Hungarian Association of Agricultural Informatics
European Federation for Information Technology in Agriculture, Food and the Environment

Journal of Agricultural Informatics. Vol. 14, No. 1
journal.magisz.org

A process-oriented data warehouse to support decision making process in crop production

János Felföldi[1], Szilvia Botos[2], Levente P. Bálint[3], Dénes Sulyok[4], Sándor Csenki[5], Viktor László Takács[6]

INFO
Received: 30.06.2023.
Accepted: 05.07.2023.
Available on-line: 07.07.2023.
Responsible Editor: László Várallyai

Keywords:
crop production, operations of technology, interval parallelism, process modelling.
JEL: C41, O13, M11

A B S T R A C T

In our article, we present the structure of a process-oriented data warehouse created according to the OLAP (online analytical processing) principle that enables simple process controlling in crop production with 1 hectare process cost and time. To identify the sequence of operations and parallelism of the processes, we used interval ranking theory, which was implemented in the Power Query M formula language.
The data available at the right time and in the right way and the economic calculations that can be produced from them, as well as the presentation of the produced information to managers, are crucial in the decision-making process for the organizations. Wide range of digital solutions has an important role to support the business activities and the decision making in the case of farm management as well.
It is important to determine the structure of the data required to base the economic calculations, build a data model and create a multidimensional database, this lays the foundation for the calculation of economic and financial indicators, which can help in making economic decisions.

1. Introduction

1.2 The role of digitalization in the decision background of the agriculture

In our article, we present the structure of a process-oriented data warehouse created according to the OLAP (online analytical processing) principle that enables simple process controlling in crop production with 1 hectare process cost and time. To identify the sequence of operations and

[1] János Felföldi
University of Debrecen, Faculty of Economics and Business, Institute of Applied Informatics and Logistics, Department of Logistics Management
felfoldi.janos@econ.unideb.hu
[2] Szilvia Botos
University of Debrecen, Faculty of Economics and Business, Institute of Applied Informatics and Logistics, Department of Business Informatics
botos.szilvia@econ.unideb.hu
[3] Levente P. Bálint
University of Debrecen
balint.peter.levente@gmail.com
[4] Dénes Sulyok
University of Miskolc
sulyok.denes@uni-miskolc.hu
[5] Sándor Csenki
University of Miskolc
Discovery Center Ltd.
info@agridron.com
[6] Viktor László Takács (orcid 0000-0001-8433-6115)
University of Debrecen, Faculty of Economics and Business, Institute of Applied Informatics and Logistics, Department of Business Informatics
takacs.viktor@econ.unideb.hu

parallelism of the processes, we used interval ranking theory, which was implemented in the Power Query M formula language. The Technology has a key role in the different operational and decision-making processes of the organizations (Panetto et al., 2020). Technologies like cloud services, Internet of Things (IoT) and Big Data analytics are highlighted by Annosi et al. (2019) that can be assumed important both in the agriculture and the food industry. Cloud services can support the virtualization of business processes to establish an ICT-based intelligent agricultural information system (Látečková et al., 2018). The contribution of IoT to the change of agrifood business processes and to the data centric management is considered relevant throught the decision support tools (Verdouw et al., 2016). The accessibility of information and knowledge is required in the agricultural sector as it may increase the productivity through a better decision making process of the farm managers (Ali-Kumar, 2011). Big Data means a tool in the Agriculture 4.0 concept that can help the development of specific marketing strategy (Lezoche et al., 2020). An additional benefit of the application of different digital solutions is a potential improvement of the cooperation between supply chain stakeholders (Singh et al., 2019). The strategic decisions on digital transformation, however, do not mean development in the performance automatically, the business model of the organization requires a complete rethink and change (Bouwman et al., 2019). In order to optimize the internal business processes of the enterprises a more intensive use of digital technologies is needed (Ahmedova, 2015). A strong and dynamic corporate entrepreneurship is needed to exploit the benefits of digital technologies by integrating them into organizational strategies and creating the right corporate culture for digital adoption and use (Yunis et al., 2017). Recent research papers analyse different aspects of digital agriculture in relation with farming systems and value and food chains (Klerkx et al., 2019).

The agricultural sector has undergone significant innovation in the last two decades, aimed at improving efficiency and competitive advantages. Farming enterprises have transformed from simple production units to multifunctional businesses. Effective farm management is crucial for financial survival and sustainability in today's competitive environment (Vasileva et al, 2020).

Challenges such as limited time for in-field monitoring, financial management difficulties, and the need for decision support (Sørensen et al, 2011) have emerged due to the increasing complexity of farming operations (Paraforos, 2017). Farmers require additional information and suitable technologies, including Big Data, to monitor and manage data collection in the field. Real-time information and informed decision-making are key components to success (Shankarnarayan-Ramakrishna, 2020).

Policy reforms promoting multifunctional rural economies have created uncertainties and pressure on farms to become more entrepreneurial and adopt new technologies (Morris et al, 2017). Technological advancements have revolutionized farming practices, enabling eco-friendly and sustainable methods. Smart farming, integrating modern technologies like big data, machine learning, and robotics, empowers farmers to make data-driven decisions (Sharma et al, 2022). Web-based tools can enhance the capabilities of agricultural enterprises to address challenges and ensure future success (Subhashree, 2023).

Financial analysis tools have evolved to predict these problems and minimize risks. Adequate accounting and analytical support principles are also crucial for effective decision-making (Bondina, 2021). Production costs and additional costs for environmentally friendly products have been studied before widely in the professional literature with great success (Peña et al, 2022).

Technology supports business models, improving production scale, cost management, and diversification. Information and communications technology provide real-time access to useful information, enhancing productivity and adaptability Web-based tools simplify complex economic analysis and support data-driven decision-making (Sopegno et al, 2016).

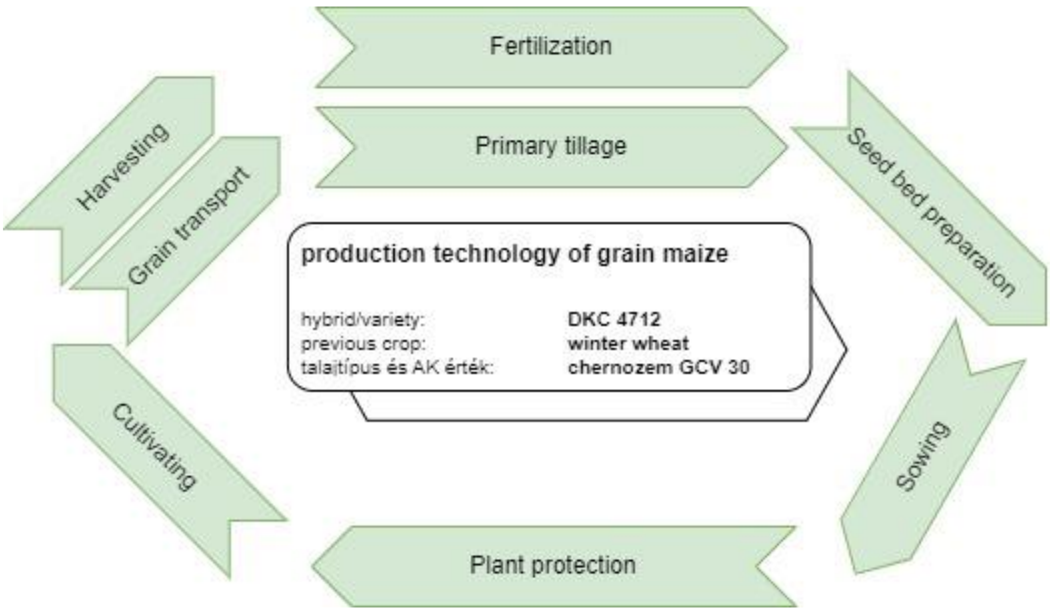

Figure 1. Production technology flow chart of grain maize
Source: Own edition

2. Methodology

In this study, we present the system that has been created by processing electronic data of the Technology process sheet and which can be used that concern managers, such as

- Share of total on-farm crop production expenditures by operation name
- Share of total working hours by operation name
- Share of production input material expenditures by material types
- Cost composition of operations for 1 hectare
- Cost composition of materials for 1 hectare
- Machinery composition for crop production on total land

The system's indicators are a simple resource cost for 1 hectare with resource quantity and time calculations for process and its operations are presented on executive dashboards using various visualisation that are clear and understandable to everyone. We used data warehouse technology for data processing.

Defining and implementing the system and its design pose considerable challenges, allowing only for an approximate understanding of its scope. Different individuals have proposed numerous definitions, one of which states, *"In a system thus we mean a whole that consists of elements, things, objects, element relations, relationships, relations, and connections that have certain properties."* (Szenteleki, 2007, p.10)

Similarly, a consistent definition of a data warehouse stays elusive. Bill Inmon and Ralph Kimball are widely recognized figures in data warehouse design. In Kimball's book, The Data Warehouse Toolkit, he defines a data warehouse as *"The conglomeration of an organization's data warehouse staging and presentation areas, where operational data is specifically structured for query and analysis performance and ease-of-use."* (For the Hungarian translation see: Szatmári, 2009, p.7)

Operational support systems within an organization can be categorized into two primary groups: transaction-oriented OLTP (Online Transaction Processing) and analysis-oriented OLAP (Online Analytical Processing) systems. The distinction between these two types lies in the fact that transaction-oriented systems facilitate daily tasks and event tracking, while OLAP systems focus on long-term information gathering and decision support. OLAP systems store and present the company's historical data. The ETL (Extract-Transform-Load) process serves as an intermediary step between these systems, involving the extraction of data from transaction systems at regular intervals, potential data cleansing, transformation to enable reporting, and finally, loading into the data warehouse structure. Transactional systems are characterized by a high number of users responsible for recording daily events, while OLAP systems are accessed by decision-makers and managers seeking information. The design of OLTP entails an individual-relationship model with detailed data resolution, while OLAP deals with processing object-oriented data previously saved. (Csepei, 2009)

Two methodologies are well-known for classical data warehouse design: data-driven (Inmon) and requirement-following (Kimball) approaches. Both methodologies have their limitations. The data-driven approach may hinder answering numerous questions from larger, more complex data sources due to tangled data. Conversely, the requirement-following method carries the risk of creating a cube that can only address a single question.

Consequently, both methodologies show advantages and disadvantages, leading to the emergence of hybrid design methodologies in the past decade and a half. These methodologies aim to use the strengths of both approaches while mitigating their respective drawbacks. When designing our system, we employed such a hybrid methodology (Takács et al., 2020).

1. Requirement Analysis: During this phase, we gather questions, metrics, and dimensional requirements related to the problem by conducting interviews with decision makers. We use a specialized structured stenography based on the problem-specific terminology to formalize these specifications. The outcome of this step is a set of formalized questions.

2. Deriving Minimal Granularity: Using the set of formalized questions, we determine the minimum level of detail required for each indicator. This step results in a set of indicators with dimensions that are detailed to the necessary extent. Deriving Ideal Schemata:

3. Ideal schemata: We establish mappings between dimensional attributes and their corresponding keys, creating the initial conceptual schemata. This step yields ideal dimensions (consisting of keys, attributes, and hierarchies) as well as ideal facts (with dimension keys for join), irrespective of the data sources.

4. Source Analysis: We break down the ideal facts into potential elementary transactional attributes and identify their presence in the source systems. The outcome of this step is the derived potential schemata.

5. Integration: We compare the ideal schemata obtained from the requirement analysis with the potential star schemata. A match is found when both schemata contain the same fact and exhibit the same dimensional characteristics at the same granularity level. During this step, we define the necessary transformations, calculate fact tables, and establish common dimensions with their associated attributes.

6. Multidimensional Modelling: We construct the cube(s) by incorporating dimensions, dimension hierarchies, and measures, resulting in a comprehensive multidimensional model.

3. Results

3.1 Requirements analysis

In the requirements analysis phase, we collect and formalize managerial questions. The question and its description are defined in a textual and a formal way. A "manager/leader" is a person whom the system provides with information. As a result, they expect the report to appear in the form of visualized data on an IT interface. Reports and dashboards must be structured according to the requirements of access rights (it is not all the same, what level of managers can access certain reports).

In the requirements analysis, we "analyze" management questions based on the following aspects of Table 1: What is the indicator? In which aggregation? What is the unit of measure? What kind of visualization would we like to see? At what resolution? What kind of slicers can be applied?

Table 1. Management question analysis

Indicator unit(s)	$\boldsymbol{I}^{\{u[,u]\}}_{\{\mathrm{af}[,\mathrm{af}]\}}$	the indicator $\boldsymbol{I}$ to be produced with u unit(s) in the upper right index and af aggregate

aggregate function(s)		function(s) in the bottom right index,
visualization	$\begin{pmatrix} vt \\ \left[\begin{Bmatrix} S \\ s \end{Bmatrix}\right] \end{pmatrix}^{v}$	the *v* visualization with the type *vt* (table, line diagram, bar graph, etc. ...) and optional *s* slicers (values can be $\boldsymbol{D}_{\{a\}}$ dimensional attribute, $\boldsymbol{D}_{\{v\}}$ subset of concrete values, or a $\boldsymbol{D}_{\{a\}}$ dimensional attribute in the *d* detail of another $\boldsymbol{I}$ indicator on the same dashboard)
slicer(s)		
detail(s)	$\left[\begin{pmatrix} \boldsymbol{D}_{\Sigma\{a\}} \\ [\boldsymbol{D}_{\Sigma\{a\}}] \end{pmatrix}^{\{d\}}\right]$	*d* details with $\boldsymbol{D}_{\{a\}}$ dimensional attribue(s), with optional $\sum\{a\}$ aggregation. *d* values e.g.: *row*, *col*umn, *cat*egory, *y* indicator

Source: Takács et al. (2020)

Formally: $\boldsymbol{I}^{\{u[,u]\}}_{\{\mathrm{af}[,\mathrm{af}]\}} \begin{pmatrix} vt \\ \left[\begin{Bmatrix} S \\ s \end{Bmatrix}\right] \end{pmatrix}^{v} \left[\begin{pmatrix} \boldsymbol{D}_{\Sigma\{a\}} \\ [\boldsymbol{D}_{\Sigma\{a\}}] \end{pmatrix}^{\{d\}}\right]$

The formal definition above covers the configuration and the display of indicators, and these are related to one diagram. Several different visualizations could arise in the manager, related to the requirement specification, so many formal descriptions will be made in this section.

Table 2. Management questions and dashboards.

Logical Description	Formally
Machinery composition for crop production on total land	$\boldsymbol{D}^{day}_{\mathrm{sum}} \begin{pmatrix} gantt \\ \mathbf{dC}_{\text{grain maize}} \end{pmatrix}^{v} (\mathbf{dP}_{\{name\}})^{parent} (\mathbf{dO}_{\{name\}})^{task} (\mathbf{dP}_{\{SD\}})^{SD} (\mathbf{dP}_{\{ED\}})^{ED}$ $\boldsymbol{D}^{day}_{\mathrm{sum}} \begin{pmatrix} waterfall \\ \mathbf{dC}_{\text{grain maize}} \end{pmatrix}^{v} (\mathbf{dO}_{\{name\}})^{cat}$
Share of total on-farm crop production expenditures by operation name Share of total working hours by operation name	$\boldsymbol{C}^{\%}_{\mathrm{sum}} \begin{pmatrix} pie \\ \mathbf{dC}_{\text{grain maize}} \end{pmatrix}^{v} (\mathbf{dO}_{\{name\}})^{legend}$ $\boldsymbol{D}^{day}_{\mathrm{sum}} \begin{pmatrix} pie \\ \mathbf{dC}_{\text{grain maize}} \end{pmatrix}^{v} (\mathbf{dO}_{\{name\}})^{legend}$
Cost composition of operations for 1 hectare	$\boldsymbol{C}^{Ft}_{\mathrm{sum}} \begin{pmatrix} waterfall \\ \mathbf{dC}_{\text{grain maize}} \end{pmatrix}^{v} (\mathbf{dO}_{\{name\}})^{cat}$ $\boldsymbol{C}^{Ft}_{\mathrm{sum}} \begin{pmatrix} waterfall \\ \mathbf{dC}_{\text{grain maize}} \end{pmatrix}^{v} (\mathbf{dR}_{\{type\}})^{cat}$ $\boldsymbol{C}^{\%}_{\mathrm{sum}} \begin{pmatrix} pie \\ \mathbf{dC}_{\text{grain maize}} \end{pmatrix}^{v} (\mathbf{dR}_{\{type\}})^{legend}$
Cost composition of materials for 1 hectare Share of production input material expenditures by material types	$\boldsymbol{C}^{Ft}_{\mathrm{sum}} \begin{pmatrix} waterfall \\ \mathbf{dC}_{\text{grain maize}} \\ \mathbf{dR}_{\text{material}} \\ \mathbf{dR}_{\text{operational}} \end{pmatrix}^{v} (\mathbf{dR}_{\{name\}})^{cat}$ $\boldsymbol{C}^{\%}_{\mathrm{sum}} \begin{pmatrix} pie \\ \mathbf{dC}_{\text{grain maize}} \\ \mathbf{dR}_{\text{material}} \\ \mathbf{dR}_{\text{operational}} \end{pmatrix}^{v} (\mathbf{dR}_{\{name\}})^{cat}$

Source: Own edition

3.2 Source analysis

In this project, our source is clearly based on the Technology process sheet and its electronic version. It shows the lead-time, resources, and its costs shown in the Figure 1. Our source system can be considered a hybrid system on the axis between a transaction-oriented and an analytics-oriented systems. It is closer to a transaction-oriented approach, while keeping a record of the steps of the plant cultivation technology, but the recording does not strictly follow the principle of recording transactions but is like filling in different cells in a form from the result as shown in the Figure 2.

production technology of grain maize		2022 Year
hybrid/variety	DKC 4712	
previous crop	winter wheat	
soil type and quality (golden crown value)	chernozem GCV 30	
comments:		

Operation 7:	harvesting	used input quantity per unit of land	unit	cost of input unit	unit	cost of per unit of land	unit
time period:	2022 10.5-10.15						
machinery:	John Deere T 660+Geringhoff Rota Disc	1	pcs			43 510	Ft/ha
power-machine	combine harvester	1	pcs	27 820	Ft/ha	27 820	Ft/ha
implement	corn header	1	pcs	4 500	Ft/ha	4 500	Ft/ha
machinery operator	driver	1	head	1 400	Ft/ha	1 400	Ft/ha
material to machinery	fuel	22	liter	445	Ft/l	9 790	Ft/ha
material to operation:						-	Ft/ha
material to operation:						-	Ft/ha
service to operation 7:						-	Ft/ha
total cost of operation:						43 510	Ft/ha

Figure 2. Technology process sheet – an example of harvesting operation
Source: Own edition

Source entities and relationships are based on the Technology process sheet are:

eComp: Company entity

- CompID: unique identifier attribute in format: C[001-999]
- Company name: string,

eCrop: Crop entity

- CropID: unique identifier attribute in format: T[000-999]
- Crop name: Hungarian equivalent of the {rice, wheat, maize, corn, etc}
- Type: Hungarian equivalent of the classification e.g. {siló, szemes, etc…}
- Hybrid/variety: specific subtype of the crop
- Precrop: like Crop
- Soil type: number of documents
- Gold crown value: numeric
- Comment: freestyle text

weProc: Process weak entity

- ProcID: unique identifier attribute in format: F[000-999]
- Process name: Hungarian equivalent of the main Process e.g. {rice, wheat, maze, corn, etc}
- Year: based ont he harvests date in format: {yyyy}
- CompID: foreign key attribute in format: C[001-999]
- CropID: foreign key attribute in format: T[000-999]

eOp: Cultivation operation entity

- OpID: unique identifier attribute in format: M[001-999]
- Operation name: string,

eRes: Resource entity

- ResID: unique identifier attribute in format: EF{A,G,M,K}[001-999], where A, G, M, K are abbreviations from Resource type
- Type: Hungarian equivalent of the classification from {material, machine, work, service}
- Subtype: Hungarian equivalent of the subclassification of the type
- Comment: specific name of the resource, too much variation
- Unit: measurement unit

weProcOp: ProcessOperation weak entity

- ProcID: foreign key attribute in format: F[000-999]
- OpID: foreign key attribute in format: M[001-999]
- OSN: operation sequence number attribute started from 1
- ProcOpID: unique identifier attribute in format: {ProcID}{OpID}-{OSN}
- StartDate: start date of the operation in {yyyy}.{mm}.{dd} format
- EndDate: end date of the operation in {yyyy}.{mm}.{dd} format

rCost: Cost realtionship

- ProcOpID: foreign key attribute in format: {ProcID}{OpID}-{OSN}
- ResID: foreign key attribute in format: EF{A,G,M,K}[001-999]
- Quantity: numeric
- Cost: numeric
- Area (hectares): area size where the cost occurred, numeric3.2. Value of identification technologies in food sector

3.3 Integration phase

Due to the hybrid nature of our source system, we didn't have much to do during the integration phase, we calculated the durations of the operations and the cost of the resources aggregated for the whole process.

The operation mesh table was developed for two purposes: to sequence process operations and to identify parallel operations based on the operations time periods. The algorithm is implemented in Power Query M formula language, and it is based on a simplified interval ranking method (Luo, B., Ye, Y., Yao, N. et al. 2021). Since its introduction by Dwyer (1951), the concept of interval numbers has garnered significant attention from scholars, leading to extensive research on methods for ranking them. Several notable achievements have been made in this area (Dubois and Prade 1987; Nakahara et al. 1992; Ganesan and Veeramani 2006). Ishibuchi and Tanaka (1990) showed a definition for the weak preference order relationship between two interval numbers in linear programming. Furthermore, Firozja et al. (2012) proposed an interval distance metric to quantify the similarity of generalized fuzzy numbers while keeping metric properties.

Definition 1: (Nakahara et al. 1992; Moore and Lodwick 2003) Assume $\tilde{a} = [a^L, a^U] = \{a | a^L \leq a \leq a^U, a^L, a^U \in \mathbb{R}\}$ is an interval number, where a^L and a^U are the upper and lower limits of $\tilde{a}$, respectively.

Definition 2: (Ishibuchi and Tanaka 1990; Ganesan and Veeramani 2005) Assume $\tilde{a} = [a^L, a^U]$ and $\tilde{b} = [b^L, b^U]$; define $\mathrm{P}(\tilde{a} \succ \tilde{b})$ as the advantage degree of $\tilde{a}$ compared with $\tilde{b}$: $\mathrm{P}(\tilde{a} \succ \tilde{b}) \in [0,1]$ and $\mathrm{P}(\tilde{a} \succ \tilde{b}) + \mathrm{P}(\tilde{b} \succ \tilde{a}) = 1$. If $\mathrm{P}(\tilde{a} \succ \tilde{b}) > 0.5$, then $\tilde{a} \succ \tilde{b}$; if $\mathrm{P}(\tilde{a} \succ \tilde{b}) = 0.5$, then $\tilde{a} = \tilde{b}$; and if $\mathrm{P}(\tilde{a} \succ \tilde{b}) < 0.5$, then $\tilde{b} \succ \tilde{a}$.

Definition 3: (Xu and Da 2002; Fan and Liu 2010) Assume $\tilde{a} = [a^L, a^U]$ and $\tilde{b} = [b^L, b^U]$; if $a^L \geq b^U$, then $\mathrm{P}(\tilde{a} \succ \tilde{b}) = 1$ and $\mathrm{P}(\tilde{b} \succ \tilde{a}) = 0$. Thus, $\tilde{a}$ and $\tilde{b}$ are separate interval numbers.

Definition 4: Let $\{\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_n\}$ be a group of arbitrary time period interval numbers, and suppose that $\tilde{p}_i$ and $\tilde{p}_j$ are elements of them and $i, j \in [1, \mathrm{n}]$; define $\cap\left(\tilde{p}_i, \tilde{p}_j\right) = min\left(p_i^U, p_j^U\right) - max\left(p_i^L, p_j^L\right)$

as intersection of $\tilde{p_i}$ compared with $\tilde{p}_j$: $\cap\left(\tilde{p}_i, \tilde{p}_j\right) \in \mathbb{R}$. If $\cap\left(\tilde{p}_i, \tilde{p}_j\right) > 0$, then $\tilde{p}_i \parallel \tilde{p}_j$; and if $\cap\left(\tilde{p}_i, \tilde{p}_j\right) \leq 0$, then $\tilde{p}_i \prec \tilde{p}_j$, if $p_i^L \geq p_j^U$; and $\tilde{p}_j \prec \tilde{p}_i$, if $p_j^L \geq p_i^U$.

Procedure for creating the operation mesh table for the sequence of operations in processes:

Based on the principle of time interval parallelism, the procedure can be summarized as follows:

1. Derive the projection "weProcOp'" from the ProcessOperation weak entity, including the attributes ProcID, OpID, StartDate, and EndDate.
2. Create the Cartesian product "IRM" from the cross join of weProcOp' with itself, matching the ProcID attribute of the entities.
3. Select the records in the IRM table where the Source OpID is not equal to the Destination OpID.
4. Extend the IRM table by adding the "Parallel" attribute based on Definition 4.
5. Select the records in the IRM table that are not parallel.
6. Create a projection of the IRM table including the attributes ProcID, Source OpID, and EndDate, Destination OpID, and StartDate.
7. Extend the IRM table by calculating the "Distance" attribute as Destination Operation StartDate minus Source Operation EndDate.
8. Create a projection of the IRM table including the attributes ProcID, Source OpID, Destination OpID, and Distance.
9. Select the records in the IRM table where the Distance is non-negative.
10. Derive the projection "minDist" table from the IRM table.
11. Calculate the minimum aggregation of the Distance attribute, grouped by ProcID, Source OpID, and assign it as the minDist attribute.
12. Join the IRM table with the minDist table, creating the Operation mesh table, where the ProcID and Source OpID attributes are equal.
13. Select the records in the Operation mesh table where the Distance attribute is equal to the minDist attribute.
14. Create a projection of the Operation mesh table including the attributes ProcID, Source OpID and Destination OpID.

3.4 Multidimensional model

The optimal physical model of our process-oriented OLAP system is shown in Figure 3. Since our source system was uploaded with a data provider form solution and the source system does not closely follow the transaction-oriented paradigm, few modifications were necessary in the integration phase. The factProcessOperation table contains a summary of operation costs and technical yMax for better visualization from the resource fact table, and the time required for the process was calculated, as well as the operation name based on the sequence of activities. For this, an operation mesh table was created in the system showing the sequence of operations per process.

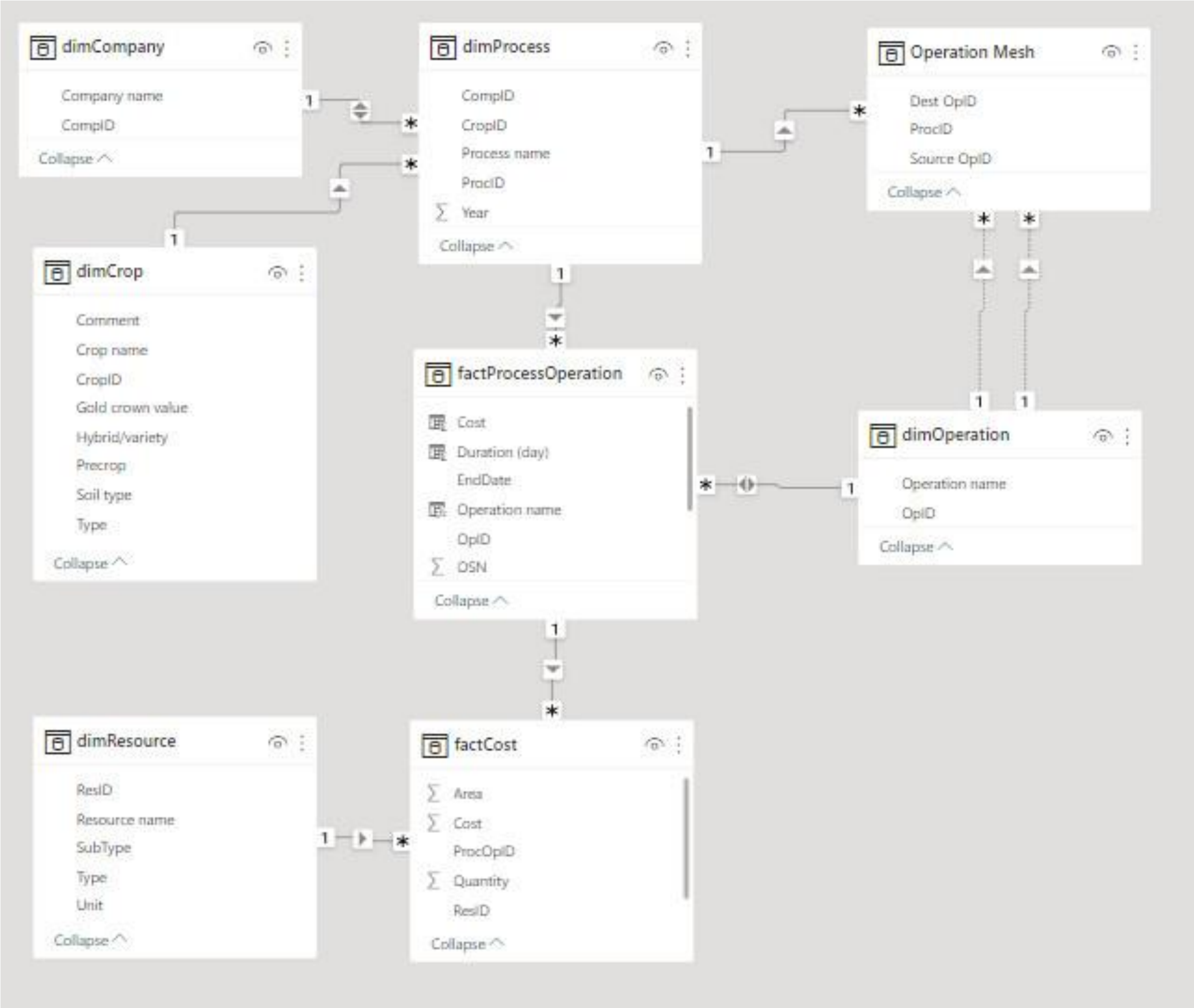

Figure 3. Multidimensional Entity Relationship model of our system
Source: Own edition

4. Conclusions

By now, Clive Humby's mathematical phrase, "Data is the new oil," has become a proverb. Then, Palmer complete: *"It's valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc., to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value."* (Palmer, 2021)

The database established for the storage of data related to crop production processes. The formalized calculations of the indicators that must be generated to answer management questions were determined, and the results were developed and visualized using Microsoft Power BI.

While our quasi-OLAP system required minimal transformations, the identification of the sequence and parallelism of process operations required interesting mathematical and relational algebra approaches in the Power Query M formula language. This method can be further improved to classify operational times as work, wait, or slack time. The method can help the implementation with the analysis that can determine the length of the critical path of the time plan, which shows the completion time of the process as a duration.

The information system created based on management questions contains key time and cost indicators and graphs that can help analyze the efficiency of production and business processes. The formal description of the visualizations and dashboards used to answer these questions was defined in the first step, the requirement analysis phase.

Investigating the possibility of expanding the prepared database with information obtained from other data collection and, if possible, integrating data from different data sources from multiple, large, distributed and heterogeneous databases and other sources of information represents a further direction of research.

Acknowledgements

Supported and realized in the framework of the project entitled „Development of a precision, resource efficient harvesting system”. (1924488750) VP3-16.1.1-4.1.5-4.2.1-4.2.2-8.1.1-8.2.1-8.3.1-8.5.1-8.5.2-8.6.1-17

References

Ahmedova, S., 2015. Factors for Increasing the Competitiveness of Small and Medium-Sized Enterprises (SMEs) in Bulgaria. *Procedia - Social and Behavioral Sciences* 195, pp. 1104-1112. doi: 10.1016/j.sbspro.2015.06.155

Ali, J., Kumar, S., 2011. Information and Communication technologies (ICTs) and farmers' decision-making across the agricultural supply chain. *International Journal of Information Man*agement 31, pp. 149-159. doi: 10.1016/j.ijinfomgt.2010.07.008

Annosi, M.C., Brunetta, F., Monti, A., Nati, F., 2019. Is the trend your friend? An analysis of technology 4.0 investment decisions in agricultural SMEs. *Computers in industry* 109, pp. 59-71. doi: 10.1016/j.compind.2019.04.003

Bondina, N., Bondin, I., Pavlova, I., 2021. Cost management models in agricultural business. *IOP Conference Series: Earth and Environmental Science*. 650, article 012022. https://doi.org/10.1088/1755-1315/650/1/012022.

Bouwman, H., Nikou, S., de Reuver, M., 2019. Digitalization, business models, and SMEs: How do business model innovation practices improve performance of digitalizing SMEs? *Telecommunications Policy* 43, 101828. doi: 10.1016/j.telpol.2019.101828

Di Tria, F., Lefons, E., & Tangorra, F., 2017. Cost-benefit analysis of data warehouse design methodologies. *Information Systems* 63, pp. 47–62.

Fan ZP, Liu Y., 2010. An approach to solve group-decision-making problems with ordinal interval numbers. *IEEE Trans Syst Man Cybern Part B (Cybernetics)* 40(5), pp. 1413–1423.

Ganesan K, Veeramani P., 2005. On arithmetic operations of interval numbers. *Int J Uncertain Fuzziness Knowl-Based Syst* 13, pp. 619–631.

Ishibuchi H, Tanaka H., 1990. Multiobjective programming in optimization of the interval objective function. *Eur J Oper Res* 48(2), pp. 219–225.

Klerkx, L., Jakku, E., Labarthe, P., 2019. A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. *NJAS - Wageningen Journal of Life Sciences* 90-91, 100315 doi: 10.1016/j.njas.2019.100315

Látečková, A., Bolek, V. and Szabo, L., 2018. “Information Systems in Agricultural Enterprises: An Empirical Study in Slovak Republic”, *Agris on-line Papers in Economics and Informatics*, Vol. 10, No. 2, pp. 49-60. DOI: 10.7160/aol.2018.100205.

Lezoche, M., Hernandez, J.E., Alemany-Díaz, M.d.M.E., Panetto, H., Kacprzyk, J., 2020 Agri-food 4.0: A suvery of the supply chains and technologies for the future agriculture. *Computers in Industry* 117, 103187. doi: 10.1016/j.compind.2020.103187

Luo, B., Ye, Y., Yao, N. et al., 2021. Interval number ranking method based on multiple decision attitudes and its application in decision making. *Soft Comput* 25, pp. 4091–4101. doi: 10.1007/s00500-020-05434-1.

Moore R, Lodwick W., 2003. Interval analysis and fuzzy set theory. *Fuzzy Sets Syst* 135(1), pp. 5–9.

Morris, W., Henley, A., Dowell, D., 2017. Farm diversification, entrepreneurship and technology adoption: Analysis of upland farmers in Wales. *Journal of Rural Studies*. 53, pp. 132-143. ISSN 0743-0167. https://doi.org/10.1016/j.jrurstud.2017.05.014.

Nakahara Y, Sasaki M, Gen M., 1992. On the linear programming problems with interval coefficients. *Int J Comput Ind Eng* 23, pp. 301–304.

Palmer, M., 2021. Data is the New Oil. *ANA Marketing Maestros*. Retrieved 16 May 2021.

Panetto, H., Lezoche, M., Hormazabal, J.E.H., Alemny-Díaz, M.d.M.E., Kacprzyk, J., 2020. Special issue on Agri-Food 4.0 and digitalization in agriculture supply chains - New directions, challenges and applications. *Computers in Industry*. 116, 103188. doi: 10.1016/j.compind.2020.103188

Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K. Ziogas, V., Sapounas, A. A., Griepentrog, H. W., 2017. Multi-level automation of farm management information systems. *Computers and Electronics in Agriculture*. 142(B), pp. 504-514. ISSN 0168-1699. https://doi.org/10.1016/j.compag.2017.11.022.

Peña, A., Rovira, M. R., Mendoza, V. J. M. F., 2022. Life cycle cost analysis of tomato production in innovative urban agriculture systems. *Journal of Cleaner Production*. 367, article 133037. ISSN 0959-6526. https://doi.org/10.1016/j.jclepro.2022.133037.

Shankarnarayan, V. K and Ramakrishna, H., 2020. Paradigm change in Indian agricultural practices using Big Data: Challenges and opportunities from field to plate. *Information Processing in Agriculture*.7(3), pp. 355-368. ISSN 2214-3173. https://doi.org/10.1016/j.inpa.2020.01.001.

Sharma, V., Tripathi, A. K., Mittal, H., 2022. Technological revolutions in smart farming: Current trends, challenges & future directions. *Computers and Electronics in Agriculture*. 201, article 107217. ISSN 0168-1699. https://doi.org/10.1016/j.compag.2022.107217.

Singh, R.K., Luthra, S., Mangla, S.K., Uniyal, S., 2019. Applications of information and communication technology for sustainable growth of SMEs in India food industry. *Resources, Conservation & Recycling* 147, pp. 10-18. doi: 10.1016/j.resconrec.2019.04.014

Sopegno, A., Calvo, A., Berruto, A., Busato, P., Bocthis, D., 2016. A web mobile application for agricultural machinery cost analysis. *Computers and Electronics in Agriculture*. 130, pp. 158-168. ISSN 0168-1699. https://doi.org/10.1016/j.compag.2016.08.017.

Sørensen, C.G., Pesonen, L., Bochtis, D.D., Vougioukas, S.G., Suomi P., 2011. Functional requirements for a future farm management information system. *Computers and Electronics in Agriculture*, 76(2), pp. 266-276. ISSN 0168-1699. https://doi.org/10.1016/j.compag.2011.02.005.

Subhashree, S. N., Igathinathane, C., Hendrickson, J., Archer, D. Liebig, M., Halvorson, J., Kronberg, S., Toledo, D., Sedivec, K., Peck, D., 2023. Forage economics calculator web tool: A decision support system for forage management. *Computers and Electronics in Agriculture*. 208, article 107775. ISSN 0168-1699. https://doi.org/10.1016/j.compag.2023.107775.

Szenteleki, K, 2007. *Információs rendszerek.* Debrecen, Hungary: Debreceni Egyetem Agrár- és Műszaki Tudományok Centruma, ISBN 978-963-9732-67-4. (In Hungarian)

Takács, V. L., Bubnó, K., Ráthonyi G. G., Bácsné Bába, É., & Szilágyi, R., 2020. Data Warehouse Hybrid Modeling Methodology. *Data Science Journal*. 19:38. pp. 1–23. doi: 10.5334/dsj-2020-038.

Vasileva, N., Astrakhantseva, I., Polyakova, I., Korkina, N., 2020. Accounting and analytical support of financial analysis of agricultural enterprises. *IOP Conference Series: Earth and Environmental Science*. 421, article 022023. https://doi.org/10.1088/1755-1315/421/2/022023.

Verdouw, C., Wolfert, S., Tekinerdogan, B., 2016. Internet of Things in agriculture. *CAB Reviews* 11, pp. 1-12. doi: 10.1079/PAVSNNR201611035

Xu ZS, Da QL., 2002. The uncertain OWA operator. *Int J Intell Syst* 17(6), pp. 569–575.

Yunis, M., Abbas, T., Kassar, A. (2017) The role of ICT and innovation in enhancing organizational performance: The catalysing effect of corporate entrepreneurship. Journal of Business Research 88, pp. 344-356. doi: 10.1016/j.jbusres.2017.12.030