to active elements without invoking specific contexts” (1995, 6-7). The field of economics that Holland was referring to is agency theory, which explains how to best organize the relationship between one party – the principal – who determines the work, and another party – the agent – who undertakes the work (Ross 1973; Grossman and Hart 1983; and for a survey, see Sappington 1991). The concept of an agent system for matchmaking in a customer channel system borrows from the conceptualization of “portals” decades earlier (for example, Schlueter Langdon & Bau 2007a, b; also, Hsin-Lu et al. 2003), which in turn evolved from “online networks” and proprietary online services, such as America Online (Schlueter Langdon & Shaw 2002, 1997).
Agent element. Figure 1 depicts the system boundaries of the agent and its core elements of matching, profiler, calculator, data exchange/ customer data platform, data factory, and data space. It is derived and adapted from (a) lessons learned from yield management in the U.S. airline business (Schlueter Langdon 2021, specifically, “Figure 3: Industrializing the Optimization of Yield per Unit of Consumption”, p. 33), (b) market-making multi-agent systems (such as Schlueter Langdon & Sikora 2006, based on Sikora & Shaw 1998), and (c) early online retailing systems (see Schlueter Langdon & Shaw 2000). In this example, the agent recommends products, such as travel suggestions, to users. The system will match a user’s product needs, such as a travel request, in terms of starting point A, destination point B, travel date, and time of day (either as departure or arrival information), and preferences such as travel speed, cost, comfort, number of stops (direct versus transfers), and more (sustainability, sheltered options for rainy weather, etc.).

Interaction between agent elements. The matching of supply with demand is accomplished by means of a matching subsystem. It uses a matching engine that includes recommendations based on machine learning (for recommendation engines, see Smith & Linden 2017), for example, to match the seat inventory on routes (seats) with user requirements and preferences (profile; for “The power of profiles – monetizing value across relationships”, see Schlueter Langdon & Bau 2007a, p. 254; “Profile power: Next generation search and economies of scope” see Schlueter Langdon & Bau 2007b, p. 721). The seat inventory is generated by a route calculator. The user profile is then created by a profiler. The route calculator determines route options using a calculation engine and is connected to third party systems so it can retrieve data on routes, schedules, and pricing. The profiler calculates profiles using a profiling engine that retrieves data from the data layer, which is made up of additional elements including a data exchange (DX), data factory (DF), and data space.

“Big data” and additional elements. With a shift in results from revenue per car to revenue per trip, a very appropriate first step would be to learn from the airlines. However, three systems may no longer be sufficient due to big data. 40 years ago, airlines faced a drought of data. Today, there is a real glut of data
available that can add critical value, and this needs to be utilized. 40 years ago, there were no smartphones. Today, nearly all adults in developed countries use one, and the device itself has evolved into one gigantic data logger (December 2018). Smartphones have been a key enabler of a trend that has been dubbed “SoLoMo” by John Doerr, a partner at influential Silicon Valley venture capitalist company Kleiner Perkins in 2010 (Guynn 2013). It summarizes the expansion of digitalization into social, local, and mobile applications, which has fueled the growth of data on consumers and the commercialization of this data for advertising and new service offerings. Companies like Facebook, Google, and Uber are key examples of this trend. Today, this data includes: (1) data on consumers (traditional demographics, government statistics), (2) data on products and services (from vendors), (3) user-product interaction data (behavioral data), and a broad range of (4) context data (see lower section of Figure 1). The latter ranges from capturing a consumer’s daily schedule and their friends and family to environmental settings such as weather and traffic conditions. All this data allows for better customization and personalization of offers by evolving from artificial and fictional “personas” with their inherent bias to true profiles cut from the real-life behavioral data of actual and potential customers (Crosby & Langdon 2014).

**Figure 2: Connecting agent into an IDS-based data space (adapted from Otto et al. 2019)**

**Data exchange.** This is where a data exchange system could add value (see DX in Figure 1). Think of it as a marketplace or “supermarket” with data products for data scientists. Today, according to meta-research, more than 80 percent of the time budget for a data analytics project is spent on data wrangling – not with algorithms (“Data is broken,” link). Companies have gone from databases to data warehouses and now to data lakes – and they seem to be drowning in them. The question is – how can all this data be consolidated, organized, and made available to data scientists? One possible solution is an internal data exchange. Instead of searching for data across departmental silos and country operations, a data scientist could “shop” for internal data in a central location.

**Data factory.** As more and more data will be generated within a company using social media, 5G, or IoT, another system will be required. Let’s call it a data factory (see DF in Figure 1). A data factory is needed
to economize on the refinement of raw data into these aforementioned data products (Schlueter Langdon & Sikora 2020). Despite the hype surrounding data analytics and artificial intelligence (AI), raw data is still confused with refined data. Machine learning and AI methods require refined data products (Crosby & Schlueter Langdon 2019). This is obvious for data scientists but few in management seem to be aware of it. There is a food analogy which can help illustrate the gap. Very few of us pick food from trees or slaughter animals; most visit a supermarket and pick food off the shelves. The food at the supermarket is processed, packaged, and labeled. Labels provide information on the product’s name, weight or volume, the vendor, ingredients, and nutritional value. For example, a “Nutrition Facts” label in the U.S. can easily exhibit 20 rows of data (U.S. FDA 2016). So, we could learn from the food industry when it comes to data: Raw data rights must be verified before any data can be ingested or harvested (rights, licensing, user consent). Then data ought to be properly labeled or tagged for it to be made discoverable through a catalog and search engines (classification). Furthermore, it must be given a rating to provide some indication of quality, because otherwise any subsequent analysis is pointless – “garbage in, garbage out” (GIGO, quality scoring). Finally, data governance mechanisms are required to ensure digital sovereignty for data owners. This problem can be tackled with a data space.

**Data space.** How can we connect it all while still ensuring data sovereignty? This is where the International Data Spaces (IDS) standard and technology comes in – it has been explicitly designed for this task (IDSA 2018). Figure 2 illustrates a basic setup between a data source (data provider) and data sink (data consumer) with the key element even referred to as an “IDS Connector” (Otto et al. 2019, p. 59). With this system, any data package or data product can be “packaged” with instructions and rules for use. Technically speaking, it is a dedicated software component which allows participants to exchange, share, and process data in such a manner that the data sovereignty of the data owner can be guaranteed. Depending on the type of configuration, the connector’s tamper-proof runtime can host a variety of system services, including secure bidirectional communication, the enforcement of content usage policies (e.g., expiration times and mandatory deletion of data), system monitoring, and the logging of content transactions for clearing purposes. The main advantage is obvious: an open standard lowers the cost for everybody and makes it easy to join in, which in turn increases the richness, quantity, and variety of data available to all data space participants. A first IDS-based data space has been implemented as the Mobility Data Space (Drees et al. 2021).

**Real-life demonstration: Travel planning app @ Reallab Hamburg**

On paper, there is no difference between theory and practice. Yet, the proof of the pudding is in the eating as they say. That’s why we put the agent system concept and data space technology to the test. We took our opportunity when participating in the RealLab Hamburg project with a dedicated subproject, “TP2: Data interaction and sovereignty”. RealLab Hamburg was initiated by the German National Platform Future for Mobility (NPM) to explore, probe, and test new mobility solutions in a controlled, real-life setting in the port city of Hamburg, Germany’s second largest urban area (“Digital Mobility Solutions”, link). RealLab Hamburg was funded by the German Federal Ministry of Transport and Digital Infrastructure (BMVI), with Hamburger Hochbahn (HHA), one of the largest transport companies in Germany, as project and consortium leader (“Startschuss für das Reallabor Digitale Mobilität Hamburg – mit Bundesminister Scheuer,” link). The project included eight subprojects with a total of 32 partners.

**Objective.** Deutsche Telekom’s Data Intelligence Hub unit (DIH, link) and the Urban Software Institute (link) teamed up for subproject TP2 to investigate how new distributed data infrastructure technology with sovereignty protection can facilitate improved mobility solutions. Specifically, the goal was to show how a federated data structure with sovereignty controls – a data space based on IDS – can create benefits for all stakeholders: faster, better travel options for citizens; new business opportunities for both
established public transport companies (e.g., convenience offers combining rail and on-demand shuttles) and new micromobility providers (e.g., connecting electric scooters to public transit).

Minimal viable demonstrator (MVD). To this end, a minimal viable demonstrator of a travel planning application was built. The app allows Hamburg residents to plan a door-to-door trip using different modes of transportation including public transport, micromobility, and shuttle services, all in one “super app”. This app was designed and implemented based on (a) the agent system concept (see Figure 1) and (b) the data space concept and technology (see Figure 2).

Data chains and data sovereignty. The integration of the agent concept with the data space is key to addressing a fundamental problem that has so far inhibited seamless intermodal travel: a lack of data sharing due to a lack of data sovereignty protection. Data sovereignty is the power to protect one’s rights to data (see for example, Federal Government of the Federal Republic of Germany 2021). In the past, this power was lost the moment data was shared. However, data sharing is imperative for intermodal travel to be seamless and for planning and booking to become a one-stop shopping experience. The intermodal travel chain requires a corresponding data chain across participating mobility providers and the end user. This data chain is not just a user interface issue either. The linking of different providers into a single app interface is only part of the story, and the easier one at that. For a “super app” to be attractive to end users, the app must also recognize a user’s existing relationships with these providers. For example, why would we bother booking a trip through a super app if it doesn’t incorporate our subscriptions and discounts with these providers. We wouldn’t want to miss out on discounts and pay more just to use the super app. But how could all of this be included in the super app? This could be easily done if all mobility providers involved were simply to share the data with the super app. As of yesterday, this has still not happened. The providers don’t trust each other, and they don’t trust the data sharing transaction itself. Anything can happen to data once it’s been shared. There is a lack of data sovereignty, the loss of power to protect one’s rights to data. Today, this is no longer the case. Data spaces, such as the Mobility Data Space built
on IDS technology (see Figure 2), are a real game changer since they can protect data sovereignty (Drees et al. 2021). It doesn’t matter if providers continue to mistrust each other – they can trust the data transaction because it protects data sovereignty, reducing the risk of data being misused once it’s been shared.

*Super app implementation.* For the travel planning app demo, the implementation of the agent concept (see Figure 1) features three different layers of user interfaces, engines, and data. Figure 3 depicts a high-level architecture view of this implementation, which reveals three different user interfaces (UI; map, widget, and digital twin), three different engines – one for each UI – and a data layer including a data space with technology from the Mobility Data Space. The calculator computes modal configurations and route options to be displayed on a map UI. The profiler aggregates end user data so it is accessible in each user’s digital travel twin UI. The matching process then computes personalized suggestions that match a customer’s preferences with route options visualized on a widget UI. This specific use case for the agent highlights the critical importance of the “matching” process for personalization. This feature is not a gimmick – far from it. It is a key feature for making intermodal travel work for residents like us. Why? Because our brains can only handle so much complexity (Iyengar & Lepper 2000). Integrating micro mobility and shuttle services with public transport dramatically increases the options for getting from point A to point B. For example, three different modes of transport (public transport, scooter, shuttle) for a three-segment trip, theoretically creates $3^3$ – or 27 – options, which is too much for us to make a quick decision. Consumer choice is good, but research demonstrates that too much of a good thing can be bad. When confronted with an overwhelming number of options, the outcome is decision paralysis (Schwartz 2005). Therefore, our app features personalization to limit recommendations to a short list of three options.

Figure 4: Three user interfaces of map, recommendation widget, and personal digital twin
Travel agent usage and first results
The intermodal travel planning app demo was installed on different screens during Intelligent Transport Systems World Congress 2021 in Hamburg (ITS WC 2021), including a fixed interactive monitor on the booth of RealLab Hamburg as well as handheld devices such as Apple iOS iPad. It was used by industry participants as well as residents during the open days of ITS WC 2021. The app demo also received much appreciation during special feedback loops with participating partners who were linked with the app using their data, including Continental (autonomous shuttle service), Sixt (electric vehicle rentals), Tier (electric scooters), and Ioki (shuttle service). These companies were linked on a test data basis and are compliant with the setting of a project funded by the Federal German government (e.g., Figure 5).

Customer journeys. From an end user perspective, interaction with the super app is very straightforward and intuitive. A user can either (a) initiate trip planning, an active scenario, or (b) be notified of an upcoming trip, a passive scenario. In the active scenario, a user opens the app and sees their digital travel twin (see “Anna’s travel twin” in Figure 4). The personal twin has all relevant personal travel data and is organized into multiple sections. The upper section allows them to enter their starting point A and destination B, as well as their start time or arrival time. The second section provides preference settings and sliders to specify the importance of travel speed, cost, and comfort. This section can be skipped to let the travel agent generate trip recommendations based on historic travel data and preferences. This data is visualized in section three, together with membership information or subscriptions with mobility companies. Here the user may find out that they have subscriptions that are rarely used or even find special offers from providers that would like to increase business with the user. If they confirm their selection, the twin will disappear, and they will automatically be switched to a map with a small window or widget hovering over it in the bottom right corner. The widget is directly linked with the digital twin; it’s a mini digital twin so to speak, or the twin’s representation on the map. Whereas the widget is part of the twin, the map could come from any provider linked with the super app. The widget highlights three travel options. The first one is immediately shown on the map (see map view on the left side of Figure 4).
Clicking on this option on the widget provides in-depth information for the route. Scrolling from option 1 to options 2 and 3 on the widget automatically switches the routes displayed on the map. The widget is like a remote control and scrolling through it lets us flip through routes on the map. Clicking the “Twin” button on the widget will make the widget disappear and pull up the twin to manually adjust the trip settings and preferences. This setup illustrates just how a user’s needs and preferences drive the market. Instead of starting on one provider’s map with the specific provider’s service territory and travel options, a user’s twin organizes the options from all different providers. It is easy to see that the twin can overcome the limitations of individual providers and cast a wider net to provide better results. All this in a single interface with the user’s travel twin, without any need to switch back and forth between different providers’ apps. This paradigm shift becomes even more visible in the user’s journey (b). They don’t have to plan anything. Instead, they simply get notified of an upcoming trip. Let’s say we’re reading the news on a tablet, and suddenly a widget icon pops up, just like any notification. When we click on it, it turns into the same widget we described earlier, and a map opens up automatically as well. We can close the widget and with it, the map, and return to reading the news. Or we can interact with the widget to investigate trip options as described before. This journey in particular highlights a new level of ease of use and speed in planning trips. Instead of a lot of back and forth, multiple logins and data entry, maintaining a profile and settings in many apps, the travel agent’s widget will just pop up with the best recommendations out there.

Figure 6: Intermodal travel speeds in Hamburg using travel agent’s calculator engine

Travel speed. The demo proves that a super app can make it easy to plan and book intermodal trips. But few of us use an app or user interface just for the sake of it, regardless how smooth and satisfying the customer journey may be. Ultimately, it’s about performance. To make the super app a success we need better results, and when it comes to traveling from A to B what matters is the speed of travel. So, is intermodal travel truly faster? One main difficulty in answering this question is the fact that there is no intermodal mobility on offer. How do you do the impossible to predict the probable? Fortunately, we can use a simulation, a proven research instrument across fields, from medicine to management (see for example Schlueter Langdon 2020a). Specifically, we use the calculator element of the travel agent to compute speeds for the different routes (see Figure 5). The calculator is connected with (a) route and scheduling data for public transport and (b) vehicle location data for electric scooters, shared bikes, shared electric vehicles, and shuttle services. The calculator’s engine is composed of (a) the open-source OpenTripPlanner for fixed public travel and (b) a proprietary engine for free-floating last mile solutions.
For travel speed calculations, travel scenarios or experiments have been carefully constructed to be representative of the city of Hamburg (for details, see Schlueter Langdon et al. 2021, Schlueter Langdon 2020b). Figure 6 presents the results of simulation runs for a total of 20 routes. These routes are exemplary but have been carefully selected to ensure representativeness. Also, for generalizability five routes for each cardinal direction have been chosen. Full intermodal travel (SPT1) is about 30 percent faster than individual car journeys.

Conclusion
Good news all round: Intermodal travel can deliver impressive speed gains, which makes urban modal shift a lot more likely. Who wouldn’t like to get from point A to point B faster? Furthermore, the super app demonstrator implementation verifies that data space technology can facilitate data sharing with sovereignty protection, which in turn enables data chains such as for mobility chains as in the RealLab Hamburg intermodal travel project. More broadly, a data space can enable new growth and revenue opportunities, lower cost, and help become compliant with new data regulation:

• Compliance and implementation of data governance: Data transactions and data chains in a data space with access controls, usage policies, and transparency rules (see Figure 2) can help companies comply with new regulation in Europe, such as the European Data Act or the German Act on Corporate Due Diligence Obligations in Supply Chains (Lieferkettensorgfaltpflichtengesetz, LkSG, link).

• Cost and scaling advantages due to a shift from a 1:1 connection (e.g., one OEM with one Tier 1 supplier) to 1-to-many relationships (e.g., one OEM with its entire supply chain), and possibly industry-wide solutions (like a club membership). Furthermore, data can be kept at the source and made available just-in-time to save the cost and confusion of duplication. In addition, the decentralized architecture and open, international standards help to avoid costly lock-ins.

• New growth and revenue: Novel data chains create new value orchestration and intermediary opportunities as illustrated by “travel agent” (see Figure 3 and Figure 4). Anna’s digital travel twin is also a symptom of the arrival of the metaverse (see for example, McKinsey 2022). While the outlines of the metaverse remain fluid it is widely seen as the next iteration of the internet that intertwines our physical with digital lives. Our intermodal travel demonstrator illustrates how a data space happens to be a key infrastructure enabler.

Acknowledgement
We are grateful to Lenny Hofmann, Tim Ganther, and Luca Loeffler, all from Deutsche Telekom, for assistance with select travel time estimations using a novel intermodal travel app demo. This demo has been funded in the RealLab Hamburg project as subproject “TP2: Data interaction and sovereignty” by the Federal Ministry of Transport and Digital Infrastructure (BMVI) under grant agreement FKZ 01MM2016HH (“Verbundprojekt: RealLabHH – Reallabor Nationale Plattform Zukunft der Mobilität; Teilvorhaben: VRUDaten”).

References