TREN/06/FP6TR/S07.55976/019851

# PROMIT

## Promoting Innovative Intermodal Freight Transport

Co-ordination Action
Priority 1.6.2 Sustainable Surface Transport

# D 4.2 Performance indicators and benchmarks for self assessment

Due Date of deliverable: December 2007
Actual submission date : December 2007

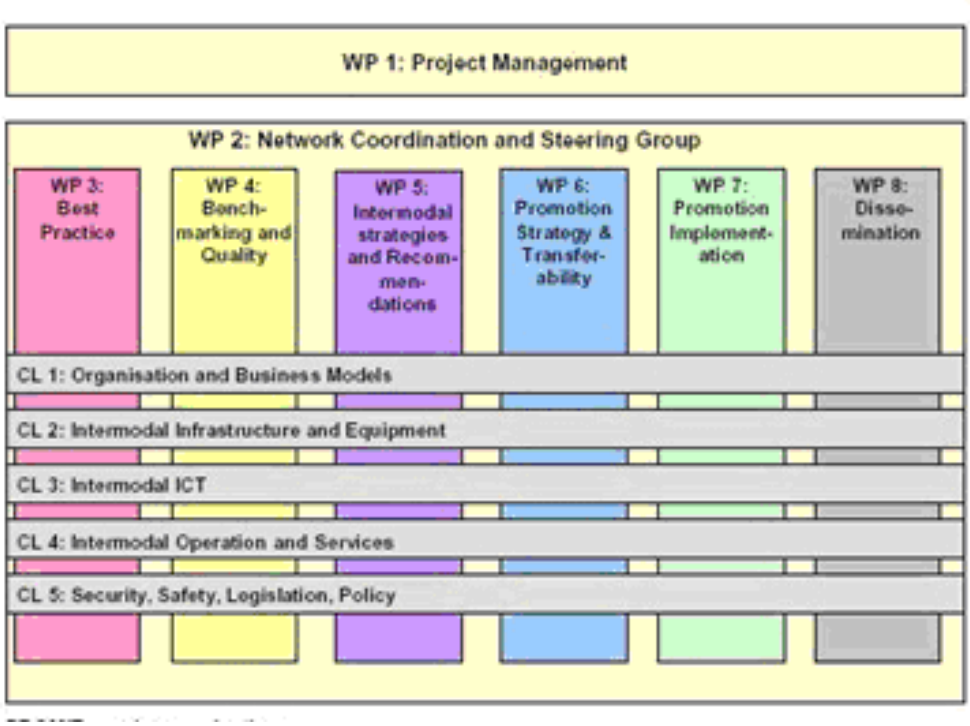
Start date of the project: March 2006 Duration: 36 months

Main Authors: E.Guis, M.Sc., M.Sc., I. Davydenko PDEng, G.Zomer, M.Sc.

| Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006) | | |
|---|---|---|
| **Dissemination level** | | |
| **PU** | Public | |
| **PP** | Restricted to other programme participants (including the Commission Services) | |
| **RE** | Restricted to a group specified by the consortium (including the Commission Services) | |
| **CO** | Confidential, only for members of the consortium (including the Commission Services) | **X** |

# 1 Introduction

This report presents work that has been done in the framework of PROMIT Work Package 4 on development of a benchmarking self assessment tool. This WP crosses all clusters in the PROMIT matrix structure. Therefore, the benchmarking self assessment tool touches aspects represented in all of the clusters. However, the main emphasis of the self assessment tool lies in continuation of the work on definition of benchmarks and producing a practical tool that can be of interest to commercial parties: users and providers of (intermodal) transport.




The tool is a logical extension of the benchmarks that have been defined in D 4.1 European Benchmarks in Intermodal transport. In that report a set of the most important KPI's has been defined as well as logical division of the intermodal transportation into transportation abstraction levels. The tool presented in the report is a mean of gathering benchmarking information.

Finding real life benchmarking information is very hard. Private companies see no interest in providing them, unless they can get something useful in return for their effort and information. The intention of the tool is to gather benchmarking information, at the same time the most important feature of the tool is that it is useful for shippers and transporters. The tool is intended to make transportation more transparent: how can a shipper know that its transportation provider offers it a good service? How can it be made aware of other transportation means such as intermodal transport? How can a shipper be assisted with a transportation mode choice and facilitated during change, overcoming the habit barrier? The tool is intended to answer all these questions.

Thus, the target user groups of the tool are shippers and transportation service providers, who transport or provide transportation on container or FTL scale. Shippers of smaller quantities, for

instance on parcel or pallet level are out of the scope. The tool will bring more transparency into important properties of transportation. If it is widely used, it will be a strong promotional tool for intermodal transportation; a good insight generator for policy makers at the EC and national levels; shift good flows from road to alternative transportation mode, thus reducing congestion problem; bring some environmental impacts as a consequence of better environmental friendliness of alternative transportation modes. The impact is not limited to the abovementioned reasons; however, transparency creation and facilitation of private shippers are the most important issues.

In order to make implementation of the tool possible, TNO has elaborated the general design of the tool. PTV has carried out actual implementation of it. This report describes the tool by explaining its design, the ideas behind it and implementation issues. The report is structured in the following way.

Chapter 2 on the tool design describes general considerations and motivates the tool, benchmarks presentation method and indicators that need to be measured. In other words, the chapter presents a top level design of the tool. Chapter 3 presents a vision on the tool and describes how it should work in practice. During design development phase TNO developed a tool prototype, which was intended to be an illustration of ideas and as a way to show that approach taken is feasible; the prototype was also a useful presentation means to show design issues in practice. The prototype included tentative implementation of all benchmarks that had been described in the 1st year report on benchmarking (D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report). However, in this report we clearly specify how those benchmarks must be measured in order to minimize misunderstanding and ambiguity as much as possible.

Chapter 4 on technical design of the tool presents information on data collection feedback generation and research functionalities. The chapter couples these three functionalities of the tool and presents information flow, technical aspects, information processing and workflow. We discuss the execution of processes within the tool on the technical level. Chapter 5 presents the implementation protocol. Implementation of the tool has been divided into 10 activities that were to be carried out (some of them independent, but the majority of the activities are sequential). The involvement of different partners within PROMIT has been discussed and defined. Finally, Chapter 6 describes actual implementation of the tool. It reports on actual implementation of the user registration, data entry and feedback generation processes and functionalities. In this chapter we also discuss differences between proposed design and actual implementation.

# 2 PROMIT benchmarking tool design

## 2.1 General considerations

The main aim of the PROMIT benchmarking tool is to provide companies with means for assessment of their logistics performance. The secondary aim of the tool is to gather benchmarking information that can be used by policy makers and for standards setting purposes. Given the secondary role of the tool, the participating companies must be assured that their benchmarking information is used properly and is not disclosed, at least not on the company aggregation level.

It is tempting to measure various indicators. However, the large number of potentially useful indicators represents a problem: the tool must be simple enough to be easily understandable by the participating companies; on the other hand, the tool must collect meaningful information. Therefore, the number of indicators must be very limited and at the same time giving a consistent overview. In this tool proposal we stick to the indicators on Level 1 and 2 from the 1st year PROMIT report.

Another issue to which attention should be paid is comparability of the measurements. For instance, if one company transports goods from A to B and the other one transports goods from C to D (or even from A to D), then a direct comparison of costs and transit times will not be meaningful, because these shipments are carried out over different distances and, possibly, within different countries. Therefore, results of the measurements must be somehow standardized, possibly by measuring cost per kilometre and transit time per kilometre (i.e. shipment speed). All these considerations are included into the design of the tool; however some extra attention should be paid to them at the development and deployment stages.

## 2.2 Presentation method

When a company enters its data into the system, it gets feedback on its performance in the form of relative position in respect to the aggregate performance of other companies. The relative position of the company is important here; a visualized relative position gives a good feeling in respect to the performance.

Moreover, the logistics performance has to be measured using several indicators. An excellent performance in respect to one indicator is often coupled with underperformance in respect to other indicator(s). This principle reflects a common wisdom in logistics: a higher service level leads to higher costs. The companies are (supposedly) interested in a balanced picture, namely to see their position on the cost/service trade-off. The following figure uses an example to represent the principle.


*Figure 1. Example of presentation method for benchmarking transport operations*

As it can be seen at the example on the picture above, the company has a good performance on transit (lead) time, excellent performance in respect to the damages; however the costs are high. This leads to a good overall performance.

## 2.3 Important issues over the presentation method

There are two major issues concerning the presentation method described above. *The first issue is related to the quality of the input*. If the system operates fully automated, an (intentionally) wrong input may severely distort averages. For example, if the average transportation cost by road is 1 Euro per container / km, and a company enters 200 Euro per container / km, it would increase average cost by the factor of 2 if the database contains 200 measurements. If the number of measurements is lower, the distortion will be even higher. As a solution, we propose to accept data only from authorized users and to have a process of "approval" of the data that is entered into the database. In case a human operator suspects a mistake, the data is not activated in the database; the company in question might be contacted for clarifications.

*The second issue is the scaling problem*. The scaling problem manifests in two ways. Suppose that we have performance indicators of two companies: company A with an annual goods flow of 1 000 ton km and company B with an annual goods flow of 1 million ton km. Should these two companies be given the same weight in the database (e.g. when the averages are calculated)? The second manifestation of this issue is calculation of overall performance of a company. The question here is whether all indicators should be given the same weight while calculating the integral performance indicator? The scaling issue can be dealt with at the later stages of development. See the chapter on implementation of the tool for the chosen solutions.

## 2.4 Indicators

In the PROMIT Synthesis Report Year 1 Work Package 4: Benchmarking, we have derived a number of Key Performance Indicators that are important in transportation logistics. Further, we have divided them over 3 transportation abstraction levels.

| **Key Performance Indicators** | | |
|---|---|---|
| **Level 1** | **Level 2** | **Level 3** |
| Price | Price | Price |
| Lead time (Transit Time) | Lead Time (Execution Time) | Transit Time (Execution Time) |
| Frequency of service | Frequency of Service | Capacity |
| Shipment compatibility | Shipment Compatibility | Capacity utilization |
| Damages | Damages | Reliability |
| Theft | Theft | Damages |
| | | Theft |
| | | Operation-Specific |
| **Additional indicators** | | |
| Flexibility | Flexibility | Schedule Convenience |
| Information Flow | Information Flow | IT (ICT) Systems |
| Invoicing Accuracy | Invoicing Accuracy | Energy Consumption |
| Relationship with shippers | Relationship with Provider | Number of Accidents |
| Market | Market | Operation-Specific |
| Energy Consumption | Energy Consumption | Frequency of Service |

Now the question is how to measure them precisely, to specify them in a way such that they are well understandable to the companies on the one hand and such that they bring meaningful results in the form of feedback and consistent database.
The implementation of the tool measures not all of the performance indicators described above. The reason for this is that for the first version of the tool a simplified approach has been takes: the tool had

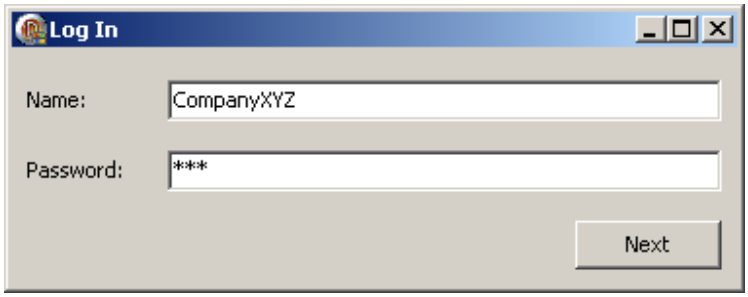
to be implemented fast. And even more important, the indicators chosen are the most basic: it increases attractiveness of the tool for the private companies and simplifies data entering process. More details on implementation can be found in the chapter on implementation of the tool.

# 3 Vision: How it should work in practice

When the design of the future benchmarking tool was being elaborated, TNO came up with a tool prototype, which showed how the benchmarking tool should work in practice. The prototype consisted of a simple windows application that included a number of screens to enter data (no database connections were necessary to implement). Below, we describe our design considerations coupled with definitions of benchmarks and input parameters. These parameters were planned at the design stage, however were only partially implemented in the current version of the tool. For more information on the current version of the tool realization, see the chapter on implementation.

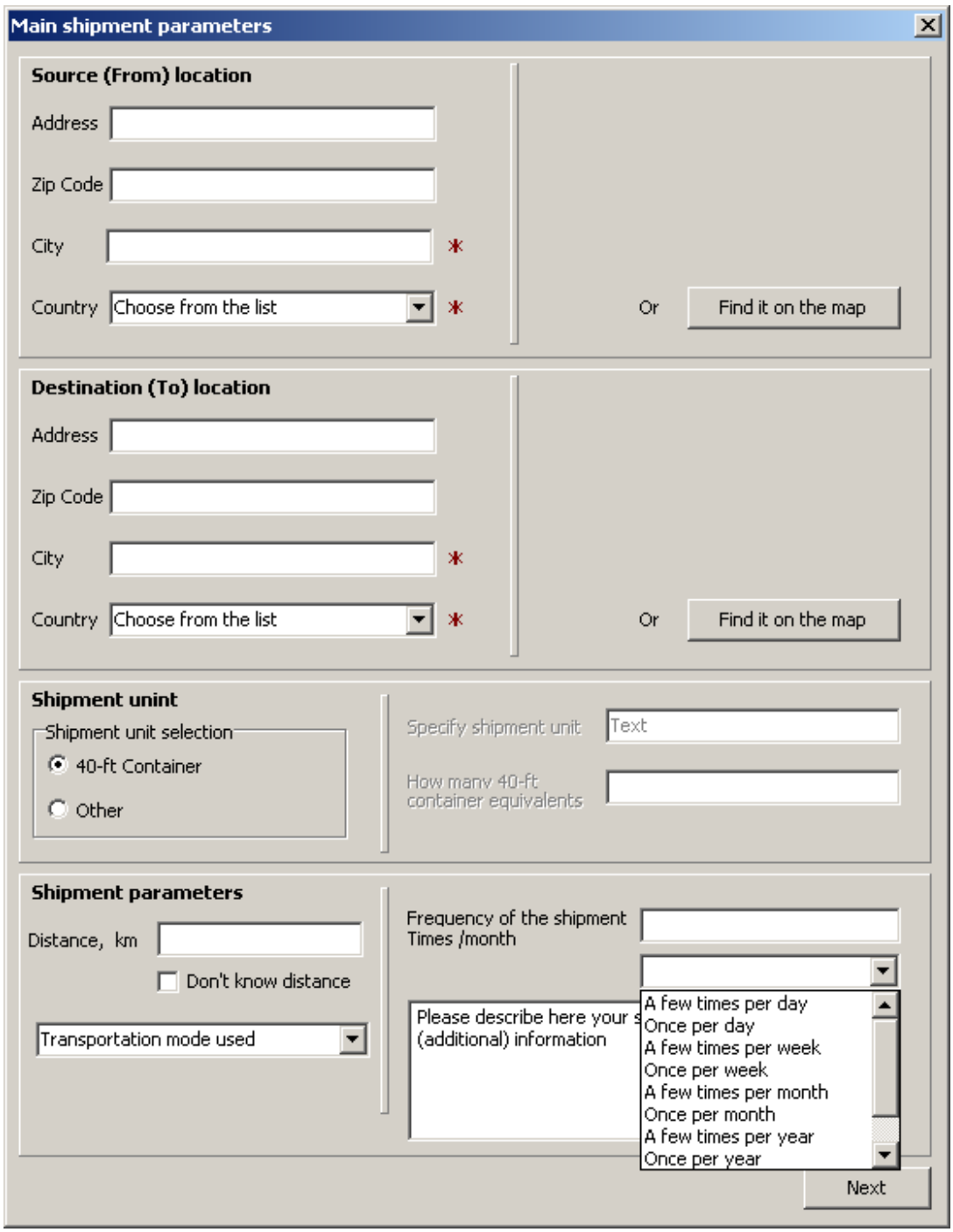
The data entrance process starts from the highest level of abstraction, namely the shipment level. First question that should be asked from participating companies is to specify the source and destination locations and what kind of shipments they conduct. Having entered general information, the companies will be asked questions on performance indicators. It should always be possible to specify a "Don't know" answer: absence of information is much better than wrong information. If the user specifies "Don't know", then his data will not be benchmarked against this specific parameter. It is also important to note that if a parameter is assessed on a scale from 1 to 10, then we assume that 1 is the lowest (worst) and 10 is highest (best) indicator. The first implementation approach would look as follows.

1. **Company identification**. Here, we just mean an identifier that is linked to a company or an entity profile (e.g. user name). The question of company identification should be discussed: how much information should a company provide at registration? Will it be allowed to stay anonymous? If a formal registration is required, then the company should first be asked to log in.




2. **Source point**. The more specific the source point is specified, the better. Minimum requirement is to specify city and country in the text form. Depending on implementation, there could be an implementation of selection on the map.
3. **Destination point**. Here apply the considerations on the Source point input.
4. **Shipment unit**. We would like to measure shipments in standardized units: 20 ft or 40 ft containers. The 20 ft container (-equivalent) or TEU is the industry standard, however, 40 ft container is more convenient measurement, because the majority of transportation is done using 40-ft containers. This is required information, i.e. without it benchmarking seems impossible.
5. **Distance.** The user specifies the distance that the shipment covers between the source point and the destination point, measured in kilometres. Given known the source and destination points, the user may choose 'don't know' option.

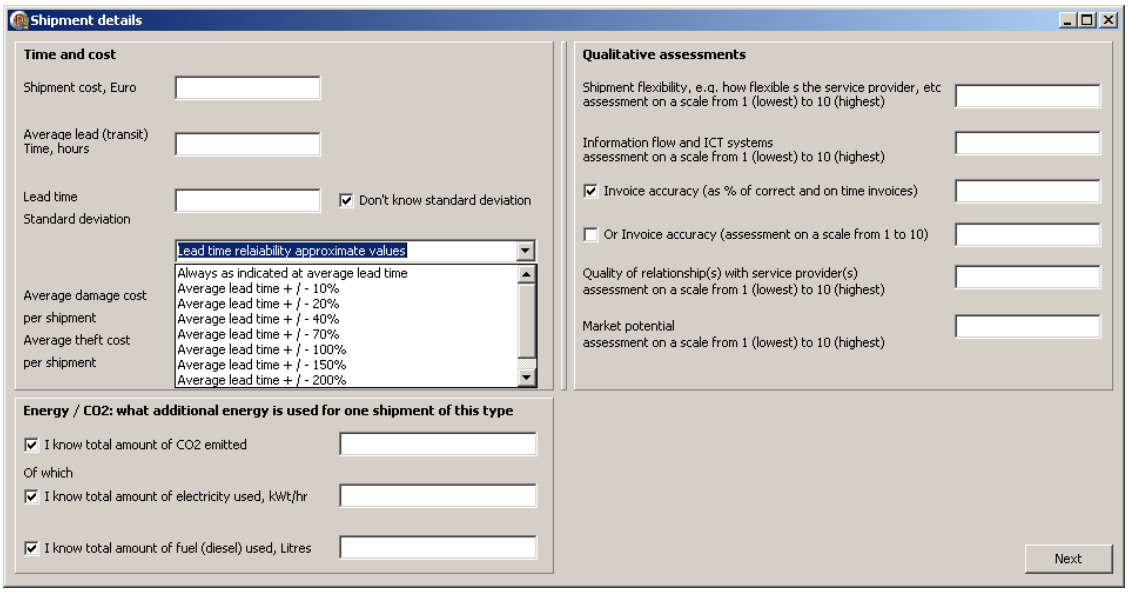
6. **Frequency of this shipment**. The user specifies how often the shipment is carried out, namely how many times per period, for certainty per month. If the shipments are not very regular, the user should be able to specify in which frequency brackets the shipments occur.
7. **General information on transportation mode**: single mode, intermodal (which modes involve), etc.



8. **Shipment cost**. Here the company specifies the cost of a single shipment of this type.
9. **Lead time**. Here the company specifies how long it takes to conduct the shipment. We are interested in the average lead time and its variation. However, it is difficult to get a well measured variation / variability of lead time. Transportation people are familiar with a less technical interpretation of reliability, which is difficult to quantify. Below, we state lead-time related parameters. The participating companies should have an option of a "Don't know" answer for each of these parameters, in case they do not have corresponding information. Thus, we propose to measure lead time in (one of) the following way:
    a. **Average lead (transit) time**. Here the company is asked to specify how long it takes to ship between the source and destination points. The measurement has to be conducted in the following way: what amount of time passes between a departure moment and an arrival moment. In other words (and as much general as possible), we define lead time as the time which is required for transportation, from the moment when the goods enter transportation until the moment the goods arrive to the destination. In a strict sense, average lead time is understood here as the algebraic mean of the lead time, measured over several shipments of this type. Hopefully, there should not be much misunderstanding among participating companies with the parameter of average lead time
    b. **Stated (promised) lead time**. This parameter specifies what lead time the transportation provider promises to the shipper. The difference with the average lead time is that average lead time is a measurement of actual performance, while the stated lead time reflects the agreement between the shipper and transporter.
    c. **Lead time variability**. The tool should provide a possibility to enter data on the following criteria. These are all measurements of the same parameter – lead time variability. We have introduced these three ways of its measurement because different companies may measure variability differently. Thus, we assume that a typical company would have information for at least one way of measurement: the companies should be able to choose the way how they measure the parameter and enter it once. There also must be a possibility to say a 'Don't know' to all three ways of measurement.
        i. **Statistical lead time variability**. This parameter specifies how much real lead time per shipment diverges from the average. In a strict way, we define lead time variability as the standard deviation on a set of real measurements. We expect that many companies may not have data on this parameter, or more probably they would not know how to calculate it.
        ii. **Perceived variation**. The question is formulated in terms of 'please estimate lead time variability as percentage of the average lead time: no variability, +/- 10%, +/-20%, +/-40%, etc. This alternative provides a way to measure variability as a coefficient of variation of lead time $c = \frac{\sigma}{\mu}$, where $\sigma$ is the standard deviation and $\mu$ is the average lead time, thus referring to percentages of average lead time.
        iii. **Reliability (stated lead time achievement)**. Reliability is measured as percent of cases when the stated lead time is satisfied. Many companies benchmark their service providers, such that assessment on such a criterion should be possible.
10. **Damages**. The company will be asked to specify average damages per shipment in terms of cost. Here we must make clear, that even though the one-off damage value may be high, but if it does not happen very often, then the cost of damages will not be very high. For instance, if a company experiences damage with value of EUR100.000 once in 10.000 shipments, then the average damage cost per shipment is EUR10.

11. **Theft**. Here the company specifies average theft value per shipment. The considerations described in the 'Damages' section, also apply for theft.
12. **Shipment flexibility**. The company specifies how flexible the service provider on the scale from 1 to 10 is. For more details, see report D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report
13. **Information flow and ICT system**. The company specifies how good the ICT system is and assesses information exchange on the scale from 1 to 10. For more details, see report D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report
14. I**nvoice accuracy**. The company assesses it on a scale from 1 to 10 or specifies percentage of correct and on time invoices (if such precise information is available). For more details, see report D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report
15. **Quality of relationship(s) with service provider(s)**. The company assesses it on a scale from 1 to 10. For more details, see report D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report
16. **Market potential**. The company assesses it on a scale from 1 to 10. For more details, see report D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report
17. **Energy used, measured as**
    a) Total amount of CO2 emitted, kg
    b) Total amount of electricity used, kWt * hr
    c) Total amount of fuel (diesel) used

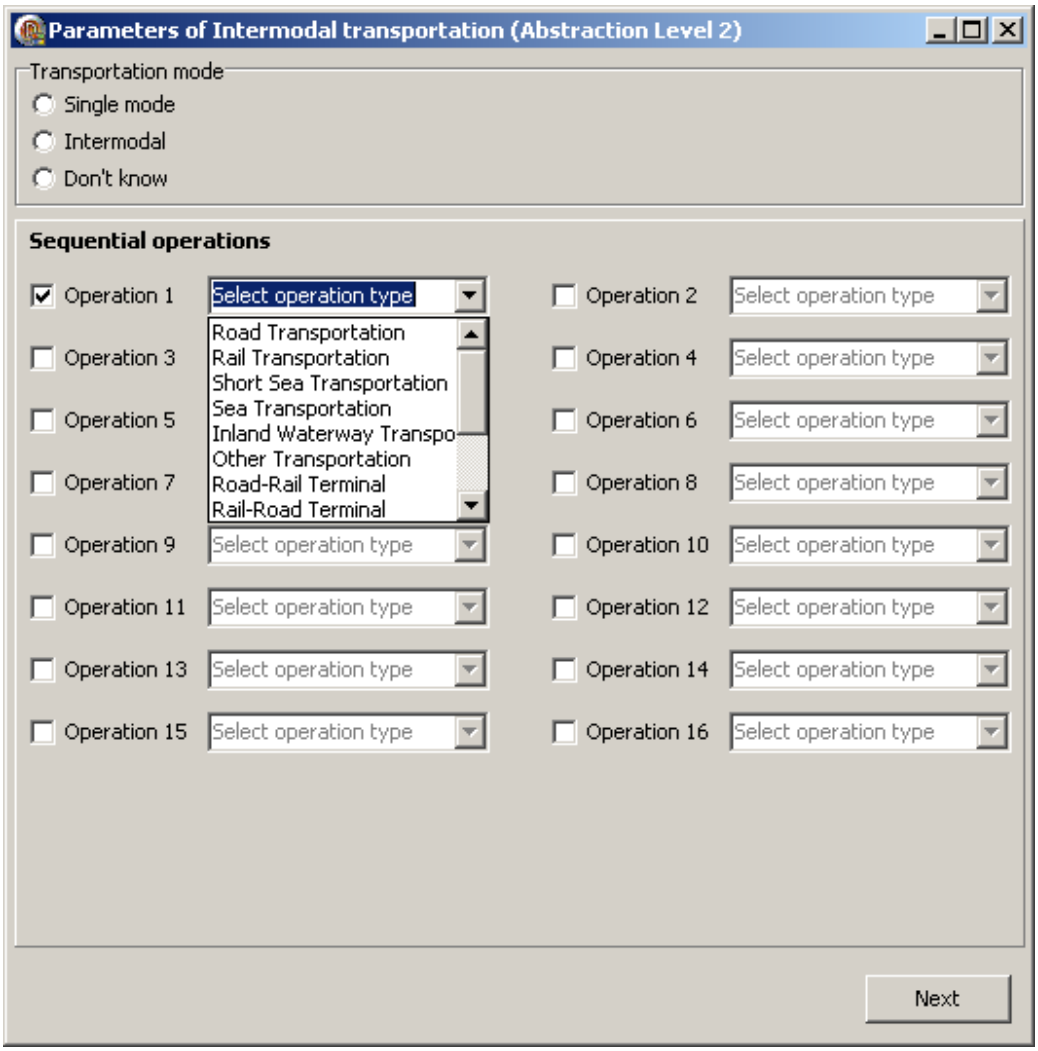
It is possible that some energy-related indicators are not applicable or there is no information on them. In this case, "Don't know" indication must be given




It is not reasonable to expect every company to have specific information on all of the parameters. Thus, we need to require an absolute minimum of information; other fields can contain a 'don't know' answer. It is in the company's best interest to provide as much information as possible (thus a good

incentive to make an effort gathering it), because more indicators can be judged, the judgment would become more correct and the overall performance of the company could be assessed more precisely.

Having entered the shipment parameters (abstraction level 1), the company then is asked to specify whether the transportation is intermodal and, if it is known, to specify a sequence of intermodal operations.




If the company specifies a sequence of intermodal operations, then it will be asked to fill in information on each operation at the second level. For each operation, the company will be asked to fill in forms similar to the forms it filled in for the first abstraction level; however, they will be more operation specific.

It seems that the data on the third abstraction level should be collected separately and in a different way. Here we connect all information to a specific shipment, and if the company knows details, then it will be possible to get information on the second abstraction level, operation specific. However, it is not reasonable to expect a shipping company to know details on operations of terminals or technical

characteristics of the vehicles used. Therefore, information on the third level should be collected by surveying service providers, for instance train and terminal operators.

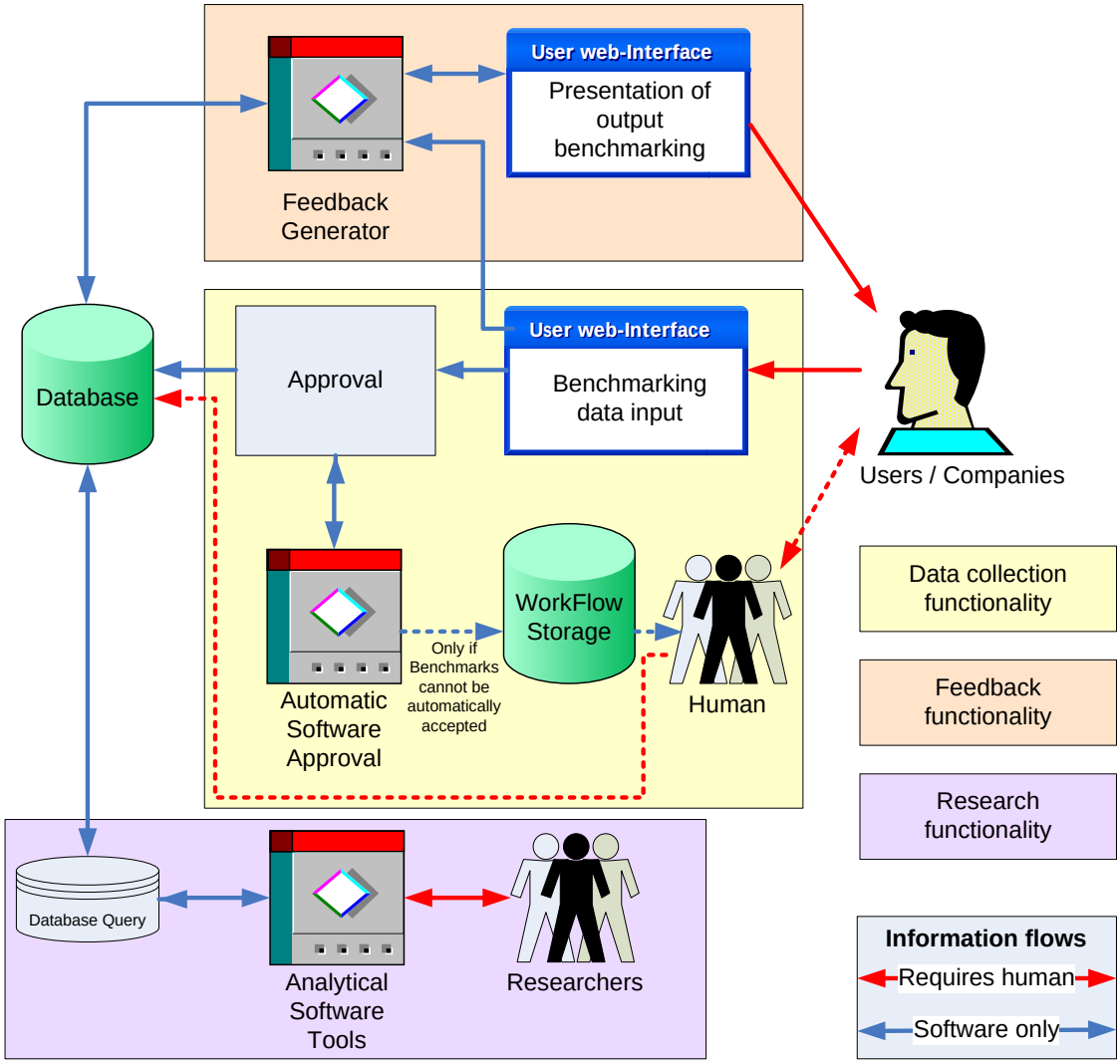
# 4 Technical design of the PROMIT benchmarking self-assessment tool

In this chapter we consider how the tool should be ideally designed and developed. Here we aim at a full-fledged and fully functional system that can handle rather big volumes of benchmarking data and provides a robust user interface for data collection and analysis. Having explained the desired functionality of the tool in the previous chapter, we describe what had to be done by TNO and what activities were to be executed by other parties, as well as what compromises functionality / usability and implementation complexity can be done.

The main principle of the PROMIT benchmarking tool is that it stores benchmarking data in database and provides users with feedback on their benchmarking data (i.e. relative benchmarking position of the company under consideration). A full-fledged tool should be based on client-server architecture: the users, using web browsers, should be able to enter their benchmarking data and get back assessments from the system. Further, in a broad sense, the tool should provide functionality in respect to the following 3 major tasks:

1. **Data collection functionality.** Users (companies) should be able to enter their benchmarking data in a convenient and unambiguous way.
2. **Feedback generation functionality.** The users (companies) should receive an immediate feedback on their data, in other words there must be implemented the principle of immediate gratification.
3. **Research functionality.** It should be possible to provide (restricted) access to the database for research purposes.

The following figure depicts the functionality and information flows of a full-fledged PROMIT self assessment tool.




In essence, the tool is an interface between users (companies), researchers and other groups of users and a database. Let's look at information flows within the tool and software pieces that are required to facilitate the flow.

1. The user enters benchmarking data of his company into a web-based form. There are many examples of such an interface: for instance online shops. They first require users to register (user identification) and then enter their data such as location, payment method, etc. Here we think of a similar process: first companies should register, and then they should be allowed to enter their benchmarking data. Benchmarking data entry process depends on amount of knowledge that the company has on its operations, thus the application guides the user while he enters the data. For more information on the data entering functionality and process, see the benchmarking tool prototype. It is also important here that the benchmarks are well documented and unambiguous. The users will be provided with help that should explain each field.
2. After entry, the data is split into two processes: approval of the data (before it is actuated in the database) and generation of the feedback (benchmarks) that will be presented to the user.

The idea behind such a split is that wrong data should not be allowed into the database; however, the users should be given feedback immediately.

Because of the possibility of misuse of the benchmarking system, there was a discussion on whether it is necessary to provide immediate feedback or later, after the entered data has been approved. There are actually two points of view on when feedback is provided to the user.

a) Feedback is provided immediately
b) Feedback is only provided when approval process is completed

For more information on approval process, see section 3 of the list, which describes the approval process and motivates it.

If there is a vigorous user-registration process, namely, not everyone is allowed to register or registering involves application process such that identities of companies (users) are checked, then feedback can be provided immediately. It is due to the fact that the companies would not want to spoil their reputation and "trustworthiness" within the system.
The second way, namely providing feedback only after an approval, is better if there is an open registration process when every company can register with the benchmarking system without any identity checks. A user (company) in this case may have as many accounts as it wants and may misuse the system, trying to study benchmarking data just entering garbage data. To prevent such a misuse, it is indeed necessary to provide feedback only after approval. At this stage TNO is more inclined towards providing feedback immediately. This is because the first selection of users will be known and trusted. When the system is fully operational and (if!) there is an open registration, then for the users that registered without background checks the feedback data could be shown only after approval.

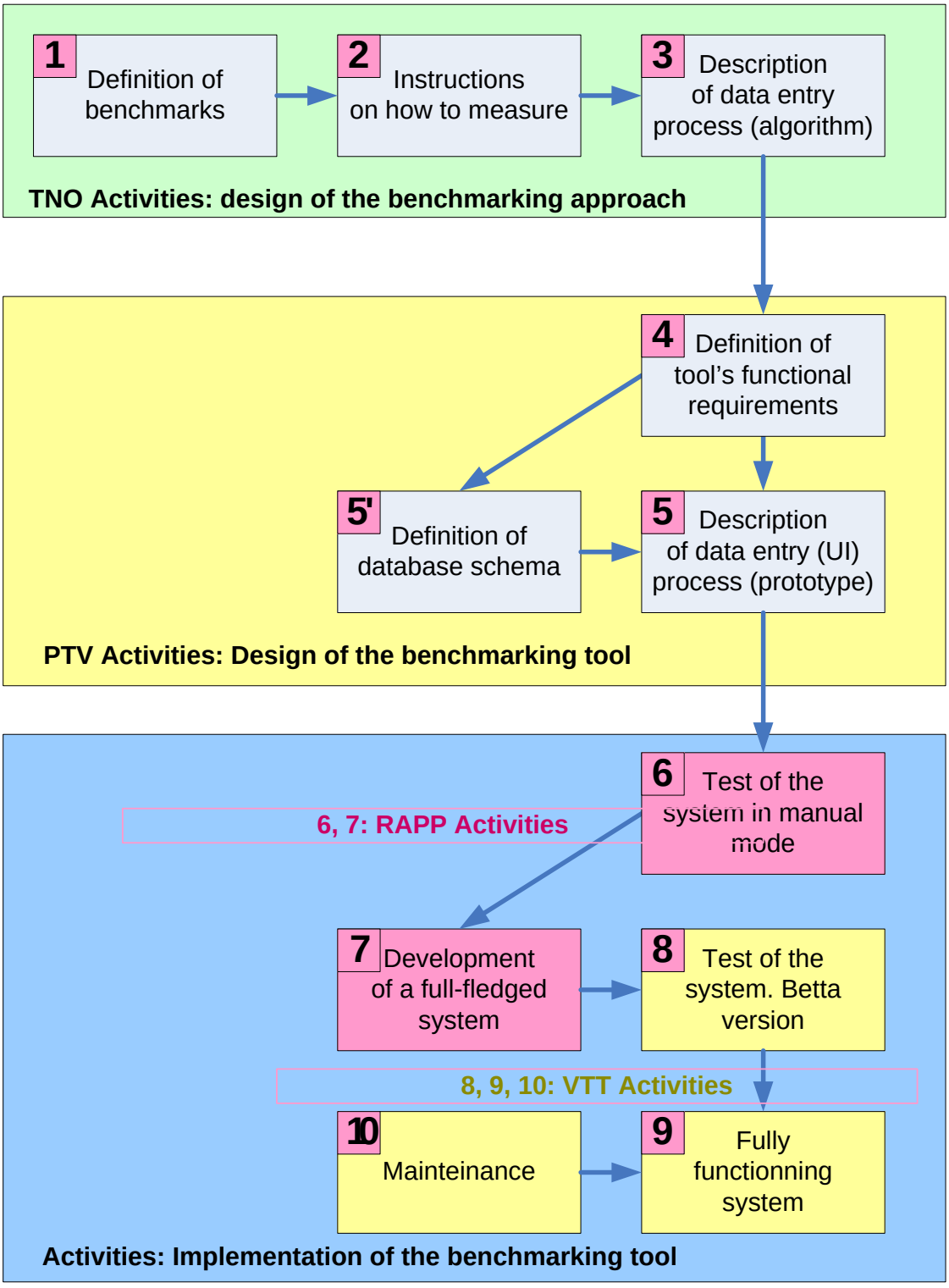
3. When the user has entered his benchmarking data, it must be approved. Unapproved data can cause serious distortions in the database benchmarking data (see Vision: how should it work in practice chapter). Thus, there must be some measures taken to prevent outrageously wrong data to be entered into the database. This can be implemented by a software filter in combination with some human operator assistance.
    3. First, the data entered is assessed by an automatic software assessment tool. The tool can be simple, for instance by allowing only data that is within some range of the averages in the database. The assessment tool can be somewhat smarter, for instance, allowing a balanced assessment of data. However, the main aim of the automatic assessment tool is to allow into the database only valid data: in more difficult cases there will be a need of a human assessment. Thus, the tool just simply reduces workload of human operators who are (in the end) responsible for correctness of the input. If data is OK according to the assessment tool, it is stored in the database. If the data is not OK or questionable, it is stored to be reviewed by humans.
       What important is that the software assessment tool sensitivity should be tilted towards false alarms (as opposed to allowing acceptance of wrong data), guaranteeing only correct data in the database.
    4. If the data is not OK according to the software assessment tool, it is stored as a workflow for human assessors. In principle, the storage physically may take place in the same database where the actual performance data is stored, however marked as needed to be approved (it is technically advisable, because the data is in the same format). The purpose of this storage is to decouple process of entering data from its confirmation, namely to make the processes asynchronous and independent (e.g. working hours, etc)
    5. Human assessors get data for approval from a workflow database. If the human assessor decides that the data is OK, it is immediately actuated in the database. If the human

assessor does not trust the data, he or she can delete it or contact the company that entered data for further clarifications.

4. As we mentioned above, after entry the data processing is split into two processes: approval and generation of feedback, thus they can be executed asynchronously. The approval process may require some time, while the users (companies) may expect an immediate gratification: getting results instantly after data entry. For this purpose, there is a feedback generator software module. In principle, it is a relatively simple piece of software; its concept is described in the Chapter 3 Vision: How it should work in practice. Main functionality here is to benchmark user's data with the averages from the database. Extended functionality may present some additional work, for instance, giving feedback while taking into account weights, different queries, and so on. The essence of the feedback generator is to perform queries and calculate the benchmarking results
5. The benchmarking results should be presented to the user: user web interface is intended to present data in a user-friendly form. This approach is discussed in the document on PROMIT tool prototype.
6. Benchmarking DATABASE stores the benchmarking data. The database should be a relational transactional database. Ideally, an industrial database should be used: Oracle, MS SQL, or others that are typically used in the IT industry.
7. The benchmarking tool should also provide research functionality. It is difficult to foresee possible research questions that can appear in the future, however, the tool should have some research and statistics gathering capabilities. It could be also necessary to restrict access to the confidential benchmarking data, namely providing access only to aggregated data, but not to the individual benchmarking records.

# 5 Proposed implementation protocol

In this chapter we describe implementation issues of the PROMIT self assessment benchmarking tool as they were foreseen at the design stage. Here TNO build up a development execution path, involving several partners, leading to a successful implementation of the benchmarking tool. The following figure describes the process of development of the tool from the concept to a full-fledged system.




In essence, the tool is a typical web-based database application. There is hardly any innovation in programming techniques in such a type of applications; they are regularly developed by the

commercial software application developers. The examples are any web-based systems, where the users must have an identity and can enter their data, for instance, internet shops.

However, the innovation is located in the dimension of functionality: how the benchmarking data has to be collected. This task has much bigger added value, and thus it is reasonable to concentrate on it.

The figure is separated into three big logical groups. The **first group** concerns the development of benchmarking concept and benchmarks themselves. It has been specified in the first year PROMIT report D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report, the description of data entry process has been described in chapter 3. The **second group** concerns the design of the software tool. There must be functional requirements defined, developed database schema, and user interface defined. The **third group** concerns implementation of the tool. First, the tool can be tested in the "manual" mode. Then a fully functional system has to be programmed and developed. After this, it must be extensively tested in adjusted. When there is a fully functional system, it will need to be maintained.

The three above described activity groups are the path to a fully functioning PROMIT self assessment tool. The following list looks at each activity in detail.

1. **Definition of benchmarks.** The benchmarks have been defined in the TNO PROMIT 1st year report D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report. The list of benchmarks is complete.
2. **Instructions on how to measure.** The benchmarks described in the TNO 1st year report (D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report) are well defined, especially in the context of structuring of transportation activities into levels. However, the tool must provide users with short and clear definitions of the benchmarks.
3. **Description of data entry process.** The data entry process is well though of, however, it needs to be proven. The proof should consist of: first, whether the used approach (first enter data for the 1st level, then repeatedly for all activities on the 2nd level) is good enough; second, whether such an approach works in practice, namely whether companies have information and patience to go trough it. There are two ways to test the process: first it should be tested conceptually, among partners and finally, the ultimate test of the approach is the popularity of the deployed tool.
4. **Definition of tool's functional requirements.** In short, the tool should be able to collect, store, process and present benchmarking data. The functionality should be user friendly and persistent. The draft functionality is discussed in the previous chapter. It is apparently the role of TNO to develop it; however, PROMIT partners might be involved in critical review of it.
5. **Definition of database schema.** The benchmarking data must be stored in a database. The database needs to be carefully thought of: which tables are in it, what is the relationship between tables, how to make it fully transactional, etc. Professional (IT) competence is desirable for this phase.
6. **Test of the system in manual mode.** Before development of a full fledged tool, it is desirable to start data collection in "manual mode". Namely, to start filling in database, without help of sophisticated software. This will reach two goals: the database will be filled in with initial data and the process and concept will be tested. To facilitate the process, the tool prototype can be made more functional. Namely, data storage can be added.

7. **Development of a full-fledged system. In essence this is development of a client-server application. This involves the following activities**
    a) Set up of a database server on an industrial platform. The database should be implemented according to the database schema described in activity 5. Additional questions are: where the server must be located, security procedures, redundancy for safety of data and other technical questions.
    b) Design of the user web interface. It includes not only implementation of web entrance forms, but various checks. For instance, there are numerical fields and there only numerical information can be entered. Further, some fields may have a range, for instance, from 0% to 100%. Some fields may become only active if an option is used. In case if a company enters data on indicators at the second level, there must be a loop of data entry process implemented. Also it includes presentation of benchmarking results to the company.
    c) Coupling of the server with web forms. This involves not only storage of the entered data, but organization of workflow for human assessors and so on. In other words, there must be some logic implemented on the server side.
    d) There must be functionality implemented for support of approval process. It will involve build up of workflow for human assessors, support of their decisions and some user interface for them (not necessarily a web interface)
8. **Test of the system (beta version)**. When the system is fully implemented, it must be extensively tested, partly in a fully operational mode. The system critically needs testing: there will be many users (if successful there will be hundreds or thousands of users). It is highly undesirable to put a very raw system on a high scale use.
9. **Fully functioning system**. After beta testing has been performed, the system is fully functional.
10. **Maintenance**. When the system is operational, it must be taken care of. The maintenance includes the following activities:
    a) Technical maintenance. This includes maintaining hardware equipment in operational state, aiming at 24/7 availability of service
    b) Functional maintenance. There could be changes in benchmarks, data collection and analysis processes, etc.
    c) Customer service. The system deals with a big number of companies, thus some of them will find it difficult to work with the system. There must be someone who helps the users of the system.

# 6 Actual implementation of the tool

In November 2007 the PROMIT benchmarking self assessment tool became operational. The actual implementation of the tool has been carried out by PTV, which developed web-based data entry interface for indicators on the levels 1 and 2. The first version of the tool does not implement all envisaged functionality. However, it allows a quick start in gathering of benchmarking data. The web interface has some logic that tests validity of the data and stores entered data in a database. Also the process of identification of users has been implemented.

The current implementation of the tool represents a compromise on the described in the previous chapters design. There are two main compromise points. The *first compromise* is in the validation of data entered: there is no validation implemented. Thus data entered by the users of the tool become active in the database immediately upon entrance. The *second compromise* is in the benchmarks: not all benchmarks that have been discussed in the PROMIT first year report (D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report) and in the Chapter 3 Vision: How it should work in practice are implemented. The main reason for the first compromise is complexity of implementation and maintenance process. The main reason for the compromise in the number of benchmarks is the shire complexity of the data entry process and, namely, complexity for the end users of the tool.

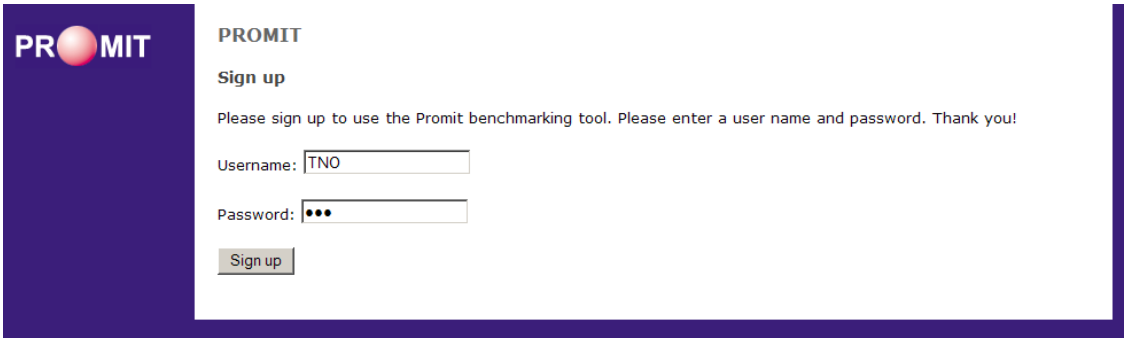
It is important to acknowledge that the tool is still in the early phases of development, consisting of a software engine that provides only basic functionality. Now it is being tested with entrance of benchmarking data and filled in with data that provides a critical mass of benchmarking information. Only when it is filled in with some useful benchmarks it will become attractive for private users and the benchmarking process will become self-sustainable, to extend analogy with nuclear reactions. At the same time, the process of entering initial benchmarking data provides a test environment for the tool and allows beta testing.

Conceptually, the tool provides functionality in 3 main areas:

1. Registration of users (account maintenance)
2. Data entry
3. Benchmark generation (feedback)

## 6.1 Registration of users

The tool allows entering of benchmarking information only by registered users. Therefore, to be able to use the tool, a user needs to create an account. In order to do that, the user has to go to the sign up page, which is located at http://www.ptv.de/promit/register.php (there is a link to it from the main page).




The current implementation of registering is very simple. Only user name and password are required to create an account. No further information is asked about company identity and other information. Such simple registration process is easy for users, because it requires minimum time to execute and there is no worry about keeping company's identity confidential. However, it is not so handy for research and integrity purposes. Thus, we propose that the following versions of the tool implement more extensive registration procedure. When an account is created, the user needs to log in at http://www.ptv.de/promit/index.php. After the user is logged in, he may proceed with the data entry process.

## 6.2 Data entry

Data entry starts at level 1 indicators. The users must enter the following parameters

1. Origin and destination, which include
    a) Zip Code
    b) Town
2. Type of shipment: the user may choose from
    a) 40-ft container
    b) Trailer
    c) Swap body (13,60 m)
3. Distance between origin and destination points, km
4. Frequency of this type of shipment, in the following brackets
    a) Once per day or more
    b) Once per day – Once per week
    c) Once per week or less
5. Shipment costs in Euros
6. Lead time in hours (from door to door)
7. Reliability of service, defined as perceived deviation from specified arrival, time in the following time brackets
    a) Less than 1 hour
    b) From 1 to 3 hours
    c) From 3 to 6 hours
    d) From 6 to 12 hours
    e) From 12 to 18 hours
    f) From 18 to 24 hours
    g) From 24 to 36 hours
    h) From 36 to 48 hours

i) From 48 to 72 hours
j) From 1 to 3 hours
k) More than 72 hours

8. Assessment of flexibility of the service provider regarding additional services on a scale from 1 to 10, where 1 is low and 10 is high
9. Assessment of quality of the information exchange with the service provider on a scale from 1 to 10, where 1 is low and 10 is high
10. Assessment of the quality of payment procedure on a scale from 1 to 10, where 1 is low and 10 is high
11. Transport modes that the transportation includes (multiple choice)
    a) Road
    b) Rail
    c) Inland waterways
    d) Short sea


If the shipment involves more than one mode, then the user can enter level 2 benchmarks. He is asked to enter parameters described in the list below. What is important here is that the tool divides the whole transportation chain into segments (operations) specified at the level 1 and asks to enter KPI's consistently. The tool, for instance, checks that the sum of distances (and the sum of costs) specified in the level 2 segments is equal to the total shipment distance / cost specified for the level 1 indicators.

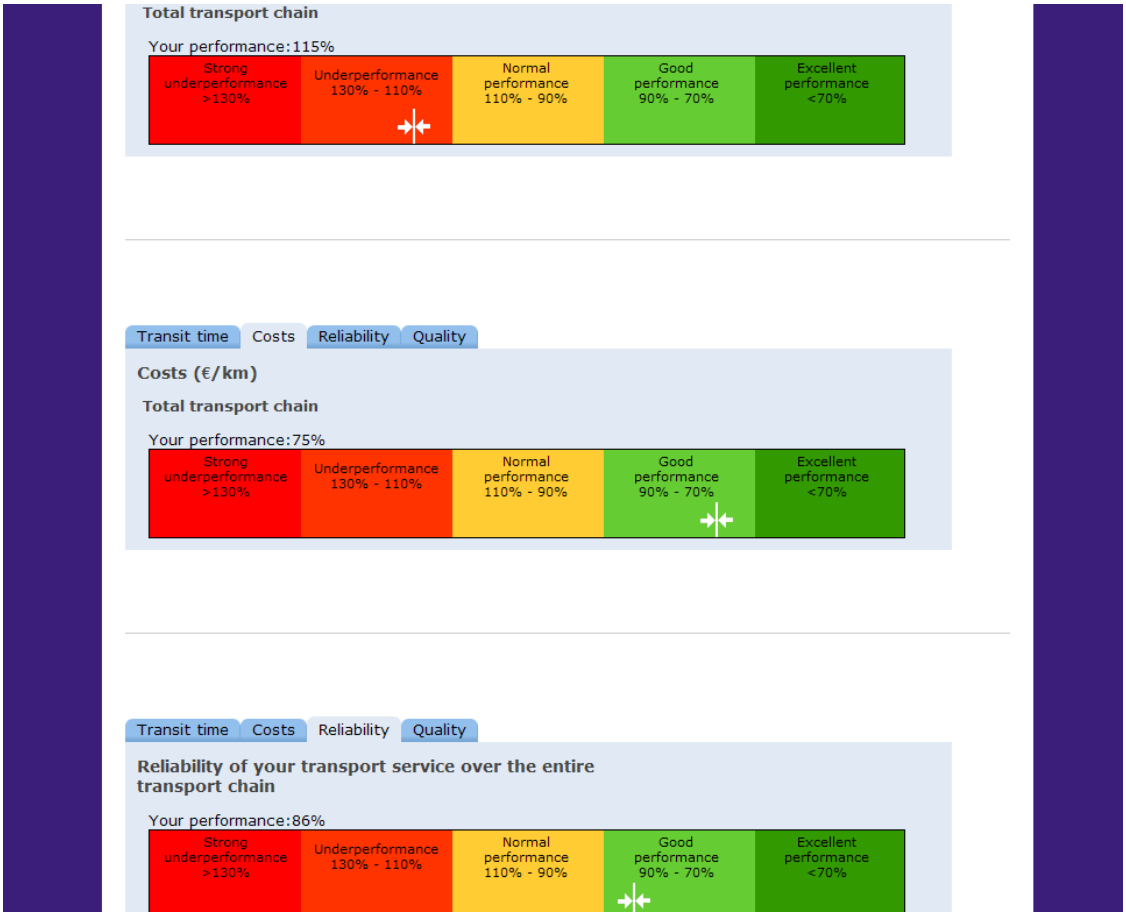
The tool also makes sure that all compulsory fields are filled in with appropriate information, for instance, it is not possible to leave them blank or unspecified and it is not possible to enter some text in the fields that require numerical information.

1. Distance
2. Transit time
3. Cost

The process of benchmark acceptance / validation is not implemented in the current version of the tool. Implementation of them could be costly and maintenance of such a process would require a dedicated person(s) who would monitor entered benchmarks. This is a trade-off that is made in order to get the first version of the tool operational. In future it is highly desirable that benchmarks entered are monitored and the acceptance process is implemented. In this way it is possible to guaranty quality of the benchmarking database. There are also some benchmarks omitted, for example energy use. There is still an ongoing discussion on what should be measured. On the one hand, it is highly desirable to measure all indicators specified in the 1st year PROMIT report D 4.1 European Benchmarks in Intermodal transport, 1st year PROMIT TNO report; on the other hand, complexity of such benchmarks could scare off potential users of the tool.

## 6.3 Benchmarking feedback generation

When the user correctly completes entering data, he immediately gets assessment of the entered benchmarks. The assessment is based on comparison per indicator to the averages of respective indicators in the database. For instance, if the average road cost per kilometre is 1,00 Euro per km and the user entered his cost of 1,20 Euro per km, then his performance is 20% worse than average and falls in the category of underperformance.




I addition to feedback generation, the tool adjusts the database with the newly entered benchmarking information. Thus, the database contains a new record(s) and following users will get their indicators compared to new averages.