Artificial intelligence in industry, a lever for change and a factor of innovation for the RATP Group

Côme Berbain,
director of Innovation, RATP Group,
&
Yohan Amsterdamer,
head of the Artificial Intelligence Program, RATP Group

Abstract:
The transportation sector has been making innovations centered on its industrial aspects for decades now. After several successive waves of automation, it is now up to artificial intelligence to bring to light new potentials for improving the quality of services and the user experience, which will be more customized and inviting. These new capacities, which are opening for the RATP Group (the Paris Area Transit Authority), are vectors of change and innovation for building models of new trends and providing knowledge that should be made concrete and operational. The subtle match between expertise in business processes, mathematics, computer science and data is the key to success.

The RATP Group (Paris Area Transit Authority) designs, runs and sees to the upkeep of various means of transportation (by bus, tram, subway, suburban rail line, boat, cable) in response to mobility needs not only in the Île-de-France region (which includes Paris) but also in other cities in France and outside the country. Given the long adaptation of its industrial infrastructure over time and the ever stronger demand for transportation, the Group must be capable of using all technological levers (such as digital technology and artificial intelligence) to improve both operating conditions and the customer experience.[1]

A program of innovation at the service of customers and transportation

Given that France and Europe want to position themselves as major actors in artificial intelligence (AI), especially as applied to transportation and "mobility", the RATP Group is conducting a strategic program for making AI a pillar of its digital transformation.

The history of "machine intelligence" at the RATP dates back. An unbroken line can be traced from the automatic control of capstans for shunting operations through the first systems of assisted locomotion and automatic ticketing, to the sensors installed along rail lines and on the rolling stock, and up to systems for planning and optimizing the supply of transportation. Automatic subway lines and then their automation — feats achieved while maintaining service as usual — have proven the Group's capabilities. In addition, they provided an opportunity for reviewing the whole production chain, from design to current operations and now up to the user experience. For all this, the RATP has relied on operational research and a first form of AI, which we might call deterministic or

[1] This article has been translated from French by Noal Mellott (Omaha Beach, France).

"symbolic". Nowadays, the potential value of data, improved algorithms and increased computing power have opened new fields to exploration in an age of statistical or "probabilistic" AI. In industry, this new AI is leading to the simulation of new classes of phenomena, a more rapid, complex, local theorization of phenomena and an amelioration of the performance of existing deterministic models.

To address these new issues, an AI program was, in 2018, designed and launched around several axes defined in relation to an analysis of employees' and customers' needs in conformity with the Group's values of responsibility and inclusion. Four axes were assigned priority:

- The increased performance of the industrial production of the supply of transportation by: *(a)* a finer, real-time perception of traffic incidents and unexpected events during operations; and *(b)* a real-time simulation (on the scale of a multimodal network) of the demand for mobility and of passenger flows in order to optimize the allocation of the means to be put to use.
- A more proactive, efficient maintenance of machinery so as to improve the availability and operation of equipment by: *(a)* consolidating local, contextualized knowledge of the operation of equipment in the network (*e.g.*, the rolling stock, infrastructures, passenger information devices, areas open to the public, etc.); and *(b)* detecting correlations between causes and effects, signs of failure and maintenance operations.
- The customer experience and the proposal to customize trips by: *(a)* understanding customers' characteristics, expectations and needs during each phase of a trip (through passenger information); and *(b)* equipping the Group's customer services to improve satisfaction and efficiency.
- Improve the feeling of safety and security among passengers and employees and at our work sites in order to detect, describe and prevent risks by using data from the real-time installations that are part of the system of video-protection.

From expert knowledge to tools and services for passengers and employees

The knowledge and know-how of transportation experts have gradually been computerized. This has made it possible to automate processes for supervising and regulating subway lines, both complicated tasks (*e.g.*, schedules for conductors) as well as simple but repetitive chores (*e.g.*, ticket sales). This computerization helps us keep up with the increasing demand for mobility and monitor customer and employee satisfaction.

This expert knowledge has been developed, acquired, stored and updated thanks to a long, empirical process of trial and error that has laid the grounds for a mastery of theory-making. At stake is the ability to "translate" this knowledge so that a machine can execute it. The RATP has applied this principle of theorization many a time to various phenomena.

How to describe and build a model of random, statistical phenomena, such as traffic disturbances and incidents, or passengers' constantly evolving needs and uses? These are topics for which AI proposes adapted solutions that are based on big data for building models of new phenomena. The key issues are how to produce, acquire, curate, consolidate and understand these data so that interpretations and operational analyses can be made. The expert's role and knowledge from the field of business processes take, once again, center stage. Domain experts intervene throughout the life cycle of data, from design and engineering to the operation and maintenance of transportation infrastructures and information systems. These same experts oversee the genesis, production and interpretation of these data, which, in turn, are used to validate the model's performance and results. Experts "teach" AI to machines. To do this, they receive help from data scientists, a new expert domain in mathematics and computer science.

This process has been used to formalize (and capitalize on) a source of knowledge that used to be restricted to individuals and stemmed from their personal experiences at the workplace. This formalization has facilitated the conveyance of knowledge and improved collective performance. One of the first projects in our AI program for regulating traffic a suburban line (RER B) offers a good example of this. To cope with a disturbance, the actions performed by switchmen and traffic supervisors are based not only on a set of regulations about rail safety but also on their personal experiences. The analysis of a traffic disturbance made by these domain experts and the decisions they choose as being the "best possible" ensue from their knowledge, which is not very formalized and is conveyed through practice. By making these experts come to grips with AI, models of good practices have been formalized, theorized, built and generalized.

Deterministic systems opened possibilities for describing, deciding, planning and optimizing, but this probabilistic AI, by using big data, makes it possible to analyze and predict. The combination of the two forms of AI is what will enable us to build tomorrow's software for assisting decision-making, overseeing current operations and maintenance, and improving the quality of both services and the passenger experience. This whole chain of functions is necessary for an industrial approach. This calls for a tighter coexistence between these two forms of AI, which should not be kept separate. The one relies on the engineer-expert whereas the other introduces the domain expert as a key player. This pair of experts oversees the training of computers for AI.

From the intelligence of data toward the knowledge of experts

The computerization and automation of transportation and mobility on a massive scale have led to developing ever more complex systems that produce ever more data. For AI, these data represent a potential for tackling new classes of phenomena that experts have been unable to make intelligible by using their current stock of know-how.

Some areas of expert knowledge are still incomplete, insufficiently explored or underused owing to a lack of robustness or of theories; or they are not yet accessible because the necessary data do not exist, are not available or are of poor quality. The ability to predict how late a train will be due to traffic is one example. Another example is to detect the warning signs of outages, to understand how systems of the new generation can fail. Furthermore, by correlating malfunctions with current operations and maintenance, a new body of knowledge is going to be accumulated faster. To cite a last example from the pandemic: the ability to build models of (and predict) passenger flows in close contact with real-time operations in the field as a function of the supply of transportation.

AI thus leads to questions about the formalization and intelligibility of this new body of knowledge. Is it always important to make a theory or interpretation, and trigger the best actions? Once AI deeply analyzes the classes of incidents related to machines and equipment, domain experts need not understand the model. They have to be able to apprehend the major explanatory variables for triggering the right maintenance operations at the right time. Eventually they will also have to act on maintenance engineering. The aforementioned project on RER B was an example of this. The formation of a traffic disturbance model relies on explicitly formulated variables, such as deviations in rush hour traffic, a typology of the current supply of transportation, and so forth.

From this capacity for a deep, local analysis, knowledge can be created in domains characterized by new forms of complexity that have not yet been explored. This is the way to move ahead in coping with certain phenomena. In some cases, probabilistic AI can accelerate model-building. However this knowledge is generated in a new form that implies a major shift in our learning and understanding. Data engineering is a domain of expertise linked to this change in the form and conception of this new body of knowledge. The difficulty stems from the duality of the pair: engineer-expert and domain expert. In practical terms, a firm must redesign careers and accelerate the recourse to “agile project management” so as to debunkerize prescriptors, developers and users.

Given these increasingly complex systems and the constant need to take into account the contexts of uses, these new “deposits” of knowledge must be tapped. They are the source that will inform and teach us so that we can attain new levels of performance and offer new services — this is the feedback that big data and AI will provide to experts!

Artificial intelligence, a learning process that fosters innovation

Artificial intelligence introduces a new way of thinking about learning processes in a firm. It proposes a virtuous circle between the learning process that experts undertake on machines and the feedback from big data to experts. Some tasks and operations can thus be automated to gain efficiency and improve transportation services, in a context of growing demand and customer requirements. This opens toward new functions and a new body of knowledge that will be of use to experts, operatives and customers. In fact, this dynamic, ongoing learning process yields benefits for current operations and for the customer experience in the transportation sector. This makes us reconsider the process of designing information systems and the chain of operations so as to switch from a static, deterministic model to a learning, probabilistic model.

There is, however, a major limitation for an industrial rollout: the access to data and the quality, qualification and even existence of data and of records about the data. This entails a change in data engineering and information systems. This change will be gradual since data might not always exist or be accessible and since expert knowledge does not always expand. For example, it is not realistic to imagine that we can have enough 2D images of each category of problems with spark plugs to feed into an automatic detection system! The balance between deterministic and probabilistic AI will always have to be found so as to come up with the right solution for the problem at hand. Even if, for instance, video data and artificial vision can be put to use, it will not always be simple or advisable to use them.

A learning process favors and stimulates innovation! The more expert and machine knowledge there is, the more innovative concepts, ideas and needs will emerge.