A Framework for Capturing the Business Benefits of Railway Digitalization

Andrew Nash
Senior Researcher
St. Pölten University of Applied Sciences
Matthias Corvinus-Straße 15, A-3100 St. Pölten, Austria
+43 676 933 0483
Email: andy@andynash.com

ORCID: https://orcid.org/0000-0002-3774-7226

Felix Laube
Manager for Traffic Management Systems
Emch+Berger AG Bern
Schlösslistrasse 23, 3001 Bern, Switzerland
Email: felix.laube@emchberger.ch

Samuel Roos
Senior Railway Planner
Emch+Berger AG Bern
Schlösslistrasse 23, 3001 Bern, Switzerland
Email: xsaro@bane.dk

Word Count: 4,441 words + 0 tables (250 words per table) = 4,441 words

Submitted August 1, 2019

ABSTRACT

This paper outlines a framework for changing railway systems and processes to help railways capture the full business benefits of digitalization. Economic research shows that businesses need to make fundamental changes to their systems and processes if they are to take full advantage of new technology. The slow implementation of digitally based signaling systems such as ETCS and PTC highlights the need for fundamental change in the railway industry to more aggressively implement new technology – and obtain the full benefits of this technology. The proposed framework integrates an improved and up-to-date understanding of customer needs with a much more efficient and customer-oriented production process. It is designed to make use of today's powerful data collection, communications and analysis technologies rather than applying new technology to old processes. The proposed framework has been developed based on earlier research results and practical experience. The paper is intended to spur discussion.

Keywords: Railways, Digitalisation, Information Technology, Management, Innovation

1. INTRODUCTION

Digitalization provides railways with an unprecedented opportunity to improve service quality and efficiency. [1] However, economic research has shown that businesses must fundamentally change their operating systems and processes to take full advantage of major technological changes such as digitalization.

Unfortunately, fundamental change is difficult for well-established and highly regulated industries such as railways. Therefore, most railways are applying digital technology to incrementally changed systems and processes which reduces benefits and slows the adoption of new technology.

A good example is the slow deployment of digital signaling systems such as the European Traffic Control System (ETCS) and positive train control. [2] These systems could provide the data for a new approach to railway operations and planning designed to significantly increase service quality and efficiency, but, instead they are often viewed today as extra costs.

In short, without fundamental change it is difficult for railways to make a strong business case for large scale information technology such as digital-based signaling systems. Therefore, the first step in capturing the full business benefits of railway digitalization is sketching a framework for fundamentally changing railway systems and processes to take full advantage of information technology. This paper outlines a proposed framework for achieving this objective.

Section 2 describes the problem of improving productivity using new technology. Section 3 describes how fundamental changes to systems and processes based on feedback can be used improve railway production. Section 4 describes the proposed framework for improving railway processes and systems designed around rapid production of railway timetables. Section 5 describes how rapid production of timetables can be also used to improve railway investment and schedule planning.

Section 6 presents conclusions and outlines implementation results. The proposed framework has been developed based on research and practical experience outlined in Section 6. The paper is intended to be provocative and spur discussion in the railway industry.

2. PRODUCTIVITY AND COMPUTERIZATION

In the early years of computerization many companies invested in new equipment, training and software, but there were only limited productivity gains. As the economist Robert Solow stated, “you can see the computer age everywhere but in the productivity statistics.” [3] However, soon it became clear that productivity gains came from redesigning systems and processes to take full advantage of new technology rather than simply applying technology to existing processes. [4] [5]

Figure 1 illustrates four stages in a company’s lifecycle. Stage 1 is the beginning, a new business. Here the company’s systems and processes can be fully aligned with technology because they can be defined specifically for the technology.

As time passes the company moves to Stage 2. The company has supplemented and refined its systems and processes, and technology has changed, but because of resistance to change in the company (e.g., due to institutional culture), there is a misalignment the company’s systems and processes no longer fit so well with the current technology. In other words, there exist technologies that are not being exploited by the company to the extent that they could be to improve its business situation.

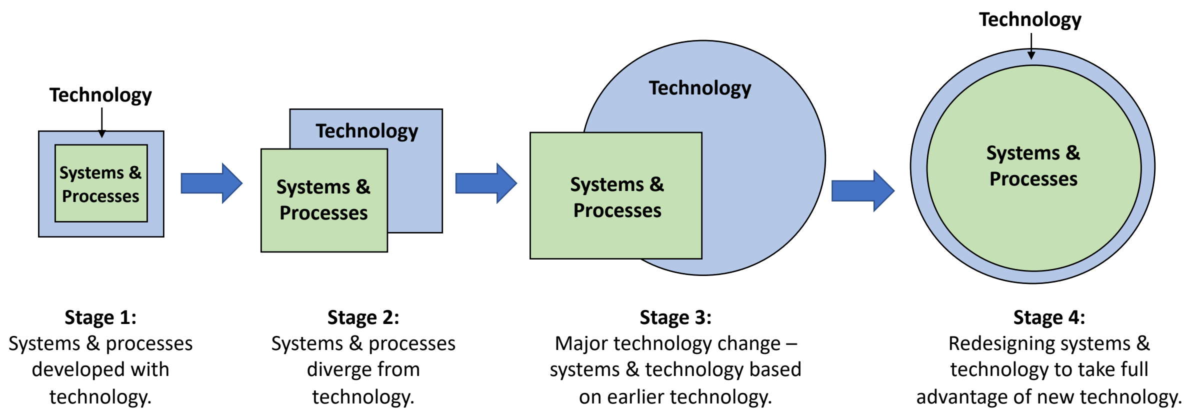

Figure 1: The relationship of business systems and processes to technology over time.

Stage 3 occurs after a significant change in technology; here the company's systems and processes become more significantly misaligned from the new technologies. Most companies in this situation adopt the new technologies but use them with their old systems and processes. For example, producing what they have always produced, only faster – rather than thinking about the opportunity new technology gives them for producing something that better meets customer needs than the old product. These companies do not go out of business immediately, but they leave the market open to new businesses created to take specific advantage of the new technology.

Some businesses are able to transition to Stage 4. These businesses reorient their systems and processes to take full advantage of new technology. They fundamentally re-invent themselves – changing how they do business, operating practices, and institutional culture. This is very difficult but necessary if companies are to thrive.

This paper argues that the railway industry is at Stage 3 of the technology process. There is a growing mismatch between railway business practices (operating practices, techniques, rules, methods, processes, culture etc.) and the possibilities of new technology, more specifically new information technology (digitalization).

A good example is the slow deployment of both the European Train Control System (ETCS) in Europe and positive train control (PTC) in the United States. [2] Both are digital based signaling and control system designed to ensure safe operation of railways. Both are quite well-known technologies, but railways have not been able to make a strong business case for their implementation.

This missing business case is analogous to the limited productivity improvement experienced following the early introduction of computers. In short, railways must embrace fundamental change by redesigning their systems and to take full advantage of the benefits of information technology (Stage 4). These benefits go well beyond safety (the main reason for ETCS and PTC) by using the digital data generated by new signaling technology to obtain business benefits of reduced costs and increased service quality.

3. FUNDAMENTAL SYSTEMS AND PROCESS CHANGE IN RAILWAYS

Railways provide transport service. Services are consumed at the same time as they are produced, which means there are no opportunities for correcting problems after production. Problems must be corrected in as close to real time as possible.

Correcting problems in real time is difficult, and for railways it is extremely difficult since they operate complex services over extensive and complicated infrastructure networks.

Closed loop systems (**Figure 2**) are used to correct problems in real time. In a closed-loop system the user's actions are monitored, compared to the desired state, and, if necessary, new actions are taken to achieve the desired state. Closed loop systems have feedback loops, enabling them to self-adjust and quickly stabilize in the desired state. This highly controlled process makes closed loop systems ideally suited for digitalization.

Applying closed loop systems to railways is a topic of growing interest. [6] [7] [8] There are two types of closed loop railway processes: task management and rescheduling, also referred to as the inner and outer control loops respectively.

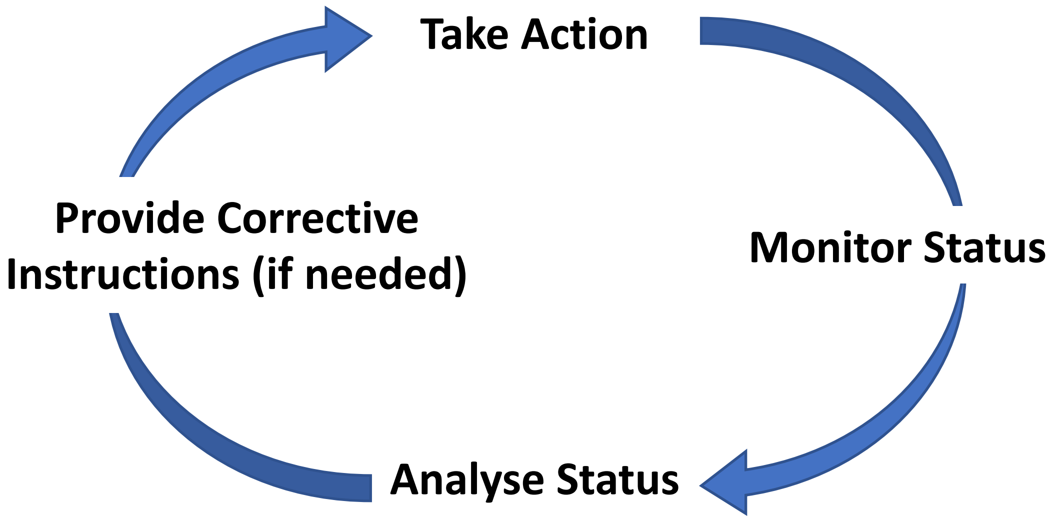

Figure 2: Generic closed-loop control process.

The railway task management and rescheduling control loops are illustrated in **Figure 3**.

The inner control loop, task management, consists of providing railway staff with information on what they are supposed to do and feedback on how they are performing so that they can take any necessary corrective actions. A good example is a driver following speed instructions from a driver guidance system.

The outer control loop, rescheduling, compares actual service data and resource status to the plan (timetable), and if actual conditions differ from planned by a set threshold, begins the process of rescheduling. Rescheduling consists of taking actions designed to return service to the plan. Rescheduling is done by dispatchers and today relies heavily on dispatcher experience. The dotted lines from the timetable box show that timetable data is used to both identify divergences from the plan and provide an objective for rescheduling.

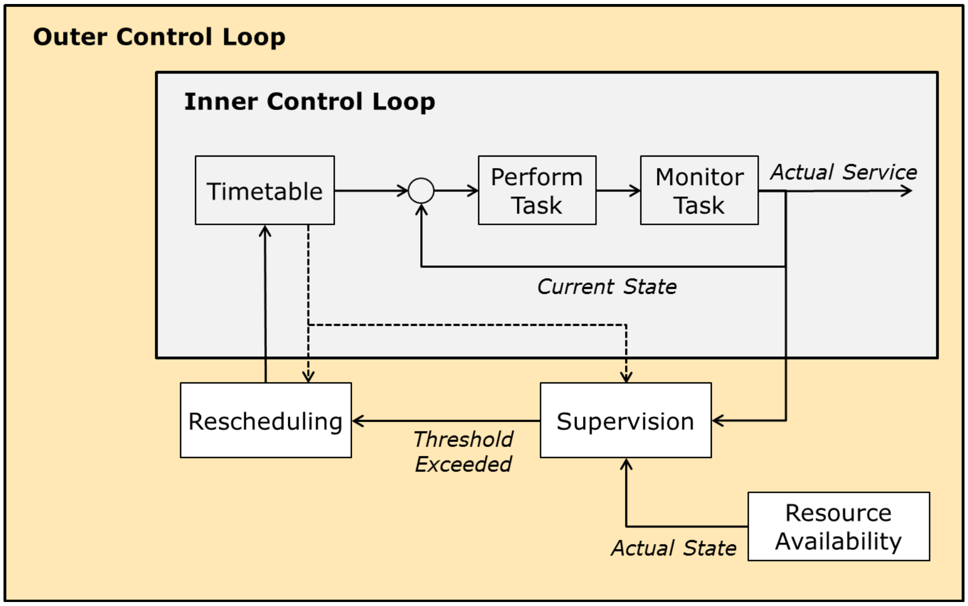

Figure 3: Railway task management and rescheduling control loops

Today's technology enables railways to fully implement these control loops throughout their operations, specifically:

- Inexpensive and accurate sensors monitor status in real-time;
- Wired and wireless technology transmits data for analysis and control; and,
- Powerful processors and data management are used to analyze data.

Applying sensors and communications technology is straightforward in the sense that they do not require fundamental changes to organization systems or processes. While they do involve extensive deployment of new technology, this technology replaces existing systems and processes. For example, today drivers use a printed timetable and knowledge of the route to determine their speed, and dispatchers use their experience, the timetable and available status information to reschedule trains. The new technology provides more and more accurate data.

The large amount of new more accurate data coming from sensors (including new digital signaling systems) could be analyzed by more powerful processors to improve railway operations by providing dispatchers and drivers with better instructions for achieving the planned timetable (incremental change).

However, having this data available also provides an opportunity for doing something different, something that would have been impossible without today's information technology, by inventing new systems and processes to take full advantage of these technologies (e.g., Stage 4).

4. VERY RAPID CREATION OF RAILWAY TIMETABLES

The system-changing innovation that will enable railways to capture the full business benefits of digitalization is the ability to rapidly develop new timetables based on highly accurate data.

Ultimately information technology should enable railways to produce new timetables almost instantly. This would enable them to revise timetables whenever there was a disruption or new customer need, thus providing precise feedback for implementation in the task management and rescheduling control loops.

It is impossible to create timetables instantly, but rapid advances in operations research and processing power will make it possible to create timetables much more quickly than today, for example in less than one-minute. The benefits of creating timetables faster could be obtained in stages by developing appropriate systems and processes for using the new timetables in the time-frame within which they are developed (e.g., if a new timetable could be developed in one-minute, how should the railway be operated during this one-minute period to best use the new timetable when it is ready).

Achieving the business benefits possible from very rapid timetable creation means developing new systems and processes to best take advantage of this speed, but also that help support actually achieving the higher speed. The rest of this section outlines how this could be done.

4.1 Scheduling Algorithm

Timetables are created using scheduling algorithms. The study of operations research and development of much more powerful information technology has led to significant improvement in scheduling algorithms. This is the increased analysis power described above.

There are two main components in a scheduling algorithm: an objective function and the operational steps (production processes to be completed). The operational steps must be carried out to achieve the objective function. Both the objective function and operational steps are determined by the organization's production methods, techniques, processes and technology.

As outlined above, when technology is first developed the organization can design their systems and processes together with the technology (hardware); however, as institutions age, it becomes difficult to adjust these systems and processes to take full advantage of new technology.

The most important part of this framework for capturing the business benefits of digital technology in the railway industry therefore, is developing improved objective functions and production processes.

4.2 Improved Objective Function

The current objective function for railway production is punctuality, defined as trains reaching their destinations (or control points) according to the planned timetable.

Punctuality is a very good objective function in the sense that it is specific and measurable. On the other hand, it doesn't measure exactly what customers want. Customers want to reach their destination on time, they don't care whether the train reaches its destination on time.

Of course, there is a huge overlap in trains arriving on time and customers arriving on time. Furthermore, it is very difficult to measure whether all customers arrive at their destinations on time (in terms of data collection and analysis).

The discrepancy between the current objective function (train punctuality) and optimal objective function (customer punctuality) is a perfect example of how technology availability defines institutional objectives: it has been possible to measure train punctuality for a long time, and it was a good substitute for customer punctuality.

However, today, train punctuality is a less good measure of customer punctuality because current best practice in railway timetabling depends on transferring between trains. In a fixed interval timetable ensuring connections can be more important than the overall punctuality of trains.

Interestingly, the importance of fixed interval timetables also reflects technological change. Originally railways were designed to provide point-to-point travel, they were used to travel or ship goods directly to a big city. Over time society's travel needs changed to a many-to-many network (caused partly by development of motor vehicles). Railways struggled until Switzerland and The Netherlands found that regular fixed interval timetables could be used to efficiently serve this many-to-many network demand. Again, implementing fixed interval timetables required changing institutional practices that had grown-up with railways, and it was a struggle (and remains so for many railways).

In addition to train punctuality being a less good measure of customer punctuality, today it is also possible to see a future when a scheduled transportation provider (e.g., a railway) could actually measure actual customer punctuality (e.g., did the customer arrive at their business meeting on-time?). This level of detailed data is not available yet, nor is the processing power available to use it in fast timetable creation, but it's a possible future.

On the other hand, it is possible to use customer behavior to eliminate the need for detailed customer punctuality data. More specifically, research shows that customers build-in a small amount of buffer time in their travel planning. This is best illustrated by arrivals at bus or tram stops: when frequency is low people tend to arrive near the vehicle departure time, but when frequency is high (every 10-minutes or less) people tend to arrive randomly. [9]

Assuming that customers include these buffer times in their travel planning, then it is likely that they are also not overly concerned if the train is a few minutes late. More important is that they are able to make their connections. (Note that the approach outlined here should also be extended to connecting scheduled transport such as buses.)

Putting these ideas together, taking full advantage of railway digitalization means changing the objective function from its current focus on train punctuality towards customer punctuality. This consists of two points:

1) Include customer needs beyond train punctuality in the objective function (e.g., ability to make connections);
2) Refine train punctuality criterion to better reflect customer behavior (i.e., don't require that trains arrive exactly on-time).

Regarding the first point, additional customer needs might include quality of service (e.g., rolling stock type). Furthermore, in addition to customer needs, the new objective function could include operating requirements (e.g., rolling stock disposition and crew assignments).

Regarding the second point, in addition to better reflecting customer behavior, refining the punctuality criterion also provides a measure of flexibility that can be used by the scheduler algorithm to develop a feasible timetable. This flexibility speeds-up the process of creating timetables.

4.3 Improved Production Processes

The second element in the scheduling algorithm is the production process. In essence, the scheduling algorithm attempts to complete a set of production processes to achieve the objective function. The production processes have input and output requirements.

Railway production processes are defined at various levels of detail depending on the analysis objective. For example, in an operational simulation an interlocking might be modelled precisely, while in long-term planning specific infrastructure such as interlockings might be ignored in favor of fixed travel times based on assumptions and standard practices.

Production processes are considered at varying levels of detail because, until now, it has been very difficult, technologically, to do otherwise. The amount of data needed to precisely analyze railway operations is huge; without today's data collection, transmission, storage and analysis systems it was impossible to prepare comprehensive detailed analyses of railway production.

Today's technology opens the possibility for drilling-down into the details of railway operating processes and using this highly detailed representation in the timetable development process. This is because these processes are "only data" and data today is easy to collect, transmit, store and analyze. Existing railway timetable planning practices are built for an age where using data was difficult; existing application of digital technology in railways is incremental: it applies the digital technology to the old practices.

In the case of production processes, taking advantage of technology means developing highly detailed descriptions of these processes. These descriptions include the tasks carried out by people that are required to produce railway service (e.g., the driver begins closing the doors) as well as the tasks required by other resources (e.g., infrastructure: setting points). Today many tasks are lumped together in multiple-task processes such as "departure process".

At first glance the advantages of such a highly detailed description of production processes might not be clear, especially given the effort required to create it. In fact, defining production processes in detail significantly improves production control and efficiency. More specifically:

1) By precisely defining the tasks required in a production process it is possible to monitor these tasks rather than overall processes (e.g., did the driver begin closing the doors on

time, rather than: did the train start moving on time). Monitoring specific tasks gives advanced warning of a disruption, which provides more options for solving the disruption satisfactorily.

2) The process of precisely defining production tasks gives railways the opportunity to closely examine and rationalize their production processes focusing on how to improve customer service and increase efficiency. Examining existing processes is particularly important for established organizations such as railways because these processes often resemble onions: a core of functions designed to serve customers, surrounded with layers of administrative, regulatory and institutional functions added over the years for one reason or another. Taking a fresh look at processes enables railways to strip away unnecessary layers and redesign processes to better meet business objectives.

Regarding the first point, precise task definition also will help improve railway operations by providing specific directions to staff. Today, especially in the case of a disruption, staff must think about (figure-out) what they need to do next rather than being able to react to the current situation. In the future staff would have precise tasks defined by time, and these tasks could be updated in case of disruptions.

Regarding the second point, process rationalization would not only give railways the opportunity to simplify processes by eliminating unnecessary tasks, but also allow railways to reduce the number of individual processes they use by creating common processes that could be combined to accomplish the objective of more specialized processes. Reexamining existing processes will also enable developers to radically simplify system architecture and automate many functions – thus increasing the speed and efficiency of timetable creation.

In summary, developing new processes based on a fundamental re-examination of customer needs and using these new processes to drive technology differs significantly from the conventional practice of simply applying new technology to existing processes.

4.4 Summary of the Proposed Framework

In summary, the proposed framework for more fully capturing the benefits of railway digitalization consists of developing a scheduling algorithm that can be used to create new timetables very quickly and that is based on a renewed understanding of customer needs and an improved description of production processes.

Clearly railways could benefit from the development of a very fast scheduling algorithm without changing the current objective of maximizing train punctuality or more detailed production processes. In fact, this is the incremental path being pursued by many railways today.

However, as shown in other industries, to fully capture the business benefits of digitalization railways must integrate a renewed understanding of customer needs and production processes into this new scheduling algorithm.

5. VERY RAPID TIMETABLE CREATION FOR RAILWAY PLANNING

The ability to create detailed timetables very rapidly and easily also means that these detailed timetables could be used in the railway planning process. Using detailed timetables for short- and long-term railway planning will improve the quality of planning by using a better objective function and a much more detailed description of production tasks.

Railway planning and production can be described in terms of three main activities:

- Long-term planning consists of developing timetables to serve future customer needs without resource constraints. This allows planners to strategically evaluate service ideas and identify the resources (infrastructure and rolling stock) needed to provide that service.
- Short-term planning consists of developing timetables with resource constraints. Short term plans can be developed for various time horizons in the future, the only requirement

is that there must be sufficient time before timetable implementation to obtain the needed resources (e.g., to build new infrastructure or obtain appropriate rolling stock).

- Production planning consists of developing timetables in real time (e.g., to address a delay, an infrastructure problem, or to provide a service such as an extra train after an event).

Today, it is difficult to create detailed railway timetables. Therefore, railways create simplified timetables for use in the planning process and only create detailed timetables late in the short-term planning process (e.g., for the annual timetable). As described earlier, it's currently impossible to create timetables fast enough for production planning.

The new scheduling algorithm proposed in this framework creates detailed timetables quickly enough so they can be used for long- and short-term planning as well as production.

Using the same timetable planning system (scheduling algorithm, objective function, production tasks) throughout the full process of long-term planning, short-term planning and production could also create a new control loop directly linking planning and production. This third control loop will help link the entrepreneurial process used in service planning directly to production, enabling production staff to provide better customer service and planners to create more implementable timetables.

Using the same timetable planning system throughout the process would also reduce technology costs since all planning and production would be supported by the same software and data structure. In addition to improving the quality and consistency of the planning-production process, this approach can reduce the number of tools and databases being used by different departments, thereby improving information technology efficiency and reducing overall costs.

Figure 4 illustrates the new planning and production loop. As shown it uses the same scheduler algorithm throughout the process as well as the new customer needs oriented objective function (Service Intentions – SI) and detailed production tasks (Production Plan – PP).

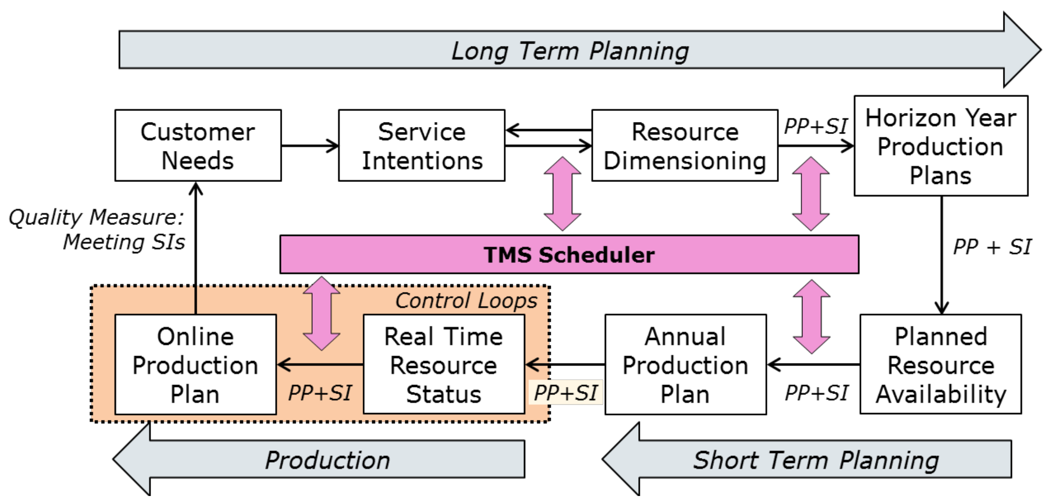

Figure 4: Railway planning and production loop

6. CONCLUSIONS AND IMPLEMENTATION EXPERIENCE

The need for fundamental change in systems and processes to take full advantage of new technology has been shown in economic research. The slow implementation of digitally based signaling systems such as ETCS and PTC highlights the need for fundamental change in railway production systems and processes. This paper outlines a framework for changing railway systems and processes designed to help capture the full business benefits of railway digitalization.

The proposed framework has been developed based on research and practical experience. It has not been fully implemented at any railway although its main building blocks have been tested successfully in research studies and an initial application of the framework is currently being developed at Banedanmark as part of its nationwide ETCS implementation. [10]

More specifically, the idea of defining production processes in detail and using this understanding in railway rescheduling was tested at the Swiss Federal Railways in their Puls90 program. [11] The specific tools developed for timetable planning at Banedanmark have been successfully used to prepare Banedanmark's Concept Timetable 2030, which supports significantly increased service and formed a basis for the ETCS design. [12]

In summary, the proposed framework is designed to help railways achieve the full benefits of digitalization by integrating an improved and up-to-date understanding of customer needs with a much more efficient and customer-oriented production process. The paper is intended to be provocative and spur discussion in the railway industry.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: A. Nash; analysis and interpretation of results: A. Nash; draft manuscript preparation: A. Nash. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

[1] Nash, Andrew, Ulrich Weidmann, and Marco Luethi; Can Information Technology help Rail Play a Greater Role in Preventing Climate Change? In Transportation Research Record No. 2139; Transportation Research Board, Washington D.C., 2009, pp 133-141.

[2[Briginshaw, David; Will digitalisation spur ETCSdeployment in Europe?; International Railway Journal; May 1, 2018; https://www.railjournal.com/opinion/will-digitalisation-spur-etcs-deployment-in-europe accessed 1 August 2019.

[3] Solow, Robert M.; We'd better watch out; New York Times Book Review; 12 July 1987.

[4] Brynjolfsson, E. and L. Hitt; Beyond Computerisation: Information Technology, Organisational Transformation and Business Performance; Journal of Economic Perspectives, Fall 2000, pp. 23-48.

[5] Harford, T.; Fifty Things That Made The Modern Economy; Little, Brown; 2017.

[6] Lüthi, M., 2009. Improving The Efficiency Of Heavily Used Railway Networks Through Integrated Real-Time Rescheduling. PhD Thesis, ETH Zurich.

[7] Corman, F., Quaglietta, E., 2015. Closing the loop in real-time railway control: framework design and impacts on operations. Transport Research Part C: Emerging Technology 54; pp 15–39.

[8] Lüthi, Marco, Giorgio Medeossi, Andrew Nash; Structure and Simulation Evaluation of an Integrated Real-Time Rescheduling System for Railway Networks; Journal of Networks and Spatial Economics, Springer Netherlands; Volume 9, Number 1, March, 2009; pp 103-121.

[9] Luethi, Marco, Ulrich Weidmann and Andrew Nash; Passenger Arrival Rates at Public Transport Stations; Presented at the Transportation Research Board Annual Meeting 2007; http://www.andynash.com/nash-publications/Luethi2007-pax-arrivals-TRB-paper.pdf accessed 8 July 2016.

[10] Laube, F. and Nash, A.; Advanced TMS will improve operating efficiency; Railway Gazette International; December 2016; pp. 35-37.

[11] Luethi, Marco, Andrew Nash, Ulrich Weidmann, Felix Laube, Raimond Wuest; Increasing Railway Capacity and Reliability through Integrated Real-Time Rescheduling; E-collection ETH Library; https://www.academia.edu/30018404 accessed 1 August 2019.

[12] Nash, Andrew, Samuel Roos and Bernd Hermann Schittenhelm; Using Banedanmark's TMS to Develop Concept Timetable 2030; US Transportation Research Board Annual Meeting 2017; http://docs.trb.org/prp/17-00359.pdf accessed 22 February 2018.