

BASIC DATA INFRASTRUCTURE (BDI)

A set of agreements for event-driven coordination in logistics

The Basic Data Infrastructure (BDI) is a set of agreements that allows participating parties to jointly (federatively) develop a specific IT network. This IT network makes it possible for data to be shared and/or retrieved from the source after authorisation. The BDI is a realisation of concepts that are being developed as part of the European FEDeRATED¹ project.

The BDI mainly focuses on event-driven coordination in the physical world where many parties together need to achieve a result: like in logistics. In logistics everything is about contractual agreements, coordination between many (sub)contractors and service providers, and about demonstrating compliance to government bodies such as Customs. The actions required to maintain the physical flow of goods are the trigger for all kinds of 'events' in the network. Knowing that relevant events are taking place and then being able to check the source is very important to all parties involved².

A set of agreements like the BDI roughly consists of agreements about:

- roles, responsibilities and rights of participants;
- mutual interaction processes;
- functional specifications of IT components;
- protocols and standards.

Participants in the BDI set of agreements subscribe to the agreements. They then invest in the necessary organisation and IT components for their role and for the part they are adding themselves. The basic assumption is that every participant chooses its own 'tech stack' to develop these IT components: this freedom is required for them to make their own choices and to be able to try out new technology.

The combination of these investments ensures that an IT network is created with special capabilities. These capabilities cannot be developed by a single party; the network is what provides access to these capabilities: the result of the federative collaboration.

These capabilities are used to devise and develop the desired services that make a difference for a group of participants. The BDI provides the blueprint for developing the network that makes these services possible. The services themselves are not part of the BDI.

The development of the BDI is aimed at providing more and more increasingly rich capabilities. In many cases 'more capabilities' will be accompanied by 'implementing more agreements and more components'. In exceptional cases, a certain capability will require a certain tech stack to be used, as there are only a few alternatives.

Not everyone will initially, or even in the long term, require all the capabilities. The extent to which a group of participants does so is their own choice. This choice depends on what a group of participants want to achieve together, which services are required and how the wishes develop over time. However, there is a minimum set of principles and agreements that always need to be applied by every participant. It forms the basis for interoperability and a migration pathway.

The minimum is currently defined as:

Principles:

1. Retrieving data from the source, machine-to-machine, as much as possible.
2. The Data Owner determines who can access the data (*Data Sovereignty*), even if it is hosted by a delegated party (*Data Holder*).

Agreements:

3. Applying the agreed semantic model, either explicitly or implicitly, with specific variations and expansions being possible depending on the use case.
4. Applying an API and/or SPARQL end point as an access point for the source.
5. Applying Identity & Authentication in the prescribed manner.
6. Applying Authorisation with permissions in the prescribed manner if the Data Owner requires Authorisation.

This minimum is necessary, but as such not enough to develop functioning data sharing networks: additional components are required (such as registers for finding APIs) to develop these. For the time being, they can still be designed and added separately for each project. As the developments progress, these components will also become available.

In practice, the development of the concept will continue, allowing more and more capabilities to be added to the network. And in the practical situation, lots of lessons will be learned about how the system can be used to maximise its value.

Digital Infrastructure for Logistics

The Digital Infrastructure for Logistics programme, which will be implemented during the period until 2027 with assistance from the National Growth Fund, accelerates the development and application of the BDI. The DIL Living Labs are mainly aimed at logistics via the main ports in the Netherlands. Projects other than DIL will also be implementing and testing the BDI agreements. This is being applauded: the more parties gain practical experience, the faster developments and adoption will be.

¹ FEDeRATED: www.federatedplatforms.eu

² The same considerations play a role in construction and in Engineering & Contracting. It is very well conceivable that these sectors will also benefit from the BDI.

BASIC DATA INFRASTRUCTURE (BDI)

Frequently asked questions (1)

① What is the BDI?

The Basic Data Infrastructure (BDI) is a set of agreements aimed at the logistics sector that allows participating parties to jointly develop an advanced IT network. This IT network makes it possible for confidential data to be shared and/or retrieved from the source after authorisation, machine-to-machine. The parties that participate can be national and international companies and government bodies.

② Is the BDI a government platform?

No. The BDI is a federative set of agreements. It is not a technical solution that runs in a data centre or in the cloud. Every party that participates invests in its own IT. The government will also develop its own IT to be able to participate in the BDI, but this is not a platform or *'man in the middle'*. However, the government is the powerful driver behind this development.

③ What is a federative system?

The word federative means that the collaborating parties are still independent, take their own decisions and make their own investments. They voluntarily choose to subscribe to the agreements and to invest in the associated IT. This provides a network that offers more options than what a party on its own can achieve.

④ So what does the word 'Infrastructure' stand for?

The IT network that is created when the agreements are applied by all participants provides basic facilities (*'capabilities'*) rather than ready-made services. Participants are free to devise and create their own services that use the unique capabilities. It is the infrastructure (the network) that offers the capabilities.

⑤ A set of agreements sounds like a very vague term. Do any functioning examples already exist?

The largest and most successful example is used by us every day: the Internet. The Internet is a *'network of networks'*. The set of agreements ensures that separate computer networks form a large virtual network together: every computer in every network can send data to and receive data from every other computer, regardless of which network that computer is part of. Every party can keep its own network, take its own investment decisions, make its own technological choices. A minimum set of agreements appears to be enough to allow more than 25,000 networks to function as a joint virtual network: the Internet. Nobody owns *'the Internet'*. The agreements about the Internet are jointly managed, modified and expanded. This joint nature creates something that no other party can develop on its own.

⑥ You talk about increasing 'capabilities' by implementing more agreements. Can you give any examples?

The Internet is once again a good example. Once all connected computers were able to exchange data packages, the need for something akin to a telephone book or *'Yellow Pages'* quickly arose. Until then, computers could only be recognised by their IP address: a series of 12 digits.

The Domain Name System was then devised: an organisation can register a *'domain name'* (tno.nl for example), which is linked to them from then onwards. And the freely accessible *'DNS records'* can be used to find the IP addresses of the public computers in a domain.

The next innovation was to use a simple prefix to identify people associated with a domain (t.jansen@tno.nl for example). Publishing the IP addresses of the mail servers for each domain in the DNS records allows everyone to email each other. Each of these extras requires additional agreements to be made and implemented in addition to the minimum. It also requires everyone to invest in hardware and software to be able to participate. However, nobody is required to implement these additional agreements: even today you can still use the Internet by only applying the minimum agreements. But it is usually convenient to implement the additional agreements.

⑦ What does the BDI add in terms of capabilities compared to what is already there?

We already have the Internet with EDI, email, chat, websites/portals and APIs on top of it: there are platforms that function as hubs. Large numbers of messages are already being sent. The reality, however, is that many human hands and eyes are still required to interpret and translate everything between systems.

The first step the BDI makes is to reverse and automate this:

- rather than receiving messages, parties are given the option to retrieve data from the source, machine-to-machine, if the source (Data Owner) allows them to:
 - data at the source;
 - data sovereignty;
- rather than that people log in to other computers and retrieve the data, computers can retrieve and process the data:
 - common semantic model;
 - automatic identification, authentication and authorisation.

BASIC DATA INFRASTRUCTURE (BDI)

Frequently asked questions (2)

The second step is to focus on the term 'events', things that happen in the physical world.

The combination is a major step forwards, in any case for logistics.

The common semantic model is one of the key building blocks of the BDI: a common terminology framework. A semantic model (ontology) is a formal description of terms and relationships between these terms in a way that allows computers to interpret the terms. It is the key foundation for software developers that need to create links and translations.

A standardised semantic model makes it easier to develop standardised APIs to request data from the source. A semantic model also allows the application of so-called Semantic Web technology (SPARQL queryable endpoint with a triple store for example). And last but not least, a semantic model is a precondition for many AI applications.

Standard ways of requesting the data from the source are essential when many parties are active in a network.

The details of all the capabilities will be announced over time once they have been developed.

8 What happens to my data if I am the source being accessed?

Professional and (commercially) sensitive data obviously needs to be protected: you may only access it if the Data Owner has given permission or if there is a legal obligation for the Data Owner to allow a government body to access it. The BDI system involves computers that query each other without human intervention.

A digital trust network for computers is essential for achieving this kind of protection: automatic identification, authentication and authorisation.

The BDI builds on the iSHARE system that has already been developed for a digital trust network.



Identification means that the organisation, the employees who are authorised to act on behalf of that organisation and their computers are known to a so-called Identity Provider, which checks this. The European EIDAS system regulates this in Europe. In the Netherlands, eHerkenning is a well-known method that is based on EIDAS.

Authentication is the checking step used when a party's computer requests information from the Data Owner/Holder: the Data Owner/Holder's computer can verify whether the alleged Identity is correct.

Authorisation is the verification whether this party has permission to access certain data (need to know, when to know).

Government bodies and companies can voluntarily make mutual arrangements about accessing more data than is legally required: this can be very useful. For private parties, mutual access is a purely commercial agreement.

9 What is the benefit of accessing data at the source?

In many cases, the principle of 'accessing data at the source when this is required' yields surprising new possibilities for collaboration.

In distribution, one infamous phenomenon is the 'Monday peak': companies and private individuals often order things at the end of the week or during the weekend, so the orders are all sent out on Monday. By allowing the scheduling packages of transporters to access the digital diaries of the recipients to check their delivery preferences they can spread the peak more and also comply with these wishes more.

Another example is following a container that arrives in a port, will be unloaded and needs to be picked up. It is very important for everyone who needs to act according to these steps to retrieve all kinds of status information from various parties involved, such as:

- Has the shipping company's bill been paid?
- Do Customs want to check the container?
- Is the container on the quay?

Quickly checking the source provides the most recent information, which is a lot more reassuring than waiting for a message.

10 You talk about 'events' like they are something special. What is the story behind this?

The semantic model introduces essential new terms, namely a number of 'events': from planned to executed events. Events are the digital representation of the result of an action in the physical world: for example 'expected arrival time issued', 'container has been unloaded'. So an event is the result of actions, not the action itself.

BASIC DATA INFRASTRUCTURE (BDI)

Frequently asked questions (3)

Take the event that a full container has been loaded onto a container ship in China. All parties involved in the Netherlands, from the port authority, the container terminal, forwarders/processors, Customs, hinterland transporters to the shipper that ordered the goods, want to track the status of the ship and this container from that moment onwards. And they create events as a result of this knowledge: a declaration, reservation of capacity, etc.



The concept of the BDI is that parties involved can subscribe to main events and the 'daughter' events: they receive new updates, a new status and new events. If required (and permitted) a party can request more data from the source.

The concept of events and subscribing to them is broadly applicable: take the schedule for a building site for example. It is highly effective for all the suppliers and subcontractors if new things or changes to the schedule are automatically identified.

Most of the development work will involve events for the time being: experience from international projects such as FEDeRATED shows that there still are many questions about the best way to create events, publish them, subscribe to them, make sure that nothing is changed, save and search events. The optimum approach still needs to be discovered.

⑪ If everyone has standardised interfaces, will it not become too tempting for the government to request more and more data?

The principle of Data Sovereignty is becoming increasingly important in European politics, both for private individuals and companies. Data is your personal property. On the other hand, there is pressure to improve supervision and to prevent fraud and crime. The choice of what kinds of data do and do not require government access is a debate that will probably continue indefinitely.

The principle of the BDI is that the Data Owner itself has control over *who* is given access to *what*. Both are fundamental starting points of the BDI.

⑫ What is the difference between the BDI and DIL (Digital Infrastructure for Logistics)?

The DIL programme is supported by the National Growth Fund. It focuses on the accelerated development and application of the BDI and on increasing the 'digital readiness' (in particular) of SME parties active in the logistics chain. The development and application of the BDI takes place in Living Labs within DIL.

These are aimed both at the development of new capabilities for the BDI and at their application in valuable use cases. The DIL Living Labs mainly focus on the import and export streams via the main ports: the existing PCSs there play a key role in this. We are now already seeing projects other than DIL recognising the benefits of the BDI and developing and testing their own use cases based on the minimum agreements. This is being applauded: the more parties gain practical experience, the faster developments and adoption will be.

⑬ What is the planning for DIL?

DIL will be launched in late 2022 and will last around 5 years.

⑭ Can my organisation already start working with the BDI?

EA number of basic 'capabilities' of the BDI (or minimum agreements) can already be used now: among other things, they are already being applied in the development of innovative services within the Logistics Top Sector, the Clean and Zero-Emission Construction programme, the Goods Transport Corridors and by RVO to retrieve energy data from companies with energy saving programmes.

The Logistics Top Sector is also setting up an Adoption Support Team: a secondary helpdesk for parties that are preparing applications of the BDI in projects. The arrival of the AST ensures that the DIL programme can go full steam ahead, while other parties also receive help in starting up.

If your organisation also wants to get started, this is applauded: the more practical experience is gained, the faster developments and adoption will be.



BASIC DATA INFRASTRUCTURE (BDI)

Frequently asked questions (4)

15 Can my organisation participate in the DIL Living Labs?

If your organisation can and wants to make a contribution, please contact Sjoerd Boot, who is the person responsible for the development of the DIL programme at the Ministry of Infrastructure and Water Management.

Email: sjoerd.boot@minienw.nl or telephone 06 - 25 20 28 33.

FEDeRATED is mainly aimed at logistics: coordination of goods flows in the physical world. The BDI uses the reference architecture of FEDeRATED as a starting point and provides a practical application in the Netherlands, which from the very start has maximum interoperability with federative infrastructures being developed in other Member States.

16 How does the BDI relate to European initiatives such as FEDeRATED, IDSA and GAIA-X?

FFEDeRATED, IDSA and GAIA-X are roughly based on the same principles: data at the source, data sovereignty, semantic modelling, digital trust. The differences are found in the specific focus given to certain capabilities: this is based on the commercial/social goal of the system.

17 When I purchase new software, should I already be considering the BDI?

DIL and the BDI will be focusing greatly on informing IT developers and IT service providers about the underlying technology. They will then decide for themselves which migration pathway would be desirable and which preparations they can make.



Ministerie van Infrastructuur
en Waterstaat

More information?

Please contact

Sjoerd Boot, responsible for the DIL programme at the Ministry of Infrastructure and Water Management.

E-mail

sjoerd.boot@minienw.nl

Telephone

06 - 25 20 28 33.

29 juni 2022