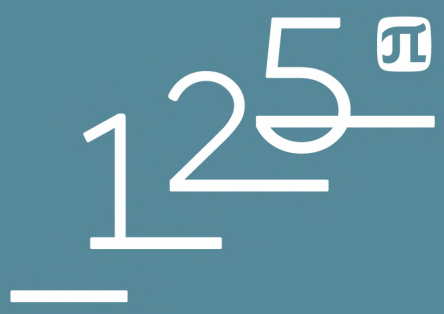
Saint Petersburg 2023

volume 2 issue 4



An international journal

TECHNO ECONOMICS

TECHNO ECONOMICS

An international journal

EDITORS OFFICE PUBLISHER
Executive Secretary: *Olga Voronova*
Development Manager: *Anastasia Levina*
Layout designer: *Dayana Gugutishvili*

PUBLISHER
Peter the Great St. Petersburg Polytechnic University
Corresponding address: 29 Polytechnicheskaya st.,
Saint-Petersburg, 195251, Russia

CONTACTS
Email: *technoeconomics @spbstu.ru*
Web: *https://technoeconomics.spbstu.ru/en*

Saint Petersburg
2023

CONTENTS

Technoeconomics. 2023, Vol. 2, No. 4 (7). Pp. 4–15.
Техноэкономика. 2023, Том 2, № 4 (7). С. 4–15.

Scientific article
UDC 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.1

STUDY OF THE CONCEPT OF BAYESIAN OPTIMIZATION AND PRACTICAL USE OF ITS ALGORITHMS IN THE PYTHON PROGRAMMING LANGUAGE

Anton Chernyagin ✉, **Sergey Svetunkov**

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ a.s.chernyagin@gmail.com

Abstract. The aim of the study is to explore the principles of Bayesian optimization and its potential for solving complex problems, including economic ones. This article presents the main aspects of Bayesian optimization such as selection of a priori distribution, estimation of posterior distribution and selection of optimal model parameters. An example of applying Bayesian optimization to find hyperparameters using the Python programming language is presented. Bayesian optimization algorithms and their application to improve machine learning models were studied. The use of Bayesian optimization algorithm for finding hyperparameters can be useful in the future for optimizing various machine learning models such as neural networks, SVM and others.

Keywords: Bayesian optimization, hyperparameters, Python, machine learning

Citation: Chernyagin A., Svetunkov S. Study of the concept of Bayesian optimization and practical use of its algorithms in the Python programming language. Technoeconomics. 2023. 2. 4 (7). 4−15. DOI: https://doi.org/10.57809/2023.2.4.7.1

Научная статья
УДК 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.1

ИССЛЕДОВАНИЕ ПОНЯТИЯ “БАЙЕСОВСКАЯ ОПТИМИЗАЦИЯ” И ПРАКТИЧЕСКОЕ ИСПОЛЬЗОВАНИЕ ЕЕ АЛГОРИТМОВ НА ЯЗЫКЕ ПРОГРАММИРОВАНИЯ PYTHON

Антон Чернягин ✉, **Сергей Светуньков**

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Россия

✉ a.s.chernyagin@gmail.com

Аннотация. Цель исследования заключается в изучении принципов байесовского подхода в различных областях, а также подробное изучение байесовской оптимизации и ее возможностей для решения сложных задач, в том числе, экономических. В данной работе представлены основные аспекты байесовской оптимизации, такие как выбор априорного распределения, оценка апостериорного распределения и выбор оптимальных параметров модели. Приведен пример применения байесовской оптимизации для нахождения гиперпараметров с помощью языка программирования Python. Были изучены алгоритмы байесовской оптимизации и их применение для улучшения моделей машинного обучения. Использование алгоритма байесовской оптимизации для нахождения гиперпараметров может быть полезным в будущем для оптимизации различных моделей машинного обучения, таких как нейронные сети, SVM и другие.

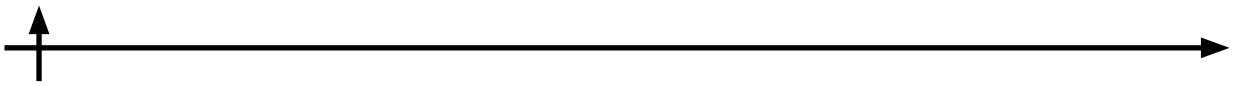
Ключевые слова: Байесовские методы, Байесовская оптимизация, гиперпараметры, Python, машинное обучение

Для цитирования: Чернягин А., Светуньков С. Исследование понятия «Байесовская оптимизация» и практическое использование ее алгоритмов на языке программирования Python // Техноэкономика. 2023. Т. 2, № 4 (7). С. 4–15. DOI: https://doi.org/10.57809/2023.2.4.7.1

Introduction

Bayesian optimization is a current and important area in the field of machine learning and statistics. It is based on the application of Bayes formula to determine the optimal model parameters given a priori knowledge and observed data. This scientific article studies the principles of Bayesian optimization, its application in various machine learning problems and the potential for improving existing methods.

The Bayesian approach allows us to account for uncertainty and a priori knowledge about model parameters, which makes it particularly useful for analyzing small data and solving complex problems. At the same time, Bayesian optimization has a theoretical underpinning and allows us to give the model parameters a meaningful interpretation. The study of Bayesian optimization is necessary and meaningful because this method can solve complex function optimization problems, considering the noise in the data and the cost of estimating the function. The study of Bayesian optimization can lead to the development of new methods and algorithms that can be applied to solve practical problems in various fields. In addition, Bayesian optimization can be used to select reasonable information that determines the whole modeling process in econometrics. Thus, the study of Bayesian optimization has great significance for various fields of science and practice.



Bayesian methods of decision making in economics consider the application of Bayesian optimization for decision making in the activities of individual economic entities. Any organization operates in the economy under conditions of uncertainty, which requires an increase in the accuracy of estimates when making economic decisions, since the financial results of the organization depend on it. Application of Bayesian optimization allows to increase the probability of making rational economic decisions.

Materials and Methods

Bayesian methods are statistical methods based on Bayes' theorem, which allows updating probabilistic estimates of events based on new data. These methods have been widely used in various fields including economics, medicine, genetics, speech recognition, space exploration, insurance, and others. They can be useful for parameter estimation, data prediction, model comparison, decision making under uncertainty, and many other tasks.

Bayesian theory and methods are named after Thomas Bayes (1702-1761), an English mathematician and clergyman who was the first to propose the use of Bayes' theorem to adjust beliefs based on updated data. His work An Essay towards solving a Problem in the Doctrine of Chances was published in 1763, two years after the author's death. However, methods using Bayes' theorem became widespread only towards the end of the 20th century, when computationally intensive calculations became possible with the development of information technology.

The principle of Bayesian methods is to use a priori knowledge of the model parameters to obtain posterior distributions of the parameters after taking into account new data. This allows uncertainty and prior experience to be taken into account when making decisions. Bayesian methods also allow models to be updated based on new data, making them flexible and adaptive.

Bayesian methods have found applications in medical diagnosis, image modeling, genetics, speech recognition, economics, space exploration, insurance, and other fields. They are used to estimate parameters, predict data, compare models, make decisions under uncertainty, and establish causal relationships.

Bayesian optimization in economics can be applied in different contexts to solve a variety of problems:

– Portfolio optimization:

Bayesian optimization can be used to find the optimal allocation of assets in a portfolio in order to maximize returns or minimize risk. The model can take into account various factors such as expected returns, volatility, and correlations between different assets (Laumanns, 2002).

– Pricing:

In business, Bayesian optimization can help in determining the optimal price of a product or service. The model can take into account data on market trends, consumer preferences, and competitive factors.

– Marketing campaigns:

When planning marketing campaigns, Bayesian optimization can be used to determine the optimal budget, timing, and distribution channels for advertising. This can help maximize the expected effect of an advertising campaign.

– Inventory management:

For businesses that sell goods, Bayesian optimization can be useful in inventory management. The model can help optimize inventory levels with respect to demand, delivery time and storage costs.

– Business Process Optimization:

Bayesian optimization can be used to optimize business processes such as manufacturing, logistics or human resource management. The model can suggest optimal parameters to im-

prove efficiency and reduce costs.

– Financial planning:

In financial planning, Bayesian optimization can help determine the optimal budget allocation between different projects or business lines.

– Risk Analysis:

A Bayesian optimization model can be used to analyze risks and select optimal risk management strategies under uncertainty.

However, it is important to note that successful application of Bayesian optimization in economics requires a good understanding of the context of the problem, proper choice of model parameters and careful interpretation of the results. Bayesian methods are a powerful tool to incorporate uncertainty and prior experience into decision making, making them an important tool in various fields of knowledge, including economics.

Bayesian optimization is a method that combines probabilistic models with optimization techniques to efficiently find optimal hyperparameters. Hyperparameters are parameters that are used to control the learning process, as opposed to model parameters that are tuned during training. Bayesian optimization allows us to select the next point to be estimated using the model of the model performance evaluation function. Bayesian optimization can be used to optimize hyperparameters in a variety of domains including machine learning, deep learning, natural language processing, and others. It can be particularly useful in tasks where model performance function estimation is expensive, such as in tasks with large amounts of data or complex models (Fujimoto, 2023).

The advantages of Bayesian optimization include efficiency and the ability to work with a black box, that is, a function that has no analytical expression. It can also be used to optimize multiple hyperparameters simultaneously. Disadvantages include the need to select an appropriate model of the model performance evaluation function and computational complexity (Downey, 2018).

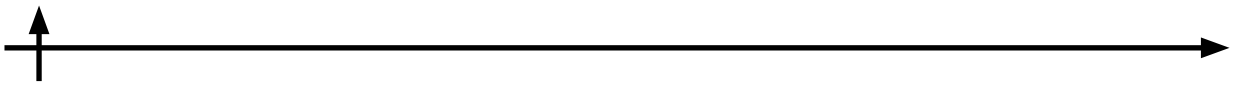
Examples of problems in which Bayesian optimization can be useful include optimization of hyperparameters of neural networks, selection of optimal parameters for machine learning algorithms, optimization of parameters in optimization problems, and others. Bayesian optimization is a powerful tool for optimizing hyperparameters in various domains. It can efficiently find the optimal values of hyperparameters using probabilistic models and optimization methods. However, the selection of a suitable model of the model performance evaluation function and computational complexity can be problems that need to be considered when using this method (Smirnova, 2022).

Bayesian optimization uses Gaussian processes to model the unknown function to be optimized. A Gaussian process is a probability distribution over functions that is updated based on new data. Bayesian optimization adaptively selects the next point to evaluate the function, which reduces the number of evaluation operations. Bayes formula is the basis of Bayesian statistics and is used to calculate the posterior probability based on a priori probability and new data. The Bayes formula is as follows:

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)} \tag{1}$$

where P(A|B) is the posterior probability, P(B|A) is the likelihood (probability of occurrence of event B given event A), P(A) is the a priori probability (probability of occurrence of event A), and P(B) is the full probability (probability of occurrence of event B).

In Bayesian optimization, Bayes formula is used to calculate the posterior probability of the distribution of functions based on observed data. The posterior distribution of functions is the



basis for selecting the next point for function evaluation. The choice of the next point is based on which point maximizes the expected improvement of the function. The expected improvement of the function is calculated based on the posterior distribution of functions and the a priori distribution of hyperparameters (Cuesta Ramirez, 2022).

A Gaussian process is a stochastic process such that every finite collection of random variables has a multivariate normal distribution, that is, every finite linear combination of these random variables has a normal distribution.

Gaussian processes can be used to model an unknown function and can be used in Bayesian optimization to build a model of the model performance evaluation function of the model. The basic properties of Gaussian processes can be defined through the covariance function. Some of these properties include stationarity, isotropy, smoothness and periodicity of the process. If the process is stationary, then the covariance function depends only on the difference between two points (Galuzzi, 2020).

Advantages of Gaussian processes include the ability to work with a black box, i.e., a function that has no analytical expression, and the ability to be used in Bayesian optimization. Disadvantages include the need to select an appropriate model of the function to evaluate model performance and computational complexity. Examples of tasks in which Gaussian processes may be useful include network traffic modeling, statistical modeling, parameter optimization in optimization problems, and others (Pico-Valencia, 2021).

Gaussian processes are a powerful tool for modeling unknown function and optimization in various domains. Gaussian process and Bayesian optimization are closely related. Bayesian optimization uses Gaussian processes to model the unknown function to be optimized. A Gaussian process is used to model the unknown function to be optimized and is a probabilistic model that describes the distribution of function values at different points. Bayesian optimization uses Gaussian processes to estimate the unknown function and select the next point to be estimated. The Gaussian process allows for uncertainty in the data and adaptively selects the next point to estimate the function, thus reducing the number of estimation operations.

The Bayesian optimization algorithm consists of two main parts:

1. Probabilistic function model: Bayesian optimization starts with an a priori distribution over the function to be optimized, which reflects the uncertainty about the function under study. With each new observation of the function, the a priori distribution is updated and a posterior distribution over the possible functions is obtained.

2. Selecting the next point to estimate: based on the posterior distribution, the next point to estimate the function is selected.

To use Bayesian optimization, the following steps should be followed:

1. Define the function to be optimized;

2. Select an a priori distribution for the function;

3. Evaluate the function at the initial points;

4. Update the a priori distribution with new observations and obtain the posterior distribution;

5. Select the next point to estimate based on the posterior distribution;

6. Repeat steps 4-5 until a stopping criterion (e.g., a given number of iterations or convergence) is reached.

Thus the Bayesian optimization algorithm solves the following problems (Feliot, 2017):

1. Optimization of complex functions: Bayesian optimization allows to find the maximum of functions with unknown structure, for example, when selecting hyperparameters for machine learning models;

2. Accounting for the cost of function estimation: in some cases, function estimation can be expensive (e.g., training a neural network). Bayesian optimization adaptively selects the next

point to be estimated given the information from previous iterations, thus reducing the number of estimation operations;

3. Noise control: the function may return different values for the same set of parameters due to noise in the data. Bayesian optimization accounts for this noise and allows finding optimal parameters given this uncertainty.

4. Balance between exploration and exploitation: the Bayesian optimization algorithm takes into account both the already known values of the function and the uncertainty in the unexplored regions of the parameter space, which allows more efficient exploration of the parameter space and finding optimal values (Garrido-Merchбn, 2020).

Bayesian optimization can be relevant for the following economic problems:

1. Decision Making by Individual Economic Entities: Bayesian Methods of Decision Making in Economics examines the application of Bayesian methods of decision making to the activities of individual economic actors

2. Monetary policy analysis: Bayesian vector autoregression model can be used to estimate the impact of various factors on the economy such as monetary policy, external shocks and other variables, which allows us to obtain robust estimates for models with a large number of variables on samples of limited size (Sheikh, 2022).

3. Estimating the impact of factors on the economy: a Bayesian approach can be used to estimate the interdependence of household income inequality and economic growth rates.

Thus, Bayesian optimization can be useful for decision making by individual economic actors, analyzing monetary policy and assessing the impact of various factors on the economy (Pozhidaeva, 2023).

What is more, Bayesian methods can help in decision-making when there is uncertainty in the data or when it is necessary to take into account previous knowledge and experience. For example, Bayesian methods can be used to predict future trends in the market, to determine the optimal price of a product or service, and to assess risk and make decisions in investment activities.

Results and Discussion

Bayesian optimization in the Python programming language in the Microsoft Visual Studio Code environment.

One type of problem often faced by scientists in both academia and industry is the optimization of black-box functions that are expensive to evaluate.

Black box functionality is a term used to refer to a system whose internal structure and mechanism of operation are very complex, unknown or unimportant within the scope of a given task. In the context of software testing, black box means that the tester does not necessarily know the internal structure of the program or system, but tests its functionality based on input and output data. This testing method is used to verify that the software performs all stated functions and customer requirements in full according to the documentation (Bishop, 2006).

In cybernetics, systems engineering, and physics, a black box is a system that is viewed as having some "input" for inputting information and an "output" for displaying the results of operation, with the processes occurring during the operation of the system unknown to the observer. The black box approach developed in the exact sciences in the 1920s and 1940s and was borrowed by other sciences, including behavioral psychology.

Thus, a black box is a system whose internal design and operating mechanism are complex, unknown, or unimportant to the task at hand, and is used in a variety of fields including software testing, aviation, and automatic control theory.

This simplifies the programming and data processing process because the programmer can

use a black-box function without having to know all the details of its implementation. Instead, he can focus on what input data he needs to provide and what results he expects to get (Sharma, 2021).

However, using a black box function can have disadvantages. For example, if the function does not work correctly, it may be difficult for the programmer to determine the cause of the error or to correct it. Also, if the function is not well documented, it may be difficult to understand exactly what input data and what output data it expects.

The concept of a black box expensive to evaluate means that it costs a lot of money or resources to perform a function or operation and its inner workings cannot be understood. A good example of a black box function expensive to evaluate is the hyperparameter optimization of a deep neural network. Each iteration of training can take up to several days, and it is impossible to analyze in advance the values of hyperparameters that will lead to the best performance of the model (Subasi, 2020).

It is possible to perform a cross-grid search of all possible hyperparameter values, but with so many training iterations to be repeated, this would result in a huge computational cost. A more efficient method is needed to find the best set of hyperparameters using the least number of iterations. Bayesian optimization can be used for this task (Gaudrie, 2020).

Bayesian optimization for the black box function in this case has 2 components (Pandita, 2020):

1. The black box function to be optimized is: f(x). We need to find a value of 'x' that globally optimizes f(x). This is a probabilistic model of the function, it is also sometimes called objective function, objective function or loss function. In the general case, we only have knowledge of the inputs and outputs of f(x) (Morice-Atkinson, 2018).

2. An acquisition function: a(x), which is used to generate new values of x to be evaluated with f(x). a(x) internally relies on a Gaussian process model N(X, y) to generate new values of x.

The optimization process itself is as follows:

1. Definition of black box function f(x), acquisition function a(x) and search space of parameter 'x'.

2. Generating several initial values of 'x' randomly and measuring the corresponding results of f(x).

3. Setting up a Gaussian process model N(X, y) on X = x and y = f(x).

4. The acquisition function a(x) then uses N(X, y) to generate new values of 'x' as follows. The model N(X, y) predicts changes in f(x) as a function of 'x'. The value of 'x' that results in the largest predicted value in N(X, y) is then offered as the next sample 'x' to estimate f(x).

5. Steps 3 and 4 should be repeated until the value of 'x' that leads to the global optimum of f(x) is obtained. At the same time, all historical values of 'x' and f(x) should be used to train the Gaussian process model N(X, y) in the next iteration - as the number of data points increases, N(X, y) becomes better at predicting the optimum of f(x).

This experiment uses the bayes_opt library to find the hyperparameter C of the SVC model trained on sklearn breast cancer data.

Support Vector Machine (SVM) is a powerful machine learning algorithm used for classification and regression tasks. Within the classification task, SVM is based on the idea of finding an optimal separating hyperplane in the feature space that maximizes the gap between classes of data.

Support Vector Classifier (SVC) model is one of the variations of SVM for classification task. It works by finding an optimal hyperplane that separates the training data into two classes. The optimal hyperplane is chosen to maximize the distance (gap) between the closest points of each

class, which are called support vectors (Lyu, 2018).

The hyperplane in SVC is defined by a set of weights (weights) and bias (bias), and training the model consists of tuning these parameters based on the training sample. However, it is important to note that in the case of nonlinear data, SVC can use so-called kernel functions, which allow the model to build nonlinear separating hyperplanes in higher dimensional space. The SVC model can also be used to solve the one-vs-all multiclass classification problem. It shows good performance for medium to large training sample sizes, although careful tuning of hyperparameters may be required to achieve optimal results. The overall performance of the SVC method and SVM in general makes it a popular choice for classification tasks in various fields such as computer vision, bioinformatics, financial analytics, and others (Popovic, 2019).

The hyperparameter C in the Support Vector Classification (SVC) model is a regularization parameter that controls the balance between maximizing the width of the separating band and minimizing classification errors. The value of C determines how much we want our model to be adapted to the noise in the data. If the value of C is very small, then the model will be more flexible and will have a larger error on the training data but a smaller error on the test data. If the value of C is very large, the model will be less flexible and will have smaller error on the training data but larger error on the test data. The value of C should be chosen optimally for the particular classification task (Nguyen, 2018).

The components of the optimizer are:

1. The black box function f(x) is the ROC AUC score that we want to maximize to get the best performing model.

2. The acquisition function a(x) is used as the upper confidence bound ("ucb") function, which is of the form: a = mean + kappa * std. Both mean and std are outputs from a Gaussian process model N(X, y). kappa is a hyperparameter of the optimizer that balances exploration and exploitation when searching for x.

The out-of-the-box Python code to perform the above optimization steps is as follows.

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.svm import SVC
from sklearn.datasets import load_breast_cancer
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split
from sklearn.metrics import roc_auc_score
from bayes_opt import BayesianOptimization, UtilityFunction
import warnings
warnings.filterwarnings("ignore")

# Prepare the data (dataset download, identify X and Y).
cancer = load_breast_cancer()
X = cancer.data
y = cancer.target
X_train, X_test, y_train, y_test = train_test_split(X, y, stratify = y, random_state = 42)
# Normalizing data.
scaler = MinMaxScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)
# Define the black box function to optimize.
def black_box_function(C):    # C - SVC hyperparameter for optimization.
    model = SVC(C = C)
    model.fit(X_train_scaled, y_train)
    y_score = model.decision_function(X_test_scaled)
    f = roc_auc_score(y_test, y_score)
    return f
# Set range of C to optimize for.
pbounds = {"C": [0.1, 10]}
# Create a BayesianOptimization optimizer for black box .
optimizer = BayesianOptimization(f = black_box_function, pbounds = pbounds, verbose = 2, random_state = 42)
optimizer.maximize(init_points = 5, n_iter = 10)
print("Best result: {}; f(x) = {}.".format(optimizer.max["params"], optimizer.max["target"]))
```

Fig. 1. Bayesian optimization of the hyperparameter

Running the above Python code produces the following output.

```
|   iter    |  target   |     C     |
-------------------------------------
| 1         | 0.9975    | 3.808     |
| 2         | 0.9979    | 9.512     |
| 3         | 0.9979    | 7.347     |
| 4         | 0.9975    | 6.027     |
| 5         | 0.9966    | 1.645     |
| 6         | 0.9981    | 8.433     |
| 7         | 0.9981    | 8.041     |
| 8         | 0.9981    | 8.237     |
| 9         | 0.9981    | 8.868     |
| 10        | 0.9981    | 8.69      |
| 11        | 0.9981    | 7.901     |
| 12        | 0.9914    | 0.1       |
| 13        | 0.9975    | 2.727     |
| 14        | 0.9975    | 4.913     |
| 15        | 0.9981    | 10.0      |
=====================================
Best result: {'C': 8.432539826625984}; f(x) = 0.9981132075471698.
```

Fig. 2. Result of running the Python code

From the above acquired results, the optimizer was able to determine that using a hyperparameter value of C = 8.432 results in the best model performance.

Conclusion

Thus, how Bayesian optimization works was studied and how it was used to find hyperparameters of a machine learning model. For small datasets or simple models, the speedup in finding hyperparameters may be negligible compared to grid search. However, for very large datasets or deep neural networks, it may not be economically feasible to test every sample on a grid, and the use of Bayesian optimization will improve the efficiency of the hyperparameter search process.

Using Bayesian optimization code to find hyperparameters of black-box functions, we can apply the knowledge gained in the following directions:

1. Machine learning: optimizing the hyperparameters of machine learning models such as neural networks, decision trees and support vector method to improve their performance and accuracy.

2. Financial research: applying Bayesian optimization to tune the parameters of econometric models used in financial analysis and forecasting.

3. Industrial optimization: using Bayesian optimization to determine the optimal parameters of processes and systems in various industries.

4. Tuning classification algorithms: determining the significance of hyperparameters and tuning classifiers with an extensive set of hyperparameters for better validation of results.

5. Comparison and analysis of results: evaluating statistical results using criteria such as the Mann-Whitney criterion to compare the performance of classical and extended Bayesian optimization.

Bayesian optimization of regression model hyperparameters used in programming can be

useful for economic purposes in several aspects:

1. Predicting future trends in the market: Bayesian optimization can help in predicting future trends in the market, which can be useful for production, investment, and strategic planning decisions.

2. Determining the optimal price of a product or service: By optimizing the hyperparameters of a regression model, more accurate forecasts of demand and prices can be obtained, which can help in making pricing decisions.

3. Risk assessment and decision making in investment activities: Bayesian optimization can help in risk assessment and decision making in investment activities by allowing uncertainty and previous experience to be taken into account in decision making.

Thus, Bayesian optimization of regression model hyperparameters applied in programming can be useful for decision making under uncertainty in economic problems such as forecasting, pricing, and investment activities.

REFERENCES

Bishop C.M. 2006. Pattern Recognition and Machine Learning. Springer, 738.

Cuesta Ramirez J. 2022. A comparison of mixed-variables Bayesian optimization approaches. Advanced Modeling and Simulation in Engineering Sciences, 9 (1), 1-29. doi:10.1186/s40323-022-00218-8

Downey A.B. 2018. Bayesian models, 182.

Feliot P.A. 2017. Bayesian approach to constrained single- and multi-objective optimization. Journal of Global Optimization, 67 (1), 97-133. doi:10.1007/s10898-016-0427-3.

Fujimoto Yu. 2023. Controller tuning with Bayesian optimization and its acceleration: Concept and experimental validation. Asian Journal of Control, 25 (3), 2408-2414. doi:10.1002/asjc.2847

Galuzzi B. 2020. Hyperparameter optimization for recommender systems through Bayesian optimization. Computational Management Science. doi:10.1007/s10287-020-00376-3

Garrido-Merchan E.C. 2020. Dealing with categorical and integer-valued variables in Bayesian Optimization with Gaussian processes. Neurocomputing, 380, 20-35. doi:10.1016/j.neucom.2019.11.004

Gaudrie D. 2020. Targeting solutions in Bayesian multi-objective optimization: sequential and batch versions. Annals of Mathematics and Artificial Intelligence, 88 (1), 187-212. doi:10.1007/s10472-019-09644-8

Laumanns M. 2002. Bayesian Optimization Algorithms for Multi-objective Optimization. Lecture Notes in Computer Science, 2439, 298-307.

Lyu W., P. Xue. 2018. An efficient Bayesian optimization approach for automated optimization of analog circuits. IEEE Transactions on Circuits and Systems. Part 1: Regular Papers, 65 (6), 1954-1967. doi:10.1109/TCSI.2017.2768826

Morice-Atkinson X. 2018. Learning from the machine: Interpreting machine learning algorithms for point- and extended-source classification. Monthly Notices of the Royal Astronomical Society, 481 (3), 4194-4205. doi:10.1093/mnras/sty2575

Nguyen T.D. 2018. Stable Bayesian optimization. International Journal of Data Science and Analytics, 6 (4), 327-339. doi:10.1007/s41060-018-0119-9

Pandita P. 2021. Computationally efficient bayesian optimization for multi-objective industrial applications. AIAA Scitech 2021 Forum, 1, 1-13. doi:10.2514/6.2021-1482

Pico-Valencia P. 2021. Bringing Machine Learning Predictive Models Based on Machine Learning Closer to Non-technical Users. Advances in Intelligent Systems and Computing, 1273, 3-15. doi:10.1007/978-3-030-59194-6_1

Popovic M. 2019. Formal verification of local and distributed python software transactional memories. Revue Roumaine des Sciences Techniques. Series Electrotechnique et Energetique, 64 (4), 423-428.

Pozhidaeva. N.A. 2023. Application of Bayesian process definition strategy for priority optimization (on the example of the banking sector). Economics: yesterday, today, tomorrow, 13 (3), 530-535. doi:10.34670/AR.2023.83.23.085

Sharma P. 2021. Role of machine learning and deep learning in securing 5G-driven industrial IoT applications. Ad Hoc Networks, 123, 102685. doi:10.1016/j.adhoc.2021.102685

Sheikh H.M. 2022. Bayesian optimization for mixed-variable, multi-objective problems. Structural and Multidisciplinary Optimization, 65 (11), 1-14. doi:10.1007/s00158-022-03382-y

Smirnova V.S., Shalamov V.V., Efimova V.A., Filchenkov A.A. 2020. Hyperparameter optimization based on combining a priori and a posteriori knowledge of the classification task. Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics, 20 (6), 828-834.

Subasi A. 2020. Practical Machine Learning for Data Analysis Using Python. Practical Machine Learning for Data Analysis Using Python, 1-520. doi: 10.1016/B978-0-12-821379-7.00008-4

СПИСОК ИСТОЧНИКОВ

Bishop C.M. 2006. Pattern Recognition and Machine Learning. Springer, 738.

Cuesta Ramirez J. 2022. A comparison of mixed-variables Bayesian optimization approaches. Advanced Modeling and Simulation in Engineering Sciences, 9 (1), 1-29. doi:10.1186/s40323-022-00218-8

Downey A.B. 2018. Bayesian models, 182.

Feliot P.A. 2017. Bayesian approach to constrained single- and multi-objective optimization. Journal of Global Optimization, 67 (1), 97-133. doi:10.1007/s10898-016-0427-3.

Fujimoto Yu. 2023. Controller tuning with Bayesian optimization and its acceleration: Concept and experimental validation. Asian Journal of Control, 25 (3), 2408-2414. doi:10.1002/asjc.2847

Galuzzi B. 2020. Hyperparameter optimization for recommender systems through Bayesian optimization. Computational Management Science. doi:10.1007/s10287-020-00376-3

Garrido-Merchan E.C. 2020. Dealing with categorical and integer-valued variables in Bayesian Optimization with Gaussian processes. Neurocomputing, 380, 20-35. doi:10.1016/j.neucom.2019.11.004

Gaudrie D. 2020. Targeting solutions in Bayesian multi-objective optimization: sequential and batch versions. Annals of Mathematics and Artificial Intelligence, 88 (1), 187-212. doi:10.1007/s10472-019-09644-8

Laumanns M. 2002. Bayesian Optimization Algorithms for Multi-objective Optimization. Lecture Notes in Computer Science, 2439, 298-307.

Lyu W., P. Xue. 2018. An efficient Bayesian optimization approach for automated optimization of analog circuits. IEEE Transactions on Circuits and Systems. Part 1: Regular Papers, 65 (6), 1954-1967. doi:10.1109/TCSI.2017.2768826

Morice-Atkinson X. 2018. Learning from the machine: Interpreting machine learning algorithms for point- and extended-source classification. Monthly Notices of the Royal Astronomical Society, 481 (3), 4194-4205. doi:10.1093/mnras/sty2575

Nguyen T.D. 2018. Stable Bayesian optimization. International Journal of Data Science and Analytics, 6 (4), 327-339. doi:10.1007/s41060-018-0119-9

Pandita P. 2021. Computationally efficient bayesian optimization for multi-objective industrial applications. AIAA Scitech 2021 Forum, 1, 1-13. doi:10.2514/6.2021-1482

Pico-Valencia P. 2021. Bringing Machine Learning Predictive Models Based on Machine Learning Closer to Non-technical Users. Advances in Intelligent Systems and Computing, 1273, 3-15. doi:10.1007/978-3-030-59194-6_1

Popovic M. 2019. Formal verification of local and distributed python software transactional memories. Revue Roumaine des Sciences Techniques. Series Electrotechnique et Energetique, 64 (4), 423-428.

Пожидаева Н.А. 2023. Применение байесовской стратегии определения процесса для первоочередной оптимизации (на примере банковского сектора). Экономика: вчера, сегодня, завтра, 13 (3), 530-535, doi: 10.34670/AR.2023.83.23.085

Sharma P. 2021. Role of machine learning and deep learning in securing 5G-driven industrial IoT applications. Ad Hoc Networks, 123, 102685. doi:10.1016/j.adhoc.2021.102685

Sheikh H.M. 2022. Bayesian optimization for mixed-variable, multi-objective problems. Structural and Multidisciplinary Optimization, 65 (11), 1-14. doi:10.1007/s00158-022-03382-y

Smirnova V.S., Shalamov V.V., Efimova V.A., Filchenkov A.A. 2020. Hyperparameter optimization based on combining a priori and a posteriori knowledge of the classification task. Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics, 20 (6), 828-834.

Subasi A. 2020. Practical Machine Learning for Data Analysis Using Python. Practical Machine Learning for Data Analysis Using Python, 1-520. doi: 10.1016/B978-0-12-821379-7.00008-4

INFORMATION ABOUT AUTHORS / ИНФОРМАЦИЯ ОБ АВТОРАХ

CHERNYAGIN Anton S. – student.
E-mail: lyamin_bm@spbstu.ru
ЧЕРНЯГИН Антон Сергеевич – студент.
E-mail: a.s.chernyagin@gmail.com

SVETUNKOV Sergey G. – Professor, Doctor of Economic Sciences
E-mail: sergey@svetunkov.ru
СВЕТУНЬКОВ Сергей Геннадьевич – профессор, д.э.н.
E-mail: sergey@svetunkov.ru
ORCID: https://orcid.org/0000-0001-6251-7644

Статья поступила в редакцию 29.11.2023; одобрена после рецензирования 18.12.2023; принята к публикации 18.12.2023.

The article was submitted 29.11.2023; approved after reviewing 18.12.2023; accepted for publication 18.12.2023.

Technoeconomics. 2023, Vol. 2, No. 4 (7). Pp. 16–24.
Техноэкономика. 2023, Том 2, № 4 (7). С. 16–24.

Scientific article
UDC 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.2

APPLICATION OF THE SIMULATION MODELING METHOD FOR SOLVING CONTENT MARKETING AUTOMATION TASKS

George Gimadeev ✉

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ medmet@yandex.ru

Abstract. The paper describes the creation of a simulation model of the activity of a content manager to increase the effectiveness of marketing, and also defines the relationship of content marketing with Internet marketing in general. Simulation modeling in the AnyLogic program was used as the main research method. The literature sources on the topic of content analysis were analyzed, the processes in the work of a content manager that are subject to automation were highlighted. A simulation model of automation of the content manager's activity in the AnyLogic program was also developed and the economic efficiency of the implemented measures was justified. The scientific novelty of the proposed method lies in the fact that the model takes into account the need to create reports on promotions, as well as the need to conduct SEO promotion and evaluate the effectiveness of advertising on different Internet platforms: yandex. direct and google.awards. The results of the research can be used by IT specialists to create and improve software products.

Keywords: simulation modeling, content analysis, Internet marketing, marketing automation, AnyLogic

Citation: Gimadeev G. Application of the simulation modeling method for solving content marketing automation tasks. Technoeconomics. 2023. 2. 4 (7). 16–24. DOI: https://doi.org/10.57809/2023.2.4.7.2

Научная статья
УДК 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.2

ПРИМЕНЕНИЕ МЕТОДА ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧ АВТОМАТИЗАЦИИ КОНТЕНТ-МАРКЕТИНГА

Георгий Гимадеев ✉

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Россия

✉ medmet@yandex.ru

Аннотация. В работе описано создание имитационной модели деятельности контент-менеджера для повышения эффективности маркетинга, а также определена связь контент-маркетинга с интернет-маркетингом в целом. В качестве основного метода исследования применялось имитационное моделирование в программе AnyLogic. Были проанализированы литературные источники по теме контент-анализа, выделены процессы в работе контент-менеджера, подлежащие автоматизации. Также была разработана имитационная модель автоматизации деятельности контент-менеджера в программе AnyLogic и обоснована экономическая эффективность внедряемых мер. Научная новизна предлагаемого метода заключается в том, что модель учитывает необходимость создания отчетов по рекламным акциям, а также необходимость ведения seo-продвижения и оценки эффективности рекламы на разных интернет-платформах: яндекс.директ и google. awards. Результаты исследования могут быть использованы it-специалистами для создания и усовершенствования программных продуктов.

Ключевые слова: имитационное моделирование, контент-анализ, интернет-маркетинг, автоматизация маркетинга, AnyLogic

Для цитирования: Гимадеев Г. Применение метода имитационного моделирования для решения задач автоматизации контент-маркетинга // Техноэкономика. 2023. Т. 2, № 4 (7). С. 16–24. DOI: https://doi.org/10.57809/2023.2.4.7.2

Introduction

Content marketing refers to marketing techniques that allow you to create, distribute and analyze information that allows you to sell products by increasing customer trust and involving them in the company's information field. The objectives of content marketing include: increasing the sales of goods, improving brand awareness of the company, informing customers about the company's services or products.

In this study, we will focus on such a part of content marketing as social media marketing (hereinafter referred to as SMM - social media marketing), however, it should be noted that this type of marketing is closely intertwined with other types of digital marketing and, above all, Internet marketing in in general (Chaffey, 2013). The line between them is very conditional. When studying content in SMM, we study content as a whole, since the same content is often used in other areas of marketing (Ahearne, 2007).

We will also be primarily interested in automating the work of a content manager, whose functions can also be attributed to the work of an Internet marketer. The work functions of a content manager and an internet marketer often overlap. For example, a content manager is often responsible for SEO promotion of content on social networks in the Yandex and Google

search engines.

The following figure shows the interaction between content marketing and social media marketing.

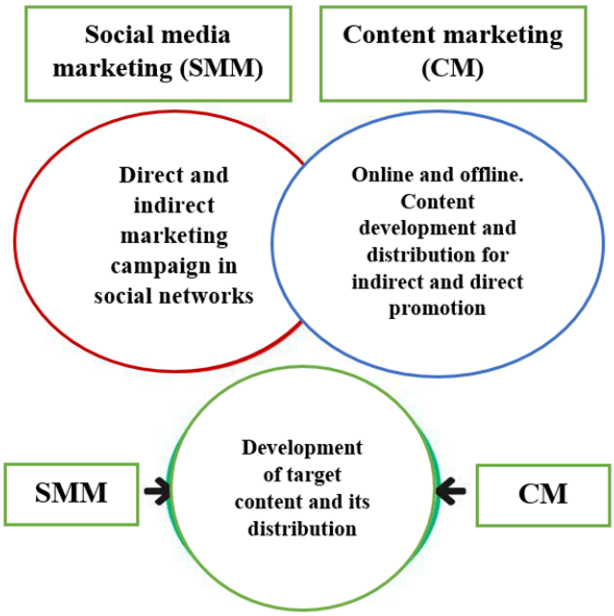


Fig. 1. Interconnection between SMM and content marketing

The distribution of content relevant to the target audience is inseparable from social networks, as a medium that is becoming increasingly popular and can compete in audience reach with television. Thus, the choice of SMM marketing for subsequent research is due to the high demand for content analysis in this area (Biemans, 2010).

Among the SMM marketing tools we can list:

1. Blogging (creating and maintaining various blogs);
2. Maintaining a thematic community;
3. Targeted advertising on social networks;
4. Direct marketing – communication with clients on behalf of the company on forums, communities, comments under posts and other possible places;
5. Monitoring mentions of the company on social networks to deal with negativity.

And also many other tools. What these tools have in common is their relatively low cost with minimal costs and long periods of time before obtaining a result, which at the same time brings a positive and long-term effect (Donath, 1995).

An article by Finnish researchers Joel Jarvinen and Heini Taiminen, "Using Automation for Content Marketing in the B2B Market," states that in large online agencies, a marketer can spend up to 17 working hours per week writing various types of reporting (Bhattacharyya, 2010).

Materials and Methods

Despite the fact that there are systems for automating the work of marketers and content managers, as a rule they do not include work on social networks or such frequently used systems as Yandex Direct or Google Awards at the same time. Whereas a content manager often has to work simultaneously in both Yandex and Google. A problem arises due to the inability to control the activity of advertising campaigns; the system itself runs ads and does not always do this effectively, which is why there is a decrease in conversion (Batt, 2012).

In addition to the need to compile numerous reports, as well as use various services for tracking the effectiveness of advertising and product promotion, content managers complain about the need to constantly monitor promotions on social networks (Buttle, 2009). Timely automated tracking of promotions or reminders about the end of these promotions would also help content managers reduce their time costs.

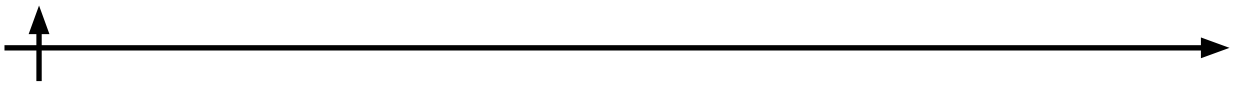
To better understand the relevance of the task of automating the work of a content manager or Internet marketer, a patent search was conducted in the field of content marketing automation. During the patent search, scientific, technical and patent information was studied:

1. Description of inventions for patents
2. Formula of invention
3. Classification
4. Drawings
5. Applications for inventions
6. Abstracts of foreign inventions
7. Commercial potential of the technology
8. Trademarks for the technology in question

About 100 patents were studied; an example of patent search work can be found in Table 1.

Table 1. Patent search results

Patent search				
Country and company owner	Patent number	Patent name	Patent Description	Search site
Google Inc, US	CA2634039A1	SEGMENT CONTENT OPTIMIZATION DELIVERY SYSTEM AND METHOD	The method includes retrieving by a processor the respective use data for the population, from the plurality of marketing systems, determining by the processor if the respective use data exceeds a threshold for particular behavioral pattern of interest, for the respective use data, determining by the processor a unique identifier for each user device of the use data (Christodoulides, 2009)	https:// worldwide. espacenet. com/
SIRIUSDECISIONS INC, US	WO2013173545A1	METHOD AND SYSTEM FOR ASSESSING BUSINESS-TO-BUSINESS SALES AND MARKETING PERFORMANCE DATA	The present invention relates to the assessment of performance data in the field of business- to-business sales and marketing. More specifically, the present invention relates to computer-implemented methods and systems for assessing performance data relating to an organization's ability create systematic, predictable and measurable demand (Montgomery, 2003).	https:// worldwide. espacenet. com/
CATHOLIC UNIV KOREA IND	KR20210071522A	B2B MARKETING INTEGRATED MANAGEMENT SYSTEM	As an example, the task to be solved according to the embodiment of the present invention can easily cope with the corporate environment, easily change and maintain permissions required for security, etc. (Beverland, 2012).	https:// worldwide. espacenet. com/



As a result of the patent search, the following conclusions were made:

1. The general patentability of the proposed method is noted. In general, marketing automation systems are present in the patent search systems studied, but it was not possible to find content marketing automation systems that would pay much attention to reducing the time for compiling reports from various systems, such as Yandex, Google Words and social networks at the same time. Although certain methods for automating the work of Internet marketers and content managers are present in the databases in the form of patents (Adamson, 2012).

2. Taking into account the patent search, as well as the cited Russian-language and foreign sources, we can note the commercial potential of the technology, which can be successfully used for large companies and Internet agencies that run marketing companies, using the entire arsenal of content creation and promotion (Jarvinen, 2016).

3. Trademarks and combinations of words for the technology in question were found with the words "marketing" and "automation", but no trademarks with the combination "marketing automation" were found (Donath, 1999).

In order to reduce the time spent by content managers and ensure the unification of their work in various areas of business, it was decided to use the simulation modeling method. Having great variability in terms of solving the problems facing the researcher, this method allows you to adapt it for various areas of business, as well as to calculate in advance the effectiveness of the implemented measures to automate the activities of marketers. The latter can be used both when creating new software products and when selecting existing ones.

The AnyLogic program was chosen as the main modeling tool. After that, it became possible to solve the main tasks of the study:

1. Select processes to be automated.
2. Determine key indicators and their values for these processes.
3. Build a model for automating the work of Internet marketers.
4. Determine the economic efficiency of the measures being implemented.

Results and Discussion

As part of solving the first task, based on the data provided, we selected the processes that we want to automate:

1. Monitoring SEO optimization services (Yandex Direct, Google Words).
2. Maintaining reports on contextual advertising on the Internet.
3. Tracking the start and end dates of promotions on social networks (Dubois, 2005).

As part of the solution of the second task, the key indicators of these processes were identified, which were collected in a table as variables of the future model, and the interrelations of the model variables were indicated. Table 2 presents the symbols of the elements used (Easton, 2010).

Table 2. Symbols of model elements

Designation	Factor
X1	Wasting time tracking SEO before implementation
X2	Wasting time tracking the CEO after implementation
X3	Reducing time in
X4	Waste of time reporting on context before implementation
X5	Wasting time on post-implementation context reporting
X6	Reducing time in
X7	Wasting time tracking the end of the promotion before implementation
X8	Wasting time tracking the end of the promotion after implementation
X9	Reducing time in
Y1	Time costs of an Internet marketer

Table 2 shows the relationship between the elements. The following are formulas for calculating dynamic variables and accumulators, as well as the values of statistical variables (Grinko, 2019). Since, on average, the implementation of a particular function (writing a report) takes 1 working hour, hours are taken as a unit of time.

Table 3. Relationship between model variables

Endogenous Variables	Exogenous variables		
Y1	X1	X2	X3

The marketer's time costs will be calculated using the formula:

Y1 =X3,X,6,X9. X=100% -100%/z, where z = x1/x2, where x1 is the time that was spent before the introduction of automation, and x2 is the time that was spent after the implementation (Dennis, 2007).

Table 3 presents the values of statistical variables. From the data presented, it can be seen that the largest items of time spent by an Internet marketer are the costs of tracking advertising campaigns in SEO (Yandex.Direct, Google Words) (Cheporova, 2019).

Table 4. Values of statistical variables of the model

Variable	Value	Unit measurements
X1	87.5	Hours
X2	16	Hours
X3	2	Hours
Y1(X3, X6, X9)	95.23 98.44 50	%

The marketer's time costs will be calculated using the formula:

Y1 =X1,X2,X3. X=100% -100%/z, where z = x1/x2, where x1 is the time that was spent before the introduction of automation, and x2 is the time that was spent after the implementation.

An Internet marketer agent and an indicator of overall marketing effectiveness have also been introduced, which directly correlates with a reduction in the time spent by an Internet

marketer. In this work, we use the concepts of Internet marketer and content manager as synonyms (Cooper, 2007; Kotik, 2020; Gimadeev and Abdukhalilova, 2023).

To solve the third research problem, a simulation model was created in the Anylogic software product.

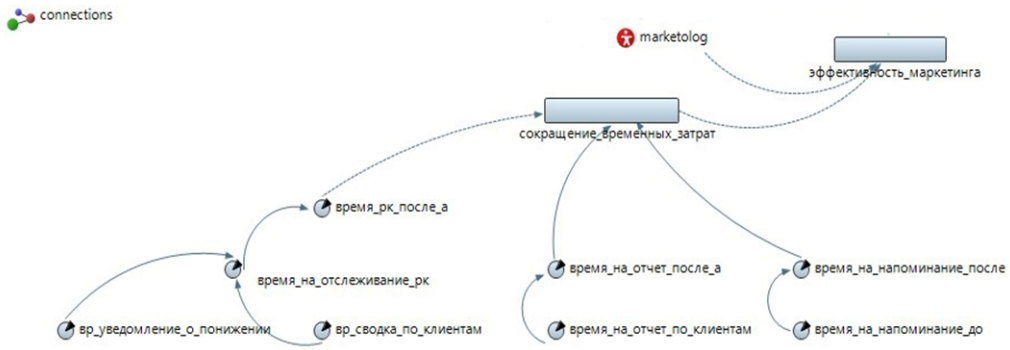


Fig. 2. Simulation model for automating the work of a content manager

Finally, as part of solving the fourth and last task of this work, the efficiency parameters of implementing a model for automating the work of a content manager or Internet marketer were calculated.

Table 5. Efficiency of implementing a content management automation system

Name	Before automation	After automation	%	Reduced/Result
Tracking the effectiveness of management of the advertisement	87.5	4.17	95.23	18 228.44
Making report	16	0.25	98.44	3 445.31
End of promotion	2	1	50	50
Other effects from the introduction of automation				Monitoring the work of Internet marketers; Increased customer loyalty

Conclusion

As part of the research, it became clear that the proposed method of automating the activities of a content manager minimizes the time spent on tracking the effectiveness of advertising campaigns in SEO and contextual advertising, drawing up reports, and also informing specialists about the end of promotions for a specific client.

The objectives of the study were successfully completed. The proposed model, thanks to the flexible software tool AnyLogic, can be supplemented with a large number of parameters and variables that can increase the effectiveness of marketing activities.

REFERENCES

Adamson B., Dixon M. 2012. The end of solution sales. Harvard Business Review, 90(7–8), 60–68.

Ahearne M., Hughes D. 2007. Why sales reps should welcome information technology: Measuring the impact of CRM-based IT on sales effectiveness. International Journal of Research in Marketing, 24(4), 336–349.

Batt P. 2012. Measures and measurement: Process and practise. Industrial Marketing Management, 41(3), 379–384.

Beverland M. 2010. What makes a good case study? A positivist review of qualitative case

research published in Industrial Marketing Management. Industrial Marketing Management, 39(1), 56–63.

Bhattacharyya B. 2014. Improving inventory demand forecasting by using the sales pipeline: A case study. Journal of Business Forecasting, 33(1), 7–11.

Biemans W., Brencic M. 2010. Marketing–sales interface confgurations in B2B firms. Industrial Marketing Management, 39(2), 183–194.

Buttle F. 2009. Customer relationship management: Concepts and technologies (2nd ed.). London: Taylor & Francis.

Chaffey D., Smith P. 2013. Emarketing excellence: Planning and optimizing your digital marketing (4th ed.). London: Taylor & Francis.

Cheporova G. E., Tsurko O. 2019. Introduction of Internet marketing in modern business. Actual problems and prospects of economic development, 272-273.

Christodoulides G. 2009. Branding in the post-internet era. Marketing Theory, 9(1), 144.

Cooper M., Budd. 2007. Tying the pieces together: A normative framework for integrating sales and project operations. Industrial Marketing Management, 36(2), 173–182.

Dennis C., Merrilees B., Jayawardhena C. 2009. E-consumer behaviour.European Journal of Marketing, 43(9/10), 1121–1139.

Donath B. 1999. Quality information leads to quality leads. Marketing News, 33(17), 11.

Donath B., Crocker R., Dixon C. 1995. Managing sales leads: How to turnevery prospect into a customer. Lincolnwood, IL: NTC Business Books.

Dubois A., Gibbert M. 2010. From complexity to transparency: Managing the interplay between theory, method and empirical phenomena in IMM case studies. Industrial Marketing Management, 39(1), 129–136.

Easton G. 2010. Critical realism in case study research. Industrial Marketing Management, 39(1), 118–128.

Gimadeev G., Abdukhalilova L. 2023. From like to sales: to the question of automation of lead generation process in social networks and lead quality assessment. Technoeconomics, 2, 3 (6), 28–43. DOI: https://doi.org/10.57809/2023.2.3.6.3

Grinko O. 2019. The content marketing funnel. Science and innovation, 9, 53—57.

Jarvinen J., Taiminen H. 2016. Harnessing marketing automation for B2B content marketing. Industrial Marketing Management. URL: https://www.researchgate.net/publication/280875581_Harnessing_marketing_automation_for_B2B_content_marketing (accessed: 19.01.2023).

Kotik V. 2020. Using AnyLogic software in simulation modeling. Young scientist, 51 (341), 13-15.

Montgomery A., Srinivasan K. 2003. Learning about customers without asking. The power of one—Leverage value from personalization technologies, 122–143.

Simulation modeling. URL: https://www.anylogic.ru/use-of-simulation (accessed: 19.01.2023).

СПИСОК ИСТОЧНИКОВ

Adamson B., Dixon M. 2012. The end of solution sales. Harvard Business Review, 90(7–8), 60–68.

Ahearne M., Hughes D. 2007. Why sales reps should welcome information technology: Measuring the impact of CRM-based IT on sales effectiveness. International Journal of Research in Marketing, 24(4), 336–349.

Batt P. 2012. Measures and measurement: Process and practise. Industrial Marketing Management, 41(3), 379–384.

Beverland M. 2010. What makes a good case study? A positivist review of qualitative case research published in Industrial Marketing Management. Industrial Marketing Management, 39(1), 56–63.

Bhattacharyya B. 2014. Improving inventory demand forecasting by using the sales pipeline: A case study. Journal of Business Forecasting, 33(1), 7–11.

Biemans W., Brencic M. 2010. Marketing–sales interface configurations in B2B firms. Industrial Marketing Management, 39(2), 183–194.

Buttle F. 2009. Customer relationship management: Concepts and technologies (2nd ed.). London: Taylor & Francis.

Chaffey D., Smith P. 2013. Emarketing excellence: Planning and optimizing your digital marketing (4th ed.). London: Taylor & Francis.

Christodoulides G. 2009. Branding in the post-internet era. Marketing Theory, 9(1), 144.

Cooper M., Budd. 2007. Tying the pieces together: A normative framework for integrating sales and project operations. Industrial Marketing Management, 36(2), 173–182.

Dennis C., Merrilees B., Jayawardhena C. 2009. E-consumer behaviour.European Journal of Marketing, 43(9/10), 1121–1139.

Donath B. 1999. Quality information leads to quality leads. Marketing News, 33(17), 11.

Donath B., Crocker R., Dixon C. 1995. Managing sales leads: How to turnevery prospect into a customer. Lincolnwood, IL: NTC Business Books.

Dubois A., Gibbert M. 2010. From complexity to transparency: Managing the interplay between theory, method and empirical phenomena in IMM case studies. Industrial Marketing Management, 39(1), 129–136.

Easton G. 2010. Critical realism in case study research. Industrial Marketing Management, 39(1), 118–128.

Gimadeev G., Abdukhalilova L. 2023. From like to sales: to the question of automation of lead generation process in social networks and lead quality assessment. Technoeconomics, 2, 3 (6), 28–43. DOI: https://doi.org/10.57809/2023.2.3.6.3

Jarvinen J., Taiminen H. 2016. Harnessing marketing automation for B2B content marketing. Industrial Marketing Management.URL:https://www.researchgate.net/publication/280875581_Harnessing_marketing_automation_for_B2B_content_marketing (accessed: 19.01.2023).

Montgomery A., Srinivasan K. 2003. Learning about customers without asking. The power of one—Leverage value from personalization technologies, 122–143.

Гринько О. 2019. Воронка контент-маркетинга. Наука и инновации, 9, 53—57.

Котик В. К. 2020. Использование программного обеспечения AnyLogic в имитационном моделировании. Молодой ученый, 51 (341), 13-15.

Чепорова Г. Е., Цурко О. Ю. 2019. Внедрение интернет-маркетинга в современный бизнес. Актуальные проблемы и перспективы развития экономики, 272-273.

Имитационное моделирование. URL: https://www.anylogic.ru/use-of-simulation (дата обращения: 19.01.2023).

INFORMATION ABOUT AUTHOR / ИНФОРМАЦИЯ ОБ АВТОРЕ

GIMADEEV George R. – PhD student.
E-mail: medmet@yandex.ru
ГИАМДЕЕВ Георгий Рустэмович – аспирант.
E-mail: medmet@yandex.ru

Статья поступила в редакцию 01.09.2023; одобрена после рецензирования 04.09.2023; принята к публикации 04.09.2023.

The article was submitted 01.09.2023; approved after reviewing 04.09.2023; accepted for publication 04.09.2023.

Technoeconomics. 2023, Vol. 2, No. 4 (7). Pp. 25–37.
Техноэкономика. 2023, Том 2, № 4 (7). С. 25–37.

Scientific article
UDC 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.3

ARCHITECTURE OF THE SOFTWARE ASSET MANAGEMENT SYSTEM

Sergey Blinov ✉

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ st068499@student.spbstu.ru

Abstract. Software Asset Management (SAM) system is a set of measures aimed at optimizing the acquisition, deployment, use and maintenance of software assets in organizations. Software asset management processes require not only a competent management system, but also support from IT services. This paper presents a way to define the SAM application architecture using a client-oriented approach. Terminological analysis was carried out, software management process was modeled, requirements for the application were compiled and the corresponding functionality of its modules was determined.

Keywords: software asset management, license management, software utilization, SAM architecture

Citation: Blinov S. Architecture of the software asset management system. Technoeconomics. 2023. 2. 4 (7). 25–37. DOI: https://doi.org/10.57809/2023.2.4.7.3

Научная статья
УДК 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.3

АРХИТЕКТУРА СИСТЕМЫ УПРАВЛЕНИЯ ПРОГРАММНЫМИ АКТИВАМИ

Сергей Блинов ✉

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Россия

✉ st068499@student.spbstu.ru

Аннотация. Система управления программными активами (SAM) - это комплекс мер, направленных на оптимизацию приобретения, развертывания, использования и сопровождения программных активов в организациях. Процессы управления программными активами требуют не только грамотной системы управления, но и поддержки со стороны ИТ-служб. В данной работе представлен способ определения и построения архитектуры приложения SAM с использованием клиент-ориентированного подхода. В ходе исследования был проведен терминологический анализ, смоделирован процесс управления программным обеспечением, составлены требования к приложению, а также определена соответствующая функциональность его модулей.

Ключевые слова: управление программными активами, управление лицензиями, использование программного обеспечения, архитектура SAM

Для цитирования: Блинов С. Архитектура системы управления программными активами // Техноэкономика. 2023. Т. 2, № 4 (7). С. 25–37. DOI: https://doi.org/10.57809/2023.2.4.7.3

Introduction

Software asset management activities in the Russian Federation are regulated by the GOST R ISO/IEC 19770-1 – 2014 standard, which is identical to ISO/IEC 19770. This standard defines the basis of a comprehensive set of software Asset Management processes (Software Asset Management, SAM), divided into levels providing for the phased implementation, evaluation and approval of SAM processes.

According to the standard, Software Asset Management is effective management, control and protection of software assets across the organization and effective management, control and protection of information about related assets necessary for the management of software assets. Therefore, within a specific enterprise, the functionality of the SAM solution is determined through a set of processes and services that implement management, control and protection. Their list depends on the scale of the organization, the type of activity, the allocated budget, the maturity of related processes, the range of software assets used, the nature of relationships with software vendors, as well as the choice between in-house development of the SAM system and the acquisition of a ready-made solution (Spewak, 1992).

SAM is a component of IT Asset Management (IT Asset Management, ITAM). ITAM is a coordinated activity of an organization to obtain value from IT assets. The following types of IT assets are distinguished:

– software;
– media (physical and digital);
– IT equipment (physical and virtual);

– licenses (including license confirmation);
– contracts;
– ITAM-IT asset management systems (including ITAM systems and tools, as well as metadata needed to manage all IT assets).

The object of the SAM system is primarily licenses, as well as the corresponding software. When expanding the functionality of the system to other IT assets, it is advisable to talk about creating an ITAM system.

Returning to the definition of SAM, it is necessary to focus on why the task of managing, controlling and protecting software assets should be solved precisely through the creation of a separate IT solution. In order to answer this question, it is necessary to reflect the specifics of software assets as an object of management.

1. Dynamic nature of software and IT assets. A software asset, both separately and as a component of an IT system, may be subject to frequent updates, corrections and user modifications. Tracking these changes is necessary for making centralized decisions on a group of software assets, monitoring usage, protecting against violations of contracts with vendors or security threats.

2. Software complexity. The software can be used as a finished product or component. The same software asset can be used at the same enterprise, but with different functionality. Differences arise due to the role model (different access levels and divisions), as well as software versions (see point 1).

3. Licensing and compliance. Software assets have complex licensing agreements that need to be monitored to ensure compliance and prevent legal problems. Their specificity is not only in controlling the number (which can be measured both by the number of devices and users, both simultaneously and during any period, etc.), but also in terms of use (right to change), updates, technical support, security, access to documentation and, of course, available functionality. There are financial and legal consequences for violating the terms of use of the Software (including due to the lack of tools for software management).

The above are only some circumstances that make it difficult to account for software assets by classical asset management tools. Therefore, a separate IT solution is needed that takes into account the specifics of SAM (Kalyatin, 2020).

The purpose is to create the architecture of the SAM application.

To achieve it, it is necessary to solve the following tasks:

1. Outline the list of and define key concepts used in the paper (mainly by analyzing regulatory documentation)

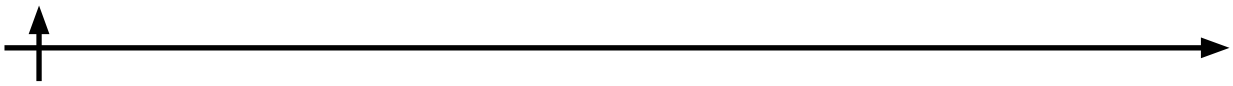
2. Define problems of AS IS situation by modeling and describing software management processes to identify pains of business users

3. Synthesis of functional requirements for the application from the pains of process participants

4. Using methodology of enterprise architecture, create an application model.

Materials and Methods

Before proceeding to the development of solution to the previously identified set of problems, it is necessary to define the key terms that will be needed in framework of the study, including characterization of the proposed SAM tools. The terms and the corresponding definitions are discussed in the section "Literature Review" and allows to form the boundaries of research, separating software assets from IT assets. To ensure the unity of terminology, GOST/ISO standards are mainly used. The functionality of the service system was formed both on the basis of scientific publications and on the basis of existing ready-made solutions in the IT



market.

After a single terminological field has been formed, it is still too early to proceed to the description of the solution. The fact is that the upper-level description of the problem does not have sufficient specifics to describe the functionality designed to ensure effective management of software assets. Therefore, the next step is to detail the problems to the level of groups of business users who are stakeholders in the process. The "pains" of the participants are formalized in the business requirements for the IT system (Clements, 2010).

The architecture of the proposed services is described below, taking into account the relationship between them and the problems corresponding to each of the components. After describing the elements of the software asset management system, a comparative analysis of the application implementation options is presented.

A software asset management system is an important tool for businesses and organizations, helping them manage the software used in their infrastructure. Within the framework of this system, the key concepts are software, application program and software architecture, which are highlighted below

Software is the main component of SAM. It includes programs, procedures, rules and any relevant documentation related to the operation of the computer system. It can be either system software necessary for computer operation, or application software designed to solve specific tasks, such as word processing or graphic editors.

An application program is a computer program designed to perform a specific task other than the one related to the operation of the computer itself, usually used by end users.

According to Clements et al., software architecture is a set of structures necessary for reasoning about a system, which includes program elements, the relationships between them and the properties of both. The software architecture of the system is a design solution related to the overall structure and behavior of the system. The architecture helps stakeholders understand and analyze how the system will provide such important qualities as modifiability, availability and security.

In the ISO standards, the author could not find a separate definition for the application architecture. Based on ISO/IEC 42010 IEEE 1471, it can be concluded that application architecture is one of several areas of architecture that form the foundations of enterprise architecture (EA).

Application architecture describes the behavior of applications used in business, focusing on how they interact with each other and with users. It is focused on the data consumed and created by applications, not on their internal structure.

The application architecture is determined based on business requirements and functionality. This includes defining the interaction between application packages, databases, and middleware systems in terms of functional coverage (Maidanova, 2020).

It answers the question: “What computer systems (applications) are required to provide the information necessary for business functions?”

There are several approaches to determining the lifecycle of software assets. So, Fadi Nouh in the article SAM Software Asset Management defines the stages of the life cycle as follows.

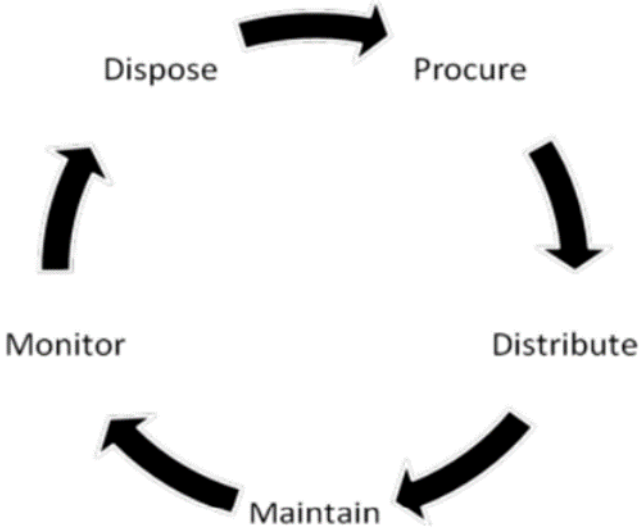


Fig. 1. Stages of the software asset life cycle (Fadi, 2016)

Michael Stone et al. in the work of IT Asset Management, proposed to use, with reservations, a typical asset lifecycle scheme when considering software assets.

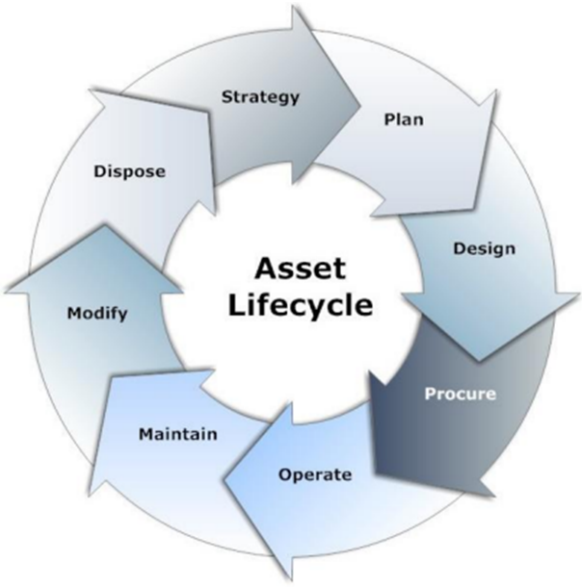


Fig. 2. Typical asset lifecycle (Stone, nd)

The emphasis in these life cycle models is on the sequence of stages and the cyclical nature of the process as a whole; the possibility of returning to previous stages before completing the full cycle is not reflected. In addition to the previous models, the following life cycle model is proposed to focus on generalizing groups of processes.

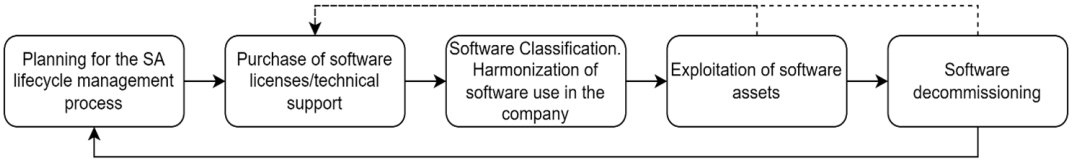


Fig. 3. Software asset management lifecycle

Results and Discussion

Challenges that stakeholders face

This work is based on the general assumption that a company with a large portfolio of software products decides to initiate the development of a SAM system in order to further improve the level of software asset management. When identifying internal stakeholders, user groups have been formed that are most affected by the potential introduction of a new informational system. The main challenges of the company's employees are presented below.

Table 1. Challenges of participants in the process of managing software asset

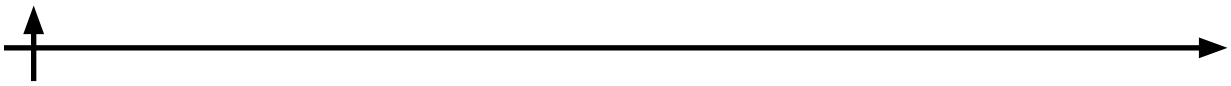
Business user	Challenges	Context
Employee	Difficulties of finding the right software Long time of software installation on workplace Long time of involvement in the workflow	Lack of a clear description of the software, lengthy approval process, possibility of installing prohibited software with administrator rights, absence of software in the catalog despite available licenses
IT Administrator	Inaccurate or incomplete list of software	Difficulty in locating and tracking all installed software, lack of automated software license inventory management, reliance on manual data entry, problems with maintaining and replicating frequent software updates and changes
Procurement Department	Lack of information on software licenses	Inability to promptly respond to changes in demand for software licenses within the organization, difficulties in justifying the position in negotiations with the supplier, violation of the basic principles of category management when working with counterparties.
IS (information security) officer	Compliance risks and audits	Potential non-compliance with software licensing agreements, difficulties in ensuring that license records are accurate and up-to-date, problems in preparing for software audits and responding to license compliance inquiries
Software expert	Inefficient allocation of resources	Problems in optimizing software license allocation, determining the percentage of license usage, constrains in budget planning and cost optimization
IT support	Significant proportion of excessive requests	Increase in number of unstructured requests regarding installation, licensing, uninstallation, compatibility and software performance in general.
Product teams	Absence of centralized software information	The risks of using specific software and the limits of its use are not clear Lack of transparency in the activities of software product teams Problems with regulatory authorities

Functional requirements

Based on the pains of participants in the process of managing software assets, as well as according to the best practices of implementing the SAM concept, the following characteristics can be distinguished that an effective SAM solution should have:

1. Software inventory management:

a. The tool should provide the ability to automatically detect and inventory software resourc-



es throughout the network and infrastructure of the organization.

b. The SAM should support tracking and recording information about software installations, versions, licenses and access rights.

c. The tool should allow manual entry or import of information about software assets for assets that cannot be detected automatically.

2. License Management:

a. The tool should provide technical support for managing and tracking software licenses, including such information as license types, terms, conditions and restrictions.

b. The SAM should provide functionality to link licenses to specific software installations or deployments.

c. The tool should support license usage monitoring and notifications to ensure compliance with license agreements.

3. Compliance and reporting:

a. The tool should generate reports and provide dashboards that allow you to track the status of compliance with software licenses throughout the organization. This should allow you to create customized reports for auditing, license reconciliation, and compliance assessment.

b. The tool should support the ability to export compliance reports in common formats (e.g. PDF, CSV) for easy sharing and documentation.

4. Procurement and contract management:

a. The tool should integrate with procurement processes to track software purchases, contracts, and renewal dates.

b. The SAM should contain alerts and notifications about contract extensions, expiration dates, and important milestones.

c. The tool should support the collection and management of vendor information, including contact details and contract terms.

5. Use and optimization of software:

a. The tool should monitor software utilization rate and provide information about software licensing model metrics, helping to identify underused or unused licenses.

b. The SAM should form optimization recommendations to improve software distribution and reduce costs.

c. The tool should support the identification of software duplication and provide recommendations for consolidation or replacement.

6. Integration with IT infrastructure:

a. The tool should integrate with existing IT infrastructure and systems, such as configuration management databases (CMDBS) and IT asset management tools.

b. SAM should support automatic synchronization of software asset information with other systems to ensure data accuracy and relevance.

c. The tool should provide integration with deployment tools software or IT Service Management Systems (ITSM) to simplify software preparation and management.

7. Security and Risk Management:

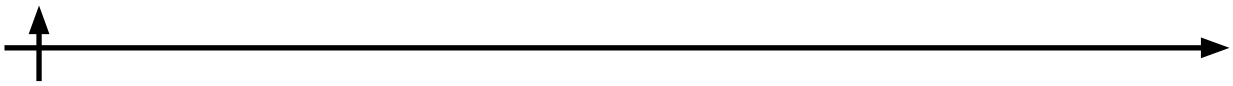
a. The SAM tool should help identify security vulnerabilities related to software assets and provide mechanisms for tracking and managing patch and update levels.

b. It should support security risk assessment by providing information about software versions with known vulnerabilities or end-of-life status.

c. The tool should ensure the implementation and enforcement of security policies related to software use and access control.

8. Scalability and performance:

a. The SAM tool must be scalable to manage the inventory of the organization's software



assets, regardless of the number of assets or distributed locations.

b. It must have effective performance capabilities that allow for quick search and retrieval of information about software assets.

9. User Access and Permissions:

a. The SAM tool should provide role-based access control to ensure that only authorized personnel can view and modify software data.

b. It should allow different levels of access and permissions depending on user roles and responsibilities.

10. User-friendly interface:

a. The SAM tool should have an intuitive and user-friendly interface that simplifies navigation, data entry and reporting functions.

b. It should offer search and filtering options to quickly search for specific software resources or create specific reports.

Based on the functional requirements, as well as the existing ready-made SAM products on the market, an overlapping list of application modules necessary for effective management of software assets in the organization was formed.

Software Registry module

The basic module of the SAM system is the main component that provides the necessary functionality. It includes the following functions:

1. Automatic standardization and ordering of software data;
2. The software catalog, which can be easily changed or expanded with new additions;
3. Classification of software into 16 groups, fixed in the standard GOST R ISO/IEC TO 12182-2002;
5. Automated software commissioning, software status display;
6. Linking software to software assets in management accounting;
7. Information about risks and limitations for OSS licenses;
8. Managing libraries and software components;
9. Mapping and automatic comparison of the business user's software list and the software catalog;
10. Categorization of access control software;
11. Definition of the software owner: definition of a financially responsible person or organization owning software assets;
12. Platform and Installation Type identification: Identification of available platforms and installation types supported by the software;
13. Reference documents related to the software.

Application Showcase module

The main functions of the module:

1. Ordering/installing software;
2. Software search;
3. Role Matrix;
4. Identification of available devices for installation;
5. Configuring Software Availability;
6. Delivery of applications in the App-V format;
7. Providing the ability to deliver and install applications;
8. The ability for managers to delete software;
9. Search for alternatives based on functions;
10. Role-playing system;
11. Publishing any software;

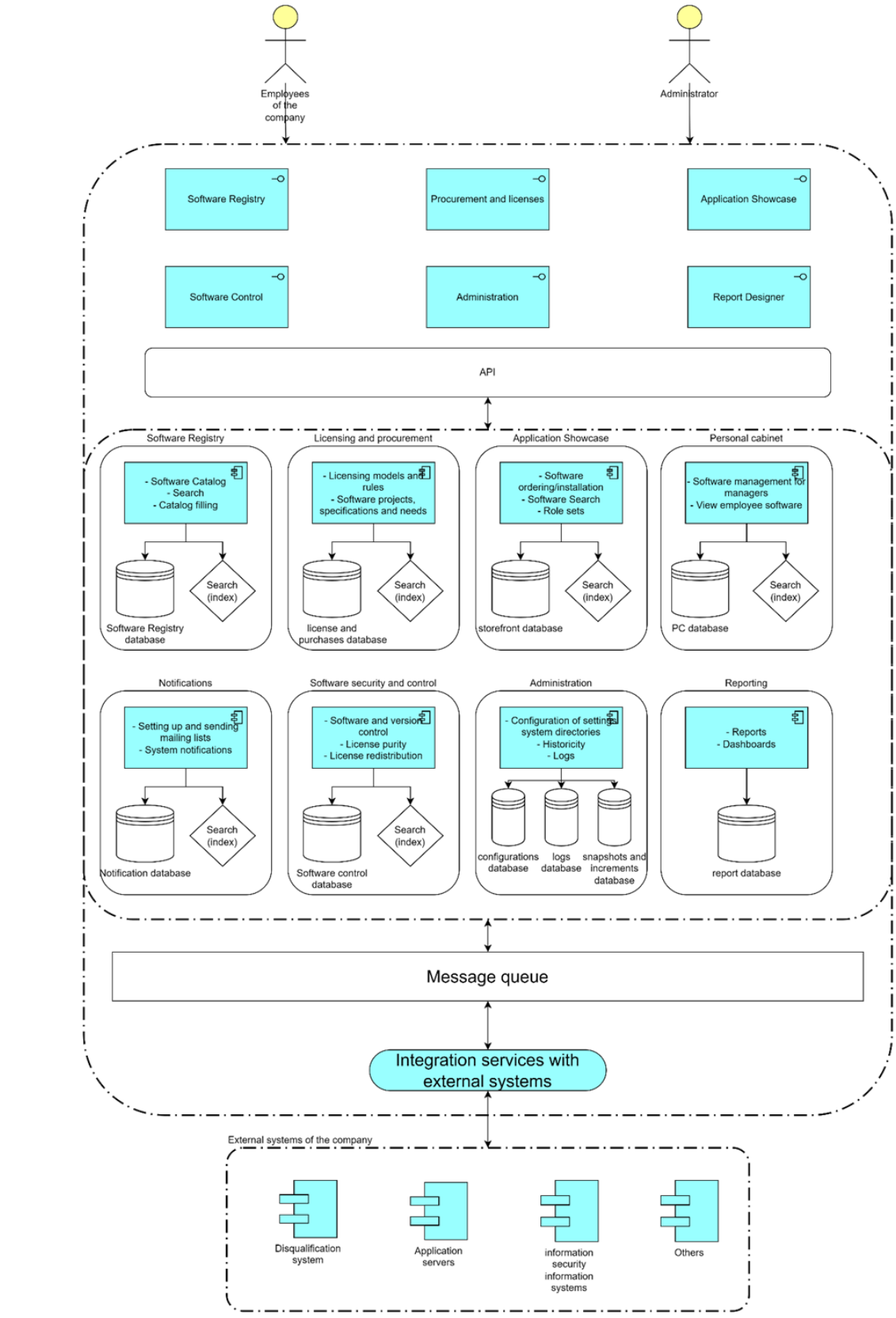
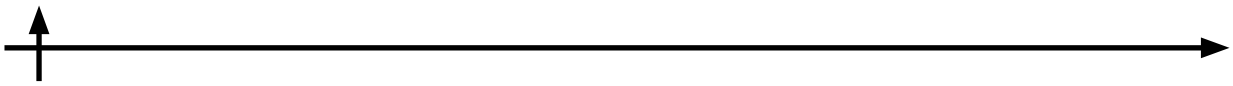


Fig. 4. SAM system architecture



12. Installing software for another employee.

Licensing and Procurement Module

Licensing models and rules: Catalog of current licensing models of software products used in the company, including:

– License types
– Allowed number of installations
– Binding to the core/device/employee/location

Software Procurement Projects and Specifications: A catalog of approved/protected software procurement projects integrated with the company's management and accounting systems, including:

– Project details and codes
– Software nomenclature
– The cost of the software at the time of purchase
– License validity period and technical support
– Contracts and Keys

Catalog of prices and suppliers: Catalog of current prices for items and software kits with information about suppliers, including:

– Current software costs
– Supplier Information
– Comparison of nomenclature names with catalog items

Justification of needs: A module for collecting information to justify needs, such as the actual use of the product for license renewal or rejected software installation requests due to lack of license.

License control and management: License expiration control, automatic distribution of license keys, flexible redistribution of paid software in case of shortage of licenses, management of license agreements, sublicensing and alienation of software.

Integration and procurement: Integration with external contract management systems to search for information, determine the cost of software for project protection, prepare specifications for investment planning, form a purchase queue and automatically determine prices based on previous purchases with correction factors.

Software security and control module

Control of portable/prohibited software is a tool to ensure operational control of portable and prohibited software in the company.

Basic definitions

– Tracking and management of prohibited and sanctioned software

Dynamic, changeable list of prohibited software

– Integration with application uninstall/blocking tools
– Notification of managers about the installation/launch of prohibited software

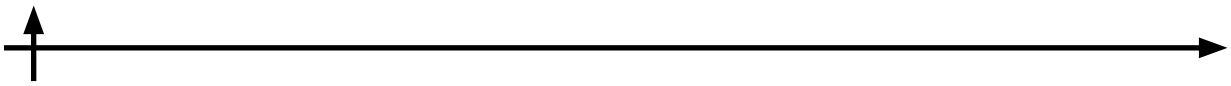
Version control is a tool for monitoring the versioning of the positions of the existing fleet of the company, as well as tracking functional and legal changes with the release of new versions.

Basic definitions:

– Use of current software versions
– Compliance with new copyright holders' requirements

Licensed cleanliness is a tool for ensuring licensed cleanliness. Comparison of data based on two variables: availability of a license for an installed copy of the software, availability of an agreed application (if required).

License redistribution is a tool that allows you to release licenses and redistribute them among company employees. For example, the release of a license due to dismissal or non-use



for more than 90 days.

Discovery and inventory: automatic acquisition of data about workstations, laptops, servers, etc. and software installed on them

Directory of software recognition rules.

Reporting module

Graphical visualization of data. The module allows you to build dynamic reports and graphs for analyzing events related to software management.

Main types of reports:

Cost of ownership of software – a report that allows you to determine the cost of ownership of software by the consolidated parameter: manufacturer, product, functional unit, etc.

Unification of software versions is an analytical report that allows you to bring a large variety of versions of one software name to a single standard

The asset lifecycle is a tool that allows you to determine the entire chronology of the movement of each license from purchase to write–off.

Examples of indicators: software usage (% utilization) taking into account licensing metrics, license purity, financial indicators, validity periods of licenses and contracts.

Administration Module

SAM operation management: configuration settings, alerts, mailing lists, system directories. The administrator role is provided to perform the functions. Ability to configure and create snapshots/increments. The configuration should be made at least in terms of the schedule, data set, storage location, etc.

Personal account module

A tool for:

– Monitoring of the installed software of subordinate employees
– Calculating the cost of software ownership within its management level
– Configurations of non-standard role sets for their employees

Software lifecycle checklist linked to the calendar for those responsible for the software

Notification Module

Mechanisms for sending system notifications, as well as mailings to users according to specified rules.

Conclusion

Effective management of software assets implies the presence of an information system supporting the relevant business processes, which complements the landscape of the enterprise architecture in that part of the functionality that is not provided by standard asset management systems, such as, for example, a procurement management system, inventory management system, security systems, etc. At the same time, the system should ensure integration with related components of the enterprise architecture to avoid duplication of tools with different software solutions.

The architecture of a software asset management system is a structural design and organization of the system itself. It defines the composition, structure and relationship of functional components, modules and data for managing software assets in an organization.

It helps to identify approaches to digitalization of tools to meet the needs of business users.

The implementation of the main functionality of the SAM application is possible within the framework of 8 interrelated modules: Software Registry, Licensing and Procurement, Application Showcase, Personal Account, Notifications, Software Security and Control, Administration, Reporting. The software management tool is aimed at increasing license utilization, minimizing legal, financial and operational risks when using software, increasing awareness

of key indicators of software asset management, improving the efficiency of software lifecycle management business processes and providing complete and structured information about the company's software assets.

REFERENCES

Clements, Wesley A. 2010. Documenting Software Architecture: Views and Beyond (2nd Edition) URL: https://www.iso.org/standard/63598.html (accessed: 05.11.2023).

Fadi N. SAM Software Asset Management. ResearchGate URL: https://www.researchgate.net/publication/291818013_SAM_Software_Asset_Management (accessed: 06.11.2023).

Kalyatin V. 2022. Establishing of Subject of Rights to Intellectual Property Created with Use of Artificial Intelligence Law. Journal of the Higher School of Economics, 4, pp. 24–50.

Maidanova S.A., Ilin I.V. 2022. Development of digital transformation strategy in the context of enterprise architecture. Technoeconomics, 2, 1 (4), 64-75.

Spewak S., Hill S. 1992. Enterprise Architecture Planning: Developing a Blueprint for Data, Applications, and Technology.

Stone M. nd. IT Asset Management. doi: 0.6028/NIST.SP.1800-5

Applications architecture. URL: https://en.wikipedia.org/wiki/Applications_architecture (accessed: 05.11.2023).

GOST 33707-2016 (ISO/IEC 2382:2015) INTERSTATE STANDARD OF Information technologies. Vocabulary. URL: https://docs.cntd.ru/document/1200139532 (accessed: 06.11.2023).

GOST R 57100-2016/ISO/IEC/IEEE 42010:2011 NATIONAL STANDARD OF THE RUSSIAN FEDERATION Systems and software engineering. Architecture description. URL: https://docs.cntd.ru/document/1200139542 (accessed: 05.11.2023)

GOST R ISO/IEC 12207-2010 IS THE NATIONAL STANDARD OF THE RUSSIAN FEDERATION Information technology. System and software engineering. Software life cycle processes. URL: https://docs.cntd.ru/document/1200082859?marker=7EI0KJ&section=text (accessed: 05.11.2023).

GOST R ISO/IEC 19770-1-2014 IS THE NATIONAL STANDARD OF THE RUSSIAN FEDERATION Information technologies. Software asset management. URL: https://docs.cntd.ru/document/1200116600 (accessed: 06.11.2023).

GOST R ISO/IEC TO 12182-2002 is the state standard of the Russian Federation Information technology. Categorization of software. URL: https://docs.cntd.ru/document/1200030161 (accessed: 05.11.2023).

IEEE Recommended Practice for Architectural Description of Software-Intensive Systems. URL: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/http://cabibbo.dia.uniroma3.it/ids/altrui/ieee1471 (accessed: 05.11.2023)

Naumen Software Asset Management. URL: https://www.naumen.ru/products/sam / (accessed: 05.11.2023).

Oxford Dictionary. URL: https://www.oxfordlearnersdictionaries.com/definition/english/application?q=application (accessed: 05.11.2023).

Software Architecture. Carnegie Mellon University. URL: https://www.sei.cmu.edu/our-work/software-architecture / (accessed: 05.11.2023).

СПИСОК ИСТОЧНИКОВ

Clements, Wesley A. 2010. Documenting Software Architecture: Views and Beyond (2nd Edition) URL: https://www.iso.org/standard/63598.html (accessed: 05.11.2023).

Fadi N. SAM Software Asset Management. ResearchGate URL: https://www.researchgate.net/publication/291818013_SAM_Software_Asset_Management (accessed: 06.11.2023).

Kalyatin V. 2022. Establishing of Subject of Rights to Intellectual Property Created with Use of Artificial Intelligence Law. Journal of the Higher School of Economics, 4, pp. 24–50.

Maidanova S.A., Ilin I.V. 2022. Development of digital transformation strategy in the context of enterprise architecture. Technoeconomics, 2, 1 (4), 64-75.

Spewak S., Hill S. 1992. Enterprise Architecture Planning: Developing a Blueprint for Data, Applications, and Technology.

Stone M. nd. IT Asset Management. doi: 0.6028/NIST.SP.1800-5

Applications architecture. URL: https://en.wikipedia.org/wiki/Applications_architecture (accessed: 05.11.2023).

ГОСТ 33707-2016 (ISO/IEC 2382:2015) Межгосударственный стандарт. Информационные технологии. Словарь. URL: https://docs.cntd.ru/document/1200139532 (дата посещения: 06.11.2023).

ГОСТ R 57100-2016/ISO/IEC/IEEE 42010:2011 Национальный стандарт Российской Федерации. Системная и программная инженерия. Описание архитектуры. URL: https://docs.cntd.ru/document/1200139542 (дата посещения: 05.11.2023)

ГОСТ R ISO/IEC 12207-2010 Национальный стандарт Российской Федерации. Информационная технология (ИТ). Системная и программная инженерия. Процессы жизненного цикла программных средств. URL: https://docs.cntd.ru/document/1200082859?marker=7EI0KJ&section=text (дата посещения: 05.11.2023).

ГОСТ R ISO/IEC 19770-1-2014 Национальный стандарт Российской Федерации. Информационные технологии (ИТ). Менеджмент программных активов. URL: https://docs.cntd.ru/document/1200116600 (дата посещения: 06.11.2023).

ГОСТ R ISO/IEC ТО 12182-2002 Государственный стандарт Российской Федерации. Информационная технология (ИТ). Классификация программных средств. URL: https://docs.cntd.ru/document/1200030161 (дата посещения: 05.11.2023).

IEEE Recommended Practice for Architectural Description of Software-Intensive Systems. URL: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/http://cabibbo.dia.uniroma3.it/ids/altrui/ieee1471 (accessed: 05.11.2023)

Naumen Software Asset Management. URL: https://www.naumen.ru/products/sam / (accessed: 05.11.2023).

Oxford Dictionary. URL: https://www.oxfordlearnersdictionaries.com/definition/english/application?q=application (accessed: 05.11.2023).

Software Architecture. Carnegie Mellon University. URL: https://www.sei.cmu.edu/our-work/software-architecture / (accessed: 05.11.2023).

INFORMATION ABOUT AUTHOR / ИНФОРМАЦИЯ ОБ АВТОРЕ

BLINOV Sergey V. – student.
E-mail: st068499@student.spbstu.ru
БЛИНОВ Сергей Валерьевич – студент.
E-mail: st068499@student.spbstu.ru

Статья поступила в редакцию 29.11.2023; одобрена после рецензирования 15.12.2023; принята к публикации 18.12.2023.

The article was submitted 29.11.2023; approved after reviewing 15.12.2023; accepted for publication 18.12.2023.

Technoeconomics. 2023, Vol. 2, No. 4 (7). Pp. 38–45.
Техноэкономика. 2023, Том 2, № 4 (7). С. 38–45.

Scientific article
UDC 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.4

DIGITAL TRANSFORMATION OF TRADE: TRENDS, STAGES AND FACTORS OF DIGITALIZATION AT THE SECTORAL LEVEL

Boburbek Giyosidinov[1], Vladimir Fedorchuk[2], Olga Voronova[2] ✉

[1] Tashkent State Transport University, Tashkent, Uzbekistan;
[2] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ iliina_ov@spbstu.ru

Abstract. This research aims to identify the key stages of digital transformation of trade. In the course of digitalization of the world economy, certain transformations are taking place at the level of all its sectors. This phenomenon is explained by the fact that the result of the introduction of digital technologies is a complex transformation of economic models, which implies the formation of new management systems, business models, types of social attitudes and consumer societies, i.e. the digital transformation of the economy consisting of many sectors. However, it is important to realize that the emergence of new digital technologies and knowledge, the possibility of their application in different ways is reflected in the development of each individual industry, which indicates the relevance of this study. When studying the state of any industry at the current stage of economic development, we face the need to consider the process of its digital transformation. In this regard, analysts in the development of trading companies need to identify and systematize data on the sequence of stages of digital transformation in trade, as well as their content. In the process of the research the main directions and trends of digitalization of trade are considered, the interrelation of the factors of digitalization of the economy with the processes of development and digital transformation of world trade at the industry level is analyzed. As a result of the study, the author's vision of the main stages of digital transformation of trade is presented.

Keywords: digital transformation, digitalization, trade, digitalization factors, digitalization stages, digital space

Citation: Giyosidinov B., Fedorchuk V., Voronova O. Digital transformation of trade: trends, stages and factors of digitalization at the sectoral level. Technoeconomics. 2023. 2. 4 (7). 38–45. DOI: https://doi.org/10.57809/2023.2.4.7.4

Научная статья
УДК 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.4

ЦИФРОВАЯ ТРАНСФОРМАЦИЯ СФЕРЫ ТОРГОВЛИ: ТЕНДЕНЦИИ, ЭТАПЫ И ФАКТОРЫ ЦИФРОВИЗАЦИИ НА ОТРАСЛЕВОМ УРОВНЕ

Бобурбек Гиёсидинов[1], Владимир Федорчук[2], Ольга Воронова[2] ✉

[1] Ташкентский государственный транспортный университет, Ташкент, Узбекинстан;
[2] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉ iliina_ov@spbstu.ru

Аннотация. Целью данного исследования является определение ключевых этапов цифровой трансформации торговли. В ходе цифровизации мировой экономики на уровне всех её отраслей происходят определённые трансформации. Данное явление объясняется тем, что результатом внедрения цифровых технологий и решений является комплексное преобразование моделей хозяйствования, что подразумевает под собой формирование новых систем управления, бизнес-моделей, типов общественных отношении и потребительских обществ, то есть цифровую трансформация экономики, состоящей из множества секторов. Однако важно понимать, что появление новых цифровых технологий и знаний, возможность их применение по-разному находит своё отражение в процессе развития каждой отдельно взятой отрасли, что свидетельствует об актуальности представленного исследования. В попытках проведения исследования состояния любой отрасли на современном этапе развития экономки мы сталкиваемся с необходимостью изучения процесса её цифровой трансформации. В связи с этим у аналитиков в области развития торговых компаний возникает необходимость в определении и систематизации данных о последовательности этапов цифровой трансформации торговли и их содержание. В процессе исследования рассмотрены основные направления и тенденции цифровизации сферы торговли, проанализирована взаимосвязь факторов цифровизации экономики с процессами развития и цифровой трансформации мировой торговли на отраслевом уровне. В результате исследования представлено авторское видение основных этапов цифровой трансформации торговли.

Ключевые слова: цифровая трансформация, цифровизация, сфера торговли, факторы цифровизации, этапы цифровизации, цифровое пространство

Для цитирования: Гиёсидинов Б., Федорчук В., Воронова О. Цифровая трансформация сферы торговли: тенденции, этапы и факторы цифровизации на отраслевом уровне // Техноэкономика. 2023. Т. 2, № 4 (7). С. 38–45. DOI: https://doi.org/10.57809/2023.2.4.7.4

Introduction

The driver of the formation of the digital economy has always been technological innovation, which due to the synergistic effect forms the foundation of the digital economy, including the transition to digital technologies and the distribution of ICT-based goods and services. At the same time, the digital transformation of the economy itself is a complex of economic and social effects, which, in turn, allows us to decompose it as a set of factors affecting the development of the global economy and its sectors. To better understand how the processes of digitalization and the development of world trade are interrelated, let us present a scheme of their relationship in Figure 1.

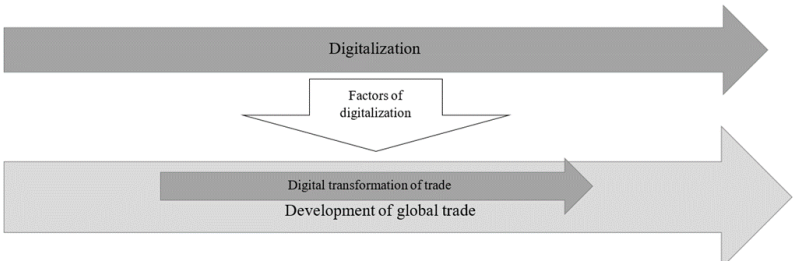


Fig. 1. The impact of digitalization on the development of global trade

According to the figure, digital transformation, being a sub-process of the process of world trade development, is influenced by digitalization factors. In other words, digitalization, including the development of information technologies, knowledge about them, data collection and processing capabilities, forms certain conditions for the development of trade, in connection with which the industry itself begins its transformation by applying these technologies and capabilities (Voronova, Kharyova, 2019). Thus, the analysis of the content of these factors in the context of digital transformation of trade will allow us to identify its main trends, which in turn will allow further research to assess their impact on the state of world trade in general and trading companies in particular (Ilyina, Mikhailova, 2013).

Materials and Methods

In this study, analytical methods were used, such as description, grouping of data, causal analysis, evaluation. The use of these methods allowed linking disparate facts about the current situation in the development of digitalization processes in trade in modern economic conditions. Following the National Program "Digital Economy of the Russian Federation", the digital transformation of trade includes activities aimed at achieving the following goals: improving the quality of goods and services, their accessibility, and the degree of public awareness. These goals were observed, described and classified as well.

Results and Discussion

Combining the listed objectives of digital transformation of trade, technological innovations and their impact on trade, as well as the very scheme of trade activity realization, we can conditionally represent the process of digital transformation of the economy with the help of a flowchart using Figure 2.

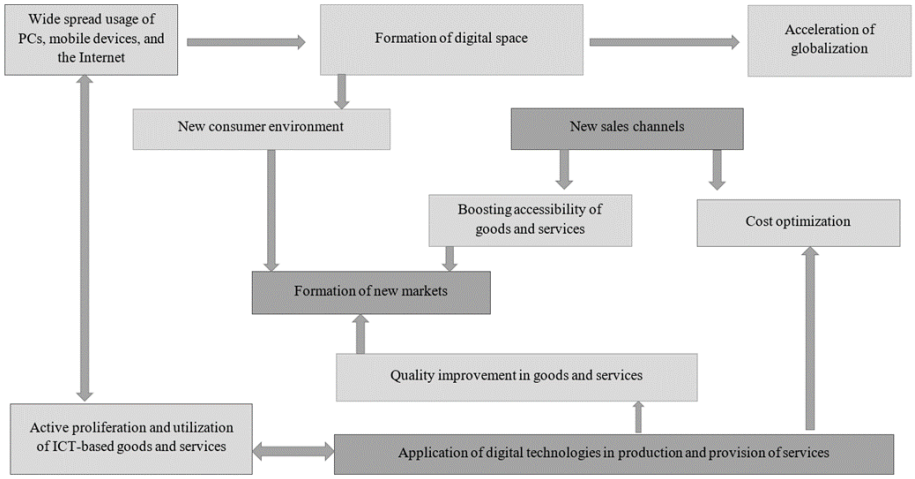
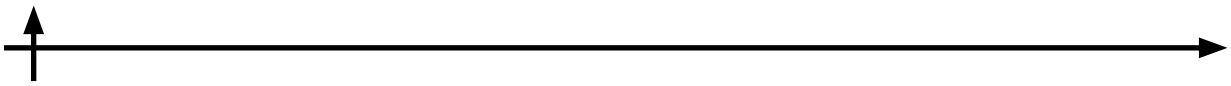


Fig. 2. Results and factors of digitalization at the sectoral level (trade)



Let us explain the meanings of the elements of the figure. As already noted, the digital economy is based on technological innovation, which at the sectoral level of trade is primarily seen as new technologies in the production and service process, as well as the proliferation of PCs, communications and the Internet. The second of the above-mentioned points allows for the formation of a digital space in the course of digitalization, which in turn can be seen as a new model of consumer environment and sales (and purchasing) channels. New opportunities in sales allow consumers to increase their awareness of their goods and services, as well as to optimize costs (marketing, transactions, etc.). On the other hand, the application of new technologies in the course of production and service activities allows, on the one hand, to improve the quality of goods and services, on the other hand, to optimize the resource and material base, and, as a consequence, to optimize this type of costs (Kokova, 2022). A striking example of cost optimization is the reduction of the role of intermediaries through the use of platform solutions in logistics, which allows to expand the possibilities of communication with consumers (Voronova, Ilyin, 2020).

Increasing availability and quality of goods and services form new markets supported by new information technologies. At the same time, it is important to note that companies themselves, in order to maintain all trade, production and service activities, and consumers - to preserve their opportunities in the new digital space, constantly require new goods and services based on newer possible ICT, which leads to the intensification of their use (Noga, 2021; Skolkovo, 2016).

In this regard, world trade is actively changing the structure of its turnover, approaches to organizational and management, marketing and operational activities, as well as to the organization of other socio-economic processes of trading companies, the key driver of development of which is the creation of a comfortable consumer environment.

The scheme presented in the figure above allows us to understand that the following main directions can be distinguished within the digital transformation process:

– Transformation of production activities through the application of digital technologies

– Transformation of logistics and sales activities in the context of new markets and sales channels;

– Transformation of marketing activities due to the formation of a new consumer environment.

Conclusion

In order to competently build the company's strategy and implement the principles of the process approach, which is an integral part of digitalization, it is necessary to understand that the first stage in the transformation will be an audit, which in the conditions of digital reality will consist of so-called digitization: modeling of business processes, the company's business model, determining the vector of development in the field of innovation and the formation of information and regulatory framework.

Based on digital modeling of the company's activities and installation of the necessary digital resources, it will be possible to move to the "qualitative stage" of transformation and start complex work to improve activities in the era of the digital economy, which includes a complex system of digital solutions. At the same time, it is important to realize that first of all, a new digital and technological infrastructure should be developed and applied, and only then should new technologies be applied, including, for example, e-commerce, cloud solutions, Big Data, etc (Barkalova, 2021; Pezzella and Pliushch, 2022).

When the infrastructure is prepared and new digital technologies are applied, new data processing capabilities are used to analyze the company's operations and the market in order to

make further objective management decisions (Krymov, 2019).

To some extent, this process can return the organization's activity to the first, "quantitative" stage and rethink the model of the enterprise's activity. When all the plans and decisions established in the beginning in the field of formation of production activities have been achieved, the equally significant stage of developing customer-oriented solutions begins (Ilyina, Kapustina, 2015).

New markets functioning in the digital space, allowed consumers to expand the choice of goods and services, to detail their requests, in this regard, offering him a convenient technology to choose and make a purchase in the electronic environment today is the main competitive advantage of trade enterprises, in this regard, the formation of a comfortable consumer environment stands out as a separate driver of industry development. This idea arises from the idea that the "new generation consumer" uses digital space not only as a place of recreation, but also as a place of work and the main channel of communication and purchase of goods.

When all models and processes have been digitized, relevant digital and technological infrastructures have been created, and customer-centric business strategies have been developed, the final - transformational - stage comes. It is at this stage that a trading enterprise, having restructured its business model, is ready to fully realize its activities in the new digital space. Based on the above, the authors have systematized the main stages of digital transformation of the trade sphere, presented in Figure 3.

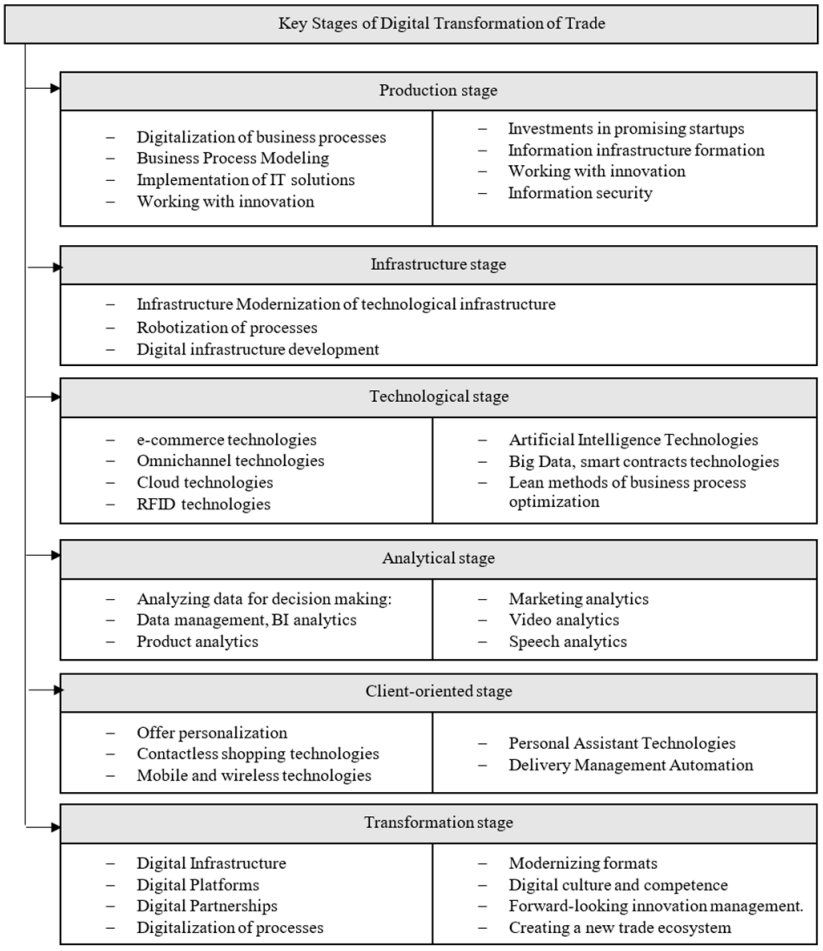


Fig. 3. Main stages of digital transformation of trade

It should be noted that due to the continuity of the digitalization process, digital transformation itself does not have a final form at this stage of economic development. The technologies and capabilities of enterprises in the digital environment are developing faster than businesses can adapt to them, so the transformation stage itself cannot be considered finite. In this regard, the most actively developing companies place great emphasis on staying ahead of innovations and continuous modernization of business formats.

REFERENCES

Barkalova N. A. 2021. Target indicators and criteria of comfortable consumer environment as a sign of balance of different formats of trade. Donetsk Readings 2021: Education, Science, Innovation, Culture and Challenges of Modernity: Proceedings of the VI International Scientific Conference.

Ilyina O.V., Kapustina I.V. 2015. Justification of methodological approaches to the definition of the system of indicators of food security in the region (on the example of St. Petersburg).

Ilyina O.V., Mikhailova G.V. 2013. Integration processes in the sphere of foreign trade. Sovremennaya nauka: actual problems of theory and practice. Series: Economics and Law, 12, 66-71.

Ilyina O.V., Mikhailova G.V. 2015. Features of state regulation of foreign trade in services in the new economic conditions set by the WTO rules. In the world of scientific discoveries, 5 (65), 127-142.

Kokova S. F., Dyshekova A. A. 2022. Digital transformation of industries: starting conditions and priorities. Journal of Applied Research, 7 (6), 577-585.

Krymov S. M. 2019. Structural changes in the chain of goods movement under the influence of new information capabilities. Intellect. Innovations. Investments. 2019, №3, c. 36-42.

Noga V. I. 2021. Features of the behavior of "digital man" in the conditions of the new reality of the world economy. Human Progress, 7 (2), 10 -18.

Pezzella E., Pliushch E. G. 2022. Digital transformation of business: use of blockchain in the oil & gas industry. Technoeconomics, 3 (3), 4-16. DOI: https://doi.org/10.57809/2022.3.3.1

Voronova O.V., Ilyin I.V., Khareva V.A. 2020. Izvestiya St. Petersburg State Economic University, 5 (125), 117-124.

Voronova O.V., Ilyin I.V., Khareva V.A. 2020. Development of the architectural model of business services of the system of interaction with consumers of network trading companies. Izvestiya St. Petersburg State University of Economics, 6, 126.

Voronova O.V., Kharyova V.A. 2019. Network retail FMCG-segment in the Russian Federation: current state and problems of development. International Scientific Journal, 2, P. 7-16.

GOST. 2015. GOST R ISO 9001 Quality management systems. Requirements (Reissued)

Government of the Russian Federation. 2017. Program "Digital Economy of the Russian Federation".

NRU HSE. 2021. Digital Transformation of Industries: Starting Conditions and Priorities: Report to the XXII April International Scientific Conference on Problems of Development of Economy and Society.

OECD. 2015. OECD Digital Economy Outlook 2015, OECD Publishing, Paris. 2015

Problems of modern economics. 2015, 4 (56), 211-214.

Skolkovo. 2016. Digital Life of Russian Megacities. Model. Dynamics. Examples.

СПИСОК ИСТОЧНИКОВ

Баркалова Н. А. 2021. Целевые индикаторы и критерии комфортной потребительской среды как признак сбалансированности различных форматов торговли. Донецкие чтения 2021: образование, наука, инновации, культура и вызовы современности: Материалы VI Международной научной конференции.

Ильина О.В., Капустина И.В. 2015. Обоснованияе методических подходов к определению системы показателей продовольсивенной безопасности региона (на

примерме Санкт-Петербурга)

Ильина О.В., Михайлова Г.В. 2013. Процессы интеграции в сфере внешней торговли. Современная наука: актуальные проблемы теории и практики. Серия: Экономика и право, 12, 66-71.

Ильина О.В., Михайлова Г.В. 2015. Особенности государственного регулирования внешней торговли услугами в новых экономических условиях, задаваемых правилами ВТО. В мире научных открытий, 5 (65), 127-142.

Кокова С. Ф., Дышекова А. А. 2022. Цифровая трансформация отраслей: стартовые условия и приоритеты. Журнал прикладных исследований, 7 (6), 577-585.

Крымов С. М. 20119. Структурные изменения в цепочке товародвижения под влиянием новых информационных возможностей. Интеллект. Инновации. Инвестиции. 2019, №3, с. 36-42.

Нога В. И. 2021. Особенности поведения «человека цифрового» в условиях новой реальности мировой экономики. Human Progress, 7 (2), 10 -18.

Pezzella E., Pliushch E. G. 2022. Digital transformation of business: use of blockchain in the oil & gas industry. Technoeconomics, 3 (3), 4-16. DOI: https://doi.org/10.57809/2022.3.3.1

Воронова О.В., Ильин И.В., Харева В.А. 2020. Развработка архитектурной модели бизнес-сервисов системы взаимодействия с потребителями сетевых торговых компаний. Известия Санкт-Петербургского государственного экономического университета, 6, 126.

Воронова О.В., Ильин И.В., Харева В.А. 2020. Методологические основы формирования системы требований к архитектуре сервисов сетевых торговых компаний. Известия Санкт-Петербургского государственного экономического университета, 5 (125), 117-124.

Воронова О.В., Харёва В.А. 2019. Сетевой ритейл FMCG-сегмента в Российской федерации: современное состояние и проблемы развития. Международный научный журнал, 2, С. 7-16.

ГОСТ. 2015. ГОСТ Р ИСО 9001 Системы менеджмента качества. Требования (Переиздание)

Правительство РФ. 2017. Программа «Цифровая экономика Российской Федерации».

НИУ ВШЭ. 2021. Цифровая трансформация отраслей: стартовые условия и приоритеты: доклад к XXII Апрельской международной научной конференции по проблемам развития экономики и общества.

OECD. 2015. OECD Digital Economy Outlook 2015, OECD Publishing, Paris. 2015

Проблемы современной экономики. 2015. № 4 (56). С. 211-214.

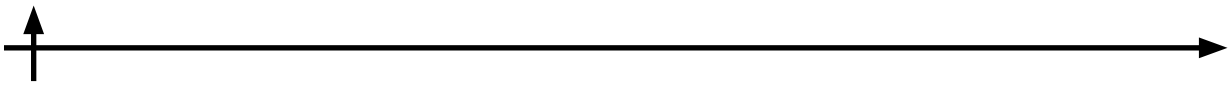
Сколково. 2016. Цифровая жизнь Российских мегаполисов. Модель. Динамика. Примеры.

INFORMATION ABOUT AUTHORS / ИНФОРМАЦИЯ ОБ АВТОРАХ

GIYOSIDINOV Boburbek B. – associate professor.
E-mail: bgiyosidinov@gmail.com
ГИЁСИДИНОВ Бобурбек Бахтиёрович – старший преподаватель.
E-mail: bgiyosidinov@gmail.com

FEDORCHUK Vladimir N. – student.
E-mail: fedvn89@gmail.com
ФЕДОРЧУК Владимир Николаевич – студент.
E-mail: fedvn89@gmail.com

VORONOVA Olga V. – Associate Professor, Candidate of Economic Sciences
E-mail: ilina.olga@list.ru
ВОРОНОВА Ольга Владимировна – доцент, к.э.н.
E-mail: ilina.olga@list.ru
ORCID: https://orcid.org/0000-0003-1032-7173



Статья поступила в редакцию 17.11.2023; одобрена после рецензирования 24.11.2023; принята к публикации 28.11.2023.

The article was submitted 17.11.2023; approved after reviewing 24.11.2023; accepted for publication 28.11.2023.

Technoeconomics. 2023, Vol. 2, No. 4 (7). Pp. 46–60.
Техноэкономика. 2023, Том 2, № 4 (7). С. 46–60.

Scientific article
UDC 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.5

DEVELOPMENT OF A MODEL FOR AUTOMATING THE SALES PROCESS OF ADVERTISING MATERIALS IN MEDIA HOLDING

Veronika Motychko ✉

OOO "Cloud Technologies", Moscow, Russia

✉ veronikamotychko@gmail.com

Abstract. In the modern digital age, media holdings face challenges and opportunities related to managing the lifecycle of sales of advertising materials. Today, the media holding market has become very dynamic and competitive. Large amounts of information, constant changes in the requirements of advertisers and consumers, as well as a variety of communication channels require media holdings to quickly adapt and be flexible in managing sales processes. This study will propose an automation model as a basis for analyzing the maturity and improvement of such a process in companies. As a result of the research, a TO BE automation model was obtained, and a list of the main digital services necessary for this kind of automation was proposed. The research is based on the existing literature and a comparative example involving the company, the media holding in the Russian Federation.

Keywords: automation, digital solution, digitalization, CRM, customer experience, LTV, big data, sales process, media holding

Citation: Motychko V. Development of a model for automating the sales process of advertising materials in media holding. Technoeconomics. 2023. 2. 4 (7). 46–60. DOI: https://doi.org/10.57809/2023.2.4.7.5

Научная статья
УДК 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.5

РАЗРАБОТКА МОДЕЛИ АВТОМАТИЗАЦИИ ПРОЦЕССА ПРОДАЖ РЕКЛАМНЫХ МАТЕРИАЛОВ В МЕДИАХОЛДИНГЕ

Вероника Мотычко ✉

ООО «Облачные Технологии», Москва, Россия

✉ veronikamotychko@gmail.com

Аннотация. В современную цифровую эпоху медиахолдинги сталкиваются с проблемами и возможностями, связанными с управлением жизненным циклом продаж рекламных материалов. Сегодня рынок медиахолдинга стал очень динамичным и конкурентным. Большие объемы информации, постоянные изменения требований рекламодателей и потребителей, а также разнообразие каналов коммуникации требуют от медиахолдингов быстрой адаптации и гибкости в управлении процессами продаж. В этом исследовании будет предложена модель автоматизации в качестве основы для анализа зрелости и улучшения такого процесса в компаниях. В результате исследования получена модель автоматизации TO BE и предложен перечень основных цифровых сервисов, необходимых для такого рода автоматизации. Исследование основано на существующей литературе и сравнительном примере с участием компании-медиа-холдинга в Российской Федерации.

Ключевые слова: автоматизация, цифровое решение, цифровизация, CRM, клиентский опыт, LTV, большие данные, процесс продаж, медиахолдинг

Для цитирования: Мотычко В. Разработка модели автоматизации процесса продаж рекламных материалов в медиахолдинге // Техноэкономика. 2023. Т. 2, № 4 (7). С. 46–60. DOI: https://doi.org/10.57809/2023.2.4.7.5

Introduction

Today, the work of any large organization is associated with the need to automate the enterprise process. Process automation is an integrated approach that describes the use of technology and systems to perform tasks and operations without significant human intervention. As Roman Okorokov, Anna Timofeeva and Tatyana Kharlamova wrote in their article: "The current stage of human development is characterized by the intensive construction of a digital economy, the parameters of which are fundamentally different from those of the previously dominant industrial economy" (Okorokov, 2019).

The topic of process automation in general is one of the key topics of modern business in various fields. As Natalia Alekseeva, Alexander Babkin, Anna Jung, Svetlana Krechko, Hanon Barabaner write: "One of the main trends in modern life is the introduction of digital technologies into people's lives, society and business". Media holdings that actively adapt to digital technologies and apply innovative solutions in their sales processes have a significant competitive advantage. In the context of rapid changes in consumer behavior and the transition to online communication channels, media holdings are forced to rethink their strategies and implement digital tools in order to remain relevant and successful in the market.

The modern media market has undergone major transformations as a result of the rapid development of information and communication technologies. In a competitive environment, Russian advertising agencies are dynamically developing and increasingly integrating into the digital environment. In recent years, the number of creative agencies and groups that operate

using new technologies has increased significantly. These include media buyers, PR agencies, online advertising agencies, web design and others (Ketova, 2020).

One of the main trends in the modern media market is the increase in the volume and diversity of data. Media holdings have access to huge amounts of data, including information about consumers, their preferences, behavior and reactions to advertising. The use of analytics and machine learning allows media holdings to extract valuable information from this data, identify trends, predict customer behavior and make more informed decisions in the sales process.

Another significant trend is the personalization of advertising. Modern consumers expect personalized and relevant content that takes into account their interests, preferences and needs. Media holdings are faced with the task of creating personalized advertising materials and adapting them to a specific consumer. Digital tools such as BPMS, low-code and no-code platforms allow media holdings to create and customize personalized advertising campaigns with minimal effort and reduced time.

The purpose of introducing digital solutions into the life cycle of sales of advertising materials of a media holding is to increase the efficiency and effectiveness of commercial activities. Automation of sales processes allows you to reduce the time of the sales cycle, speed up order processing and improve the quality of customer service. This helps to increase sales and customer satisfaction, as well as reduce costs and improve the operational efficiency of the media holding.

To improve the quality of automation of the media holding sales process, it is necessary to solve a number of tasks, including identifying the prerequisites for implementation, analyzing technologies for automation, studying the process, as well as optimizing such a process taking into account current technologies. Automation of the process will help to reduce the time for processing orders, reduce the number of errors and manual labor. Therefore, automation of the sales process of advertising materials is of particular importance and practical importance.

The purpose of the study is to propose a model for automating the sales process of advertising materials in a media holding as a basis for analyzing the maturity and improvement of such a process in companies.

Selected tasks:

– to analyze articles on the topic of automation of the sales process;
– to analyze the technologies relevant to the automation of the sales process;
– to consider the business process of selling advertising materials in a media holding;
– to develop an optimized process model based on the results obtained during the analysis.

Materials and Methods

For a detailed consideration of the topic, it is necessary to analyze what automation is, its prerequisites, as well as the possibilities of modern digital technologies for automating the sales process of advertising materials.

In order to maintain their place in the market, advertising organizations need to constantly develop and adapt their activities to the trends of behavior of the audience and business entities.A.V. Chekhov notes that currently the main activity of media holdings can be divided into three groups:content production, a combination of content that occurs in the format of printed publications, Internet resources, radio distribution of media products in the form of print media, Internet resources, radio and TV (Chekhov, 2021).

The current research will mainly focus on the TV group and the sale of TV advertising materials.

One of the key tools to achieve this goal is the implementation of a customer relationship

management (CRM) system. CRM systems provide an integrated approach to managing customer data and sales processes, allowing media holdings to effectively interact with customers, improve service and increase sales. This paper will consider the prerequisites for the introduction of CRM into the sales processes of media holding in Russia in 2023, taking into account the withdrawal of their current CRM from the market.

Customer relationship management (CRM) is a specialized technology that enables firms to capture, store, access, share and analyze large quantities of customer data. The potential benefits of using CRM systems include higher customer loyalty, improved marketing effectiveness, better customer service and support and lower costs through improved efficiency.

Prerequisites for the introduction of CRM into sales processes:

Outdated CRM system: One of the main prerequisites for the introduction of a new CRM system is an outdated or poorly functioning current CRM. If a media holding is faced with problems in processing customer information, difficulties in analyzing data, or an inability to provide personalized customer service, replacing the CRM system becomes a necessity.

Expanding the customer base: Media holdings may face a growing customer base and an increase in sales. In such circumstances, a powerful and flexible CRM system is required, capable of efficiently processing large amounts of data, tracking customer interaction and providing analytical tools for making informed decisions.

The need for personalization: In the face of increasing competition and changing consumer preferences, it is important for media holdings to offer personalized content and promotional materials. The CRM system allows you to collect and analyze data about customers, their preferences and behavior, which allows you to create personalized offers and improve the customer experience.

Integration with other systems: Media holdings often use different IT systems to manage different aspects of their business, such as content management, analytics, finance, etc. The integration of the CRM system with other systems allows you to create a single information space, improve data exchange, automate processes and provide a single overview for the entire business.

Improved communication: CRM systems allow you to strengthen internal communication within the media holding and ensure more effective interaction between sales, marketing, analytics and other key functional areas. This helps to improve coordination and cooperation, which has a positive effect on the quality of customer service and sales performance.

Advantages of CRM for business:

1. Optimization of customer experience.
2. Increase the level of customer retention.
3. Increase in sales revenue.
4. Improving the efficiency of processes.
5. Improving the rationality of operations and the level of cooperation.

The main functions of CRM

1. Contact management.
2. Interaction management.
3. Managing potential customers.
4. Automation of work processes.
5. Customer analytics.
6. CRM integration.
7. Secure mobile CRM system.

According to a study by SAP, CRM includes the following functions [6].

Table 1. The main functions of CRM according to SAP research

The main functions of CRM	
Main Sales	Customer Service
Contact management	Contact Center
Interaction management	Omnichannel
Managing potential clients	Email newsletters
Automation of work processes	Engagement
Customer analytics	Surveys
CRM integration	Contact Center
Secure mobility	

As Sikora Melanie Hodgkinson et al. says, the changed trends in the work of the CRM system are aimed at improving the quality of customer service (customer experience) and improving the efficiency of working with customers and interacting with them (Sykora, 2022).

Customer Experience management: One of the main trends in the CRM system is the transition from simple to complex. customer experience management (Holmlund, 2020). CRM systems are installed using software to collect and analyze data related to customer engagement with the campaign. This allows organizations to better understand the needs and preferences of customers, as well as offer personalized solutions and services that meet their expectations (Spewak, 1992).

In general, modern trends in CRM systems are aimed at improving customer interaction, increasing the level of personalization, automating processes and providing deeper data analysis. The introduction of modern CRM systems can help media holdings in Russia to increase the efficiency of sales of advertising materials, improve customer service and achieve a competitive advantage in the market.

Customer expirience within the framework of the implemented CRM:

1. Collection of metrics.
2. Analysis of the customer experience situation (metrics analysis).
3. Visualization of customer expirience reports.
4. Automation of work processes.
5. Adjustment of the service, products and services, taking into account customer feedback.
6. Expected result: improvement of the LTV indicator (LTV is an indicator of the profit that the business receives for all the time it works with the client).

According to Mongo's research, LTV (Lifetime Value) is the customer's lifetime value. LTV shows the profit from the relationship with the client for the entire period — from the moment when he saw the first advertisement or registered on the site, to the last purchase.

LTV is calculated by multiplying the average revenue received from a client over a certain period by the expected duration of cooperation with him. This indicator helps companies understand how much they are willing to invest in attracting and retaining customers (Okorokov, 2019).

The use of the LTV indicator in the automation of the sales process has a number of advantages:

1. Profitability forecasting: LTV calculation allows companies to predict the potential profit that can be obtained from each client. This allows you to make informed decisions about how much resources and effort should be invested in attracting and retaining customers, as well as

optimize marketing and sales strategies.

2. Determining the pricing policy: Knowing the customer's LTV, the company can determine the optimal price of a product or service. If the LTV is high, then the company can set a higher price based on the expected long-term profit. In case of low LTV, the pricing policy can be adjusted to attract and retain customers.

3. Improving Customer Experience: Understanding LTV helps companies identify which customers are most valuable and how to provide them with a higher quality of service. By focusing on customers with high LTV, the company can offer personalized services, bonus programs or additional benefits, which increases customer loyalty and prolongs their life cycle (Provost, 2013).

4. Data-based decision making: LTV calculation is based on the analysis of data on customer behavior, purchases and interaction history. This helps companies make informed decisions based on factual information rather than assumptions. Automation of the sales process using CRM and analytical tools allows you to collect and analyze data for a more accurate calculation of LTV.

5. Optimization of marketing costs: Knowing the customer's LTV, the company can determine how much it can spend on attracting a new customer. This helps to optimize marketing costs and distribute the budget between different channels for attracting customers. The company can also focus on retaining existing customers who have a high LTV, which saves resources and increases the overall profitability of the business.

The use of the LTV indicator in the automation of the sales process allows companies to more accurately analyze and predict the results of their sales, optimize marketing strategies, improve customer service and improve overall business profitability.

The introduction of CRM into the processes of such a large company also requires technology that allows you to work with large amounts of data. As Holmlund M Van Vaerenbergh YCiuchita R et al say, in today's fast-paced digital economy, big data analytics (BDA) has huge potential to enhance customer experience.

The introduction of Big Data technology into the CRM system of media holding has a number of advantages and applicability:

1. Processing and analysis of large amounts of data: Big Data technologies allow you to effectively process and analyze huge amounts of data, including information about customers, advertising campaigns, sales and other important indicators. This allows the media holding to gain a more complete understanding of its customers, their preferences and needs, as well as make more accurate forecasts and make informed decisions.

2. Personalization of offers: with the help of Big Data technology, the media holding can analyze large amounts of data about customers, their interests, preferences, purchase history and behavior. This allows you to create personalized offers and recommendations for customers, taking into account their individual needs and preferences. This approach increases the likelihood of successful sales and improves the customer experience.

3. Optimization of marketing campaigns: Big Data analysis allows the media holding to optimize its marketing campaigns. Using customer data, the behavior of their interaction with advertising materials and sales results, the media holding can determine the effectiveness of various marketing strategies and tactics, as well as optimize budget allocation based on data and insights.

4. Forecasting and analytics: Big Data technologies allow for more accurate forecasting and analytics in the field of sales of advertising materials. The analysis of historical data, customer information and external factors allows the media holding to predict demand, determine the most effective sales strategies, manage inventory and make informed decisions to achieve the

best results.

5. Improving customer retention: Through Big Data analysis, the media holding can better understand its customers, their needs and preferences. This allows us to provide a more personalized and high-quality service, which helps to increase customer satisfaction and retain them on a long-term basis.

Machine Learning Technology (MO) is a branch of artificial intelligence that allows computer systems to learn and improve based on experience and data, without explicit programming.

It is based on algorithms and models that are able to analyze large amounts of data, identify patterns, make predictions and make decisions based on this data.

The use of machine learning technology in the implementation of a CRM system in a media holding can significantly increase the effectiveness of sales of advertising materials. It allows you to predict demand, personalize offers, automate sales processes, analyze data and improve customer retention. As a result, the media holding can increase its competitiveness, improve the quality of customer service and achieve higher profit margins.

The use of machine learning technology in the implementation of a CRM system in a media holding has significant potential and can lead to improved efficiency and effectiveness of sales of advertising materials. The main applications of machine learning in this context are discussed below:

1. Demand forecasting: Machine learning allows you to analyze historical sales data and customer behavior to predict future demand for advertising materials. This allows the media holding to adapt its sales strategies, optimize inventory and manage resources more efficiently.

2. Personalization of offers: With the help of machine learning, you can analyze data about customers, their preferences, purchase history and behavior to create personalized offers. This allows the media holding to offer customers promotional materials that best meet their needs and interests, which increases the likelihood of successful sales.

3. Sales Process Automation: Machine learning can be applied to automate various stages of the sales process, such as lead qualification, customer segmentation, and determining optimal communication strategies. This allows you to reduce the time and effort spent on routine tasks and focus on the more strategic aspects of sales (Kalyatin, 2020).

4. Data analysis and prediction of results: Machine learning can be used to analyze large amounts of data about customers, sales and marketing activities. This helps to identify hidden patterns, trends and patterns that can be useful for making decisions about sales and marketing strategies. Machine learning can also be used to predict sales results and evaluate the effectiveness of marketing campaigns.

5. Improving customer retention: Machine learning allows you to analyze data about customers and their behavior, which helps identify factors that affect customer retention. This allows the media holding to take targeted measures to retain customers, for example, to offer personalized discounts or bonus programs (Bishop, 2006).

The use of machine learning technology in the implementation of a CRM system in a media holding can significantly increase the effectiveness of sales of advertising materials. It allows you to predict demand, personalize offers, automate sales processes, analyze data and improve customer retention. As a result, the media holding can increase its competitiveness, improve the quality of customer service and achieve higher profit margins (Clements, 2010).

This study used a mixed review method combining critical review, synthesis and deduction.

A critical review of the literature in the current study contains a discussion of the prerequisites for the introduction of automation into the sales process in the studies identified in the reviewed studies. In addition, the critical review allowed us to consider and discuss the appli-

cability of technologies in such a process.

The generalization in the current study contains a combination of knowledge about the technologies used in the process, in their applicability in various fields that were considered in the study.

Deduction made it possible, by extracting hypotheses from the studies reviewed, to draw a conclusion in the current study.

The study also used the method of modeling the architecture of processes using the Archi digital solution. Archi is a freely distributed open source cross–platform tool for modeling at all levels of enterprise architecture in terms of the ArchiMate language. Archi is developed and is a registered trademark of Philip Beauvoir. The Orchid software product is based on the Eclipse Rich Client Platform (RCP) framework using the Eclipse IDE integrated development environment.

The basis of the Archi tool is ArchiMate. ArchiMate is an open source enterprise architecture modeling language standard developed by the Open Group consortium. The ArchiMate language is supported by the tools of various vendors and is actively used by consulting companies. It is fully consistent with the TOGAF enterprise architecture model, also supported by the Open Group consortium. ArchiMate supports the description, analysis and visualization of enterprise architecture.

Results and Discussion

After analyzing all the technologies, they were assembled into a single concept to automate the sales process of advertising materials in the media holding. As a result, we get the following solution concept:

The concept of the solution describes how CRM will be integrated with the current systems used in media holding. The client's path from the entry point to processing in the company's systems is shown. In the diagram, you can see the entire client path as it looks in the system. Customer data gets into the CRM, from there via the 1C data bus for billing and contract processing, into the BI system for analytics, etc.

One of the important stages in the client's journey is the client's login to the system. This is where the client's initial contact with the media holding takes place, and information about the client begins to be collected and entered into the CRM system. This may include information about the customer, their contact information, preferences, interaction history, and other important information (Maidanova, 2022).

The next step is data processing in the company's systems. The CRM system interacts with other systems such as 1C (billing and contract processing system) and BI-system (analytics system). The data bus provides information transfer between these systems so that customer data is available and used in the right systems (Hastie, 2009).

Integration with 1C allows you to automate the billing and contract processing processes. The customer data collected in the CRM system is transferred to 1C for generating and billing customers, as well as for processing the necessary contracts. This simplifies and speeds up the process of interacting with customers and increases the accuracy and reliability of financial transactions.

Integration with the BI system allows you to analyze customer and sales data. The data collected in the CRM system is transferred to the BI system for various analytical operations. This may include analyzing the effectiveness of marketing campaigns, tracking sales and conversions, researching customer behavior, and other types of analytics. The results and insights obtained help the media holding to make informed decisions, optimize sales strategies and improve the

overall customer experience (Kalyatin, 2022).

Thus, the integration of CRM with current media holding systems allows you to create a complete system where customer and sales data is efficiently transmitted and processed between different systems. This reduces manual labor, improves data accuracy, increases process efficiency, and enriches analytical capabilities (Ketova, 2022). This integration ensures a harmonious and modern CRM implementation in the media holding, improving the work of the sales department of advertising materials and the overall customer experience.

For the analysis, a was collected and the architecture of the AS IS process was described.

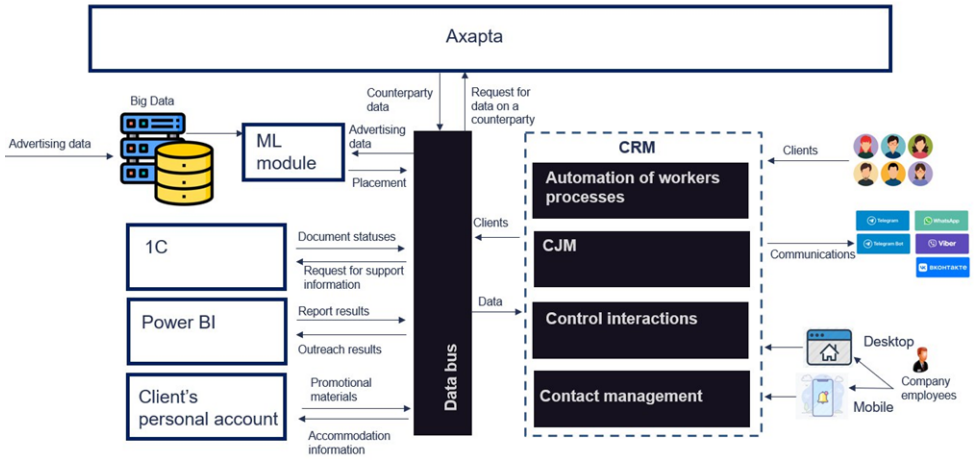


Fig. 1. Solution scheme

As a result of the analysis of the AS IS process, several key problems were found that impede the effectiveness of the advertising sales department in the media holding (Alekseeva, 2020). These problems are related to the lack of integration between the systems, the lack of maintenance processes in the current CRM and the separation of customer information in two different databases. The conclusions drawn from the analysis are presented below:

1. Lack of integration between systems: In the current situation, the sales department is forced to use the mail service to communicate with other departments, such as support. This creates inconveniences and delays in the transmission of information, as well as increases the likelihood of errors and data inconsistencies. This lack of connectivity between the systems makes it difficult to respond promptly to customer requests and reduces the overall effectiveness of the sales process.

2. Lack of maintenance processes in the current CRM: The current CRM system has functionality for pre-sale events, but there are no maintenance processes. This limits the ability to track and support customers after the transaction, which can negatively affect long-term customer relationships and repeat sales. It was decided not to finalize the current CRM system, as it is leaving the market, which requires the search for an alternative solution to ensure maintenance processes.

3. Separation of customer information in two databases: In the media holding, there are two customer databases that are located in different systems. This leads to duplication of data, makes it difficult to access up-to-date information and increases the likelihood of errors in processing client information. The sales department needs to interact with the support department to obtain reliable information about contractors, which requires additional efforts and increases the response time to customer requests.

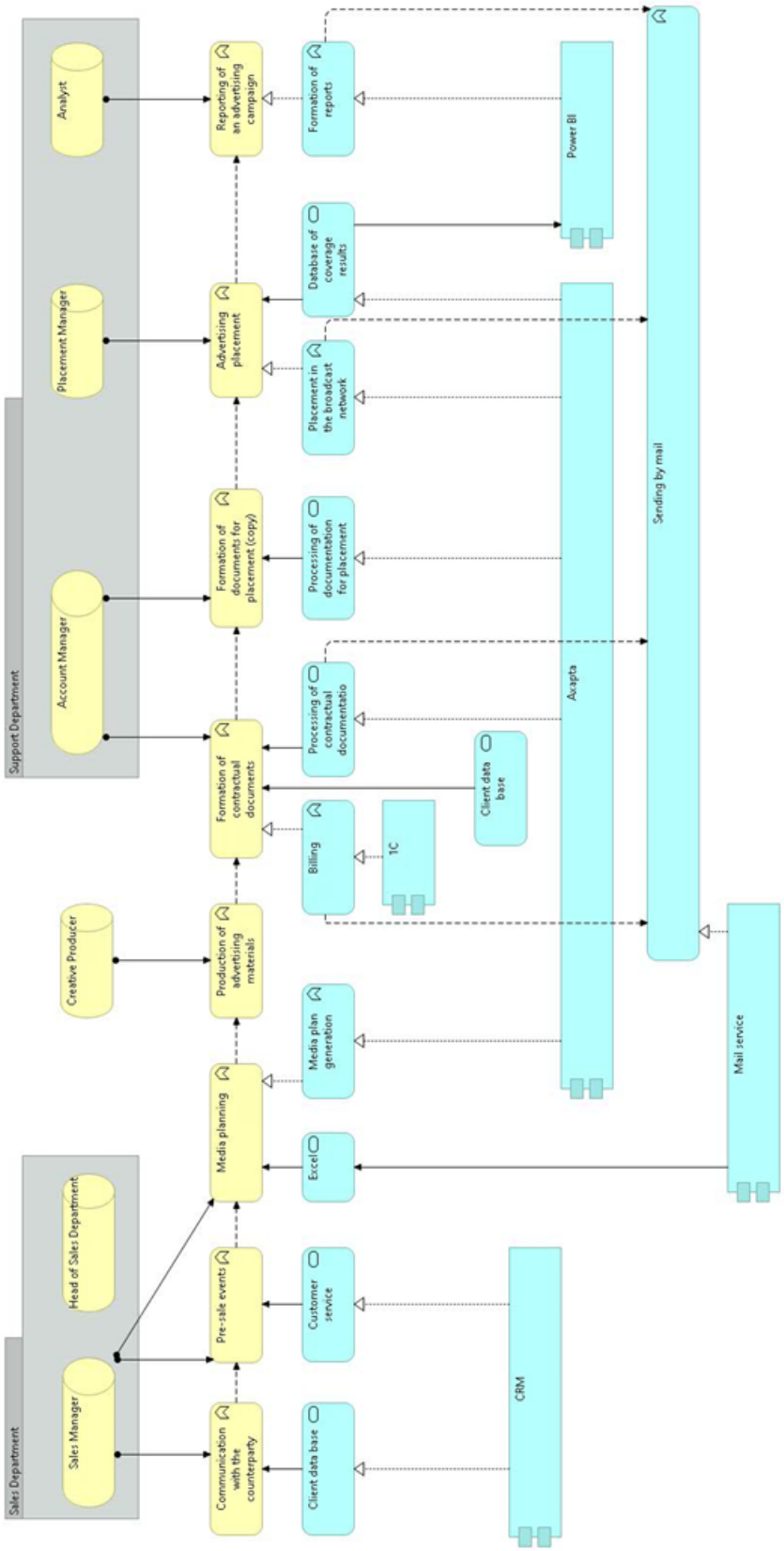
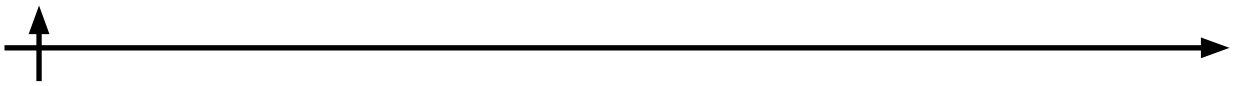


Fig. 2. AS IS scheme



To solve these problems and improve the efficiency of the sales process in the media holding, it was proposed to introduce a new CRM system. This will achieve the following goals:

1. Integration of existing systems: The new CRM system will be integrated with other media holding systems, such as the tracking system and the billing system. This will provide a single repository of customer data, reduce manual processing of information and increase data accuracy.

2. Support for customer support processes: The new CRM system will provide functionality for customer support processes after the transaction is completed. This will allow you to track customer interactions, provide support and improve the level of service, which contributes to customer satisfaction and repeat sales.

3. Centralization of customer data: The introduction of a new CRM system will allow you to combine customer data in a single database. This will simplify access to up-to-date customer information, prevent data duplication, and improve the overall accuracy and reliability of information.

Thus, the introduction of a new CRM system will solve the identified problems and increase the efficiency of the sales process of advertising materials in the media holding. It will provide integration between systems, support for maintenance processes and centralization of customer data, which will lead to improved customer service, optimization of the sales department and strengthening the position of the media holding in the market.

After the implementation of the CRM system and all other related systems described earlier, the process was optimized and brought to the following form TO BE.

After the successful implementation of the CRM system and other related systems, the sales process of advertising materials in the media holding was optimized and brought to a new improved state TO BE. The following changes have been implemented:

1. Unified system for the sales department and the support department: As a result of the implementation of the CRM system, the sales department and the support department now work in a single system. This greatly simplifies communication and information exchange between departments, since all the necessary data and artifacts are available within one system. The sales department receives up-to-date information about customers and the status of transactions, and the support department can provide proper customer support after the transaction is completed (Cao, 2020).

2. Automated media plan generation: The new system has implemented the functionality of automatic media plan generation for all types of sales. This eliminates the possibility of manual errors when generating documentation and improves the accuracy and reliability of creating media plans. Now the process of forming media plans has become more efficient and faster, which saves time and resources of the sales department.

3. Unified customer database: The introduction of a CRM system made it possible to combine customer databases into a single one. Now the sales department has access to a common database, which reflects all the clients of the media holding. This eliminates the need to communicate with the support department by mail to obtain information about the reliability of contractors. The sales department can promptly receive information about customers and make decisions based on up-to-date data, which improves the quality of customer service and increases the efficiency of the department.

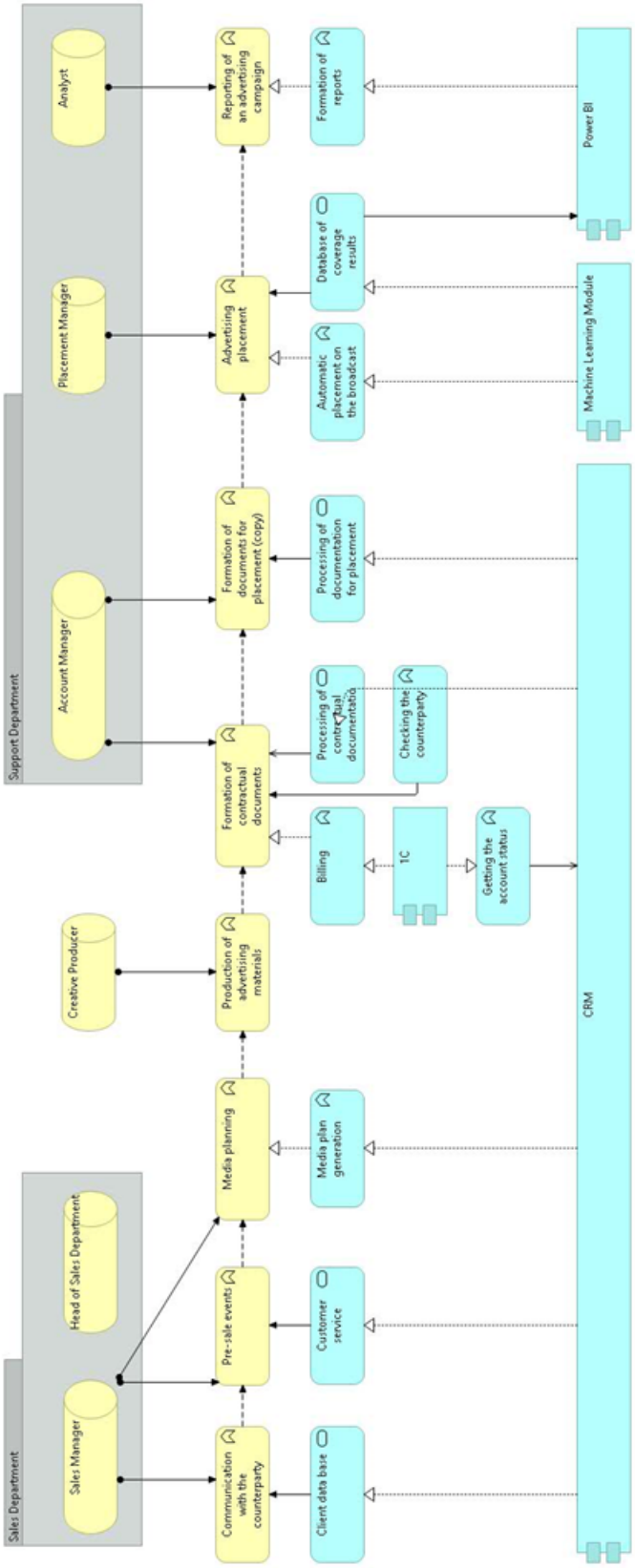
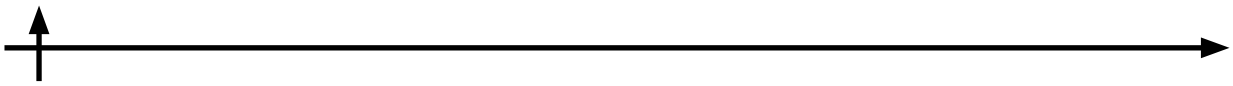


Fig. 3. TO BE scheme



As a result of these changes, the sales process of advertising materials in the media holding has become more advanced and optimized. The sales department and the support department work in a single system, which simplifies communication and information exchange. Automation of media plan generation and a unified customer database allow you to avoid errors, reduce manual operations and improve the accuracy and reliability of data. As a result, the media holding receives advantages in the form of increased efficiency of the sales process, improved customer service and strengthening its position in the market.

Conclusion

The study examined an important topic of automation of the sales process and the introduction of digital solutions in media holding. The purpose of this work was to optimize and improve the sales process of advertising materials, as well as to create an effective system for managing customer data and communications between sales and support departments.

By analyzing the current state of the sales process and identifying its weaknesses, key problems were identified, such as lack of integration between systems, lack of maintenance processes and separated customer databases. The need to solve these problems has become the main motivation for the introduction of a CRM system and other digital solutions.

As a result of the implementation of the CRM system and other related systems, the following goals have been achieved:

1. A single system for the sales department and the support department: The introduction of a CRM system made it possible to combine the work of the sales department and the support department within one system. This ensures more effective communication and information exchange between departments, simplifies work processes and increases the overall efficiency of the media holding. The sales department receives up-to-date information about customers and the status of transactions, and the support department can provide appropriate support to customers after the transaction is completed.

2. Automated media plan generation: The introduction of a CRM system has made it possible to automate the process of generating media plans for all types of sales. This eliminates the possibility of manual errors in the formation of documentation and increases the accuracy and reliability of creating media plans. Now the process of forming media plans has become more efficient and faster, which allows you to save time and resources of the sales department and improve the quality of customer service.

3. Unified customer database: The introduction of a CRM system made it possible to combine customer databases into a single one. Now the sales department has access to a common database, which reflects all the clients of the media holding. This eliminates the need to communicate with the support department by mail to obtain information about the reliability of contractors. The sales department can promptly receive information about customers and make decisions based on up-to-date data, which improves the quality of customer service and increases the efficiency of the department.

Thus, the introduction of digital solutions, including a CRM system, into the sales processes of advertising materials of the media holding has significantly optimized and improved the work of the sales department and the support department. A unified system for both departments ensures more effective communication and information exchange, automated media plan generation reduces the likelihood of errors, and a unified customer database improves the availability and quality of customer information. As a result of the achieved goals, the media holding receives advantages in the form of increased efficiency of the sales process, improved customer service and strengthening its position in the market.

REFERENCES

Alekseeva N., Babkin A., Yung A., Krechko S., Barabaner H. 2020. Digital Transformation Impact on the Intellectual Capital of an Innovatively Active Industrial Cluster. Proceedings of the International Scientific Conference - Digital Transformation on Manufacturing, Infrastructure and Service (DTMIS '20), Association for Computing Machinery, New York, USA, 1–7. doi: https://doi.org/10.1145/3446434.3446442

Bishop C. M. 2006. Pattern Recognition and Machine Learning. Springer.

Cao G., Tian N. 2020. Enhancing customer-linking marketing capabilities using marketing analytics, Journal of Business & Industrial Marketing, 35 (7), 1289-1299.

Chekhov A.V. 2020. Media holdings in the information environment. URL: https://cyberleninka.ru/article/n/mediaholdingi-vinformatsionnoy-srede (accessed: 05/07/23).

Clements, Wesley A. 2010. Documenting Software Architecture: Views and Beyond (2nd Edition) URL: https://www.iso.org/standard/63598.html (accessed: 05.12.2023).

Hastie T., Tibshirani R., Friedman J. 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.

Holmlund M., Vaerenbergh Y., Ciuchita R. 2020. Customer experience management in the age of big data analytics: A strategic framework. Journal of Business Research, 116, 356–65. doi: 10.1016/j.jbusres.2020.01.022

Kalyatin V. 2022. Establishing of Subject of Rights to Intellectual Property Created with Use of Artificial Intelligence Law. Journal of the Higher School of Economics, 4, pp. 24–50.

Ketova N.P., Zunde V.V., Shashnev P.D. 2020. Transformation of the advertising market as a factor in activating the behavior of advertising agencies in the conditions of digitalization. Creative Economy, 14 (5), 953-968.

Maidanova S.A., Ilin I.V. 2022. Development of digital transformation strategy in the context of enterprise architecture. Technoeconomics, 2, 1 (4), 64-75.

Okorokov R., Timofeeva A., Kharlamova T. 2019. Building intellectual capital of specialists in the context of digital transformation of the Russian economy. In IOP Conference Series: Materials Science and Engineering, 497 (1), 012015. doi:10.1088/1757-899X/497/1/012015

Provost F., Fawcett T. 2013. Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media.

Spewak S., Hill S. 1992. Enterprise Architecture Planning: Developing a Blueprint for Data, Applications, and Technology.

Sykora M., Elayan S. 2022. The Power of Emotions: Leveraging User Generated Content for Customer Experience Management». Journal of Business Research, 144. doi: 10.1016/j.jbusres.2022.02.048

LTV: what it is, how to calculate the indicator of lifetime value, what it is needed for in marketing. URL: https://www.mango-office.ru/products/calltracking/glossary/ltv/ (accessed: 24.08.23).

Modeling enterprise architecture with Archi. URL: https://ekonomika.snauka.ru/2014/11/6308 (accessed: 24.08.23).

The Open Group Library. URL: https://publications.opengroup.org / (accessed: 25.08.23).

TOGAF. Www.Opengroup.Org. URL: https://www.opengroup.org/togaf (accessed: 24.08.23).

What is CRM: Definition of customer relationship management. B.D. SAP. 2023. URL: https://www.sap.com/central-asia-caucasus/products/crm/what-is-crm.html (accessed: 23.08.23).

СПИСОК ИСТОЧНИКОВ

Alekseeva N., Babkin A., Yung A., Krechko S., Barabaner H. 2020. Digital Transformation Impact on the Intellectual Capital of an Innovatively Active Industrial Cluster. Proceedings of the International Scientific Conference - Digital Transformation on Manufacturing, Infrastructure and Service (DTMIS '20), Association for Computing Machinery, New York, USA, 1–7. doi: https://doi.org/10.1145/3446434.3446442

Bishop C. M. 2006. Pattern Recognition and Machine Learning. Springer.

Cao G., Tian N. 2020. Enhancing customer-linking marketing capabilities using marketing

analytics, Journal of Business & Industrial Marketing, 35 (7), 1289-1299.

Chekhov A.V. 2020. Media holdings in the information environment. URL: https://cyberleninka.ru/article/n/mediaholdingi-vinformatsionnoy-srede (accessed: 05/07/23).

Clements, Wesley A. 2010. Documenting Software Architecture: Views and Beyond (2nd Edition) URL: https://www.iso.org/standard/63598.html (accessed: 05.12.2023).

Hastie T., Tibshirani R., Friedman J. 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.

Holmlund M., Vaerenbergh Y., Ciuchita R. 2020. Customer experience management in the age of big data analytics: A strategic framework. Journal of Business Research, 116, 356–65. doi: 10.1016/j.jbusres.2020.01.022

Kalyatin V. 2022. Establishing of Subject of Rights to Intellectual Property Created with Use of Artificial Intelligence Law. Journal of the Higher School of Economics, 4, pp. 24–50.

Кетова Н.П., Зундэ В.В., Шашнев П.Д. 2020. Трансформация рынка рекламы как фактор активизации поведения рекламных агентств в условиях цифровизации. Креативная экономика, 14 (5), 953–968.

Maidanova S.A., Ilin I.V. 2022. Development of digital transformation strategy in the context of enterprise architecture. Technoeconomics, 2, 1 (4), 64-75.

Okorokov R., Timofeeva A., Kharlamova T. 2019. Building intellectual capital of specialists in the context of digital transformation of the Russian economy. In IOP Conference Series: Materials Science and Engineering, 497 (1), 012015. doi:10.1088/1757-899X/497/1/012015

Provost F., Fawcett T. 2013. Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media.

Spewak S., Hill S. 1992. Enterprise Architecture Planning: Developing a Blueprint for Data, Applications, and Technology.

Sykora M., Elayan S. 2022. The Power of Emotions: Leveraging User Generated Content for Customer Experience Management». Journal of Business Research, 144. doi: 10.1016/j.jbusres.2022.02.048

LTV: what it is, how to calculate the indicator of lifetime value, what it is needed for in marketing. URL: https://www.mango-office.ru/products/calltracking/glossary/ltv/ (accessed: 24.08.23).

Modeling enterprise architecture with Archi. URL: https://ekonomika.snauka.ru/2014/11/6308 (accessed: 24.08.23).

The Open Group Library. URL: https://publications.opengroup.org / (accessed: 25.08.23).

TOGAF. Www.Opengroup.Org. URL: https://www.opengroup.org/togaf (accessed: 24.08.23).

What is CRM: Definition of customer relationship management. B.D. SAP. 2023. URL: https://www.sap.com/central-asia-caucasus/products/crm/what-is-crm.html (accessed: 23.08.23).

INFORMATION ABOUT AUTHOR / ИНФОРМАЦИЯ ОБ АВТОРЕ

MOTYCHKO Veronika – analyst.
E-mail: veronikamotychko@gmail.com
МОТЫЧКО Вероника – аналитик.
E-mail: veronikamotychko@gmail.com

Статья поступила в редакцию 15.11.2023; одобрена после рецензирования 20.11.2023; принята к публикации 21.11.2023.

The article was submitted 15.11.2023; approved after reviewing 20.11.2023; accepted for publication 21.11.2023.

Technoeconomics. 2023, Vol. 2, No. 4 (7). Pp. 61–71.
Техноэкономика. 2023, Том 2, № 4 (7). С. 61–71.

Scientific article
UDC 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.6

APPLYING MACHINE LEARNING METHODS IN ELECTRONIC DOCUMENT MANAGEMENT SYSTEMS

Laylo Abdukhalilova[1], Oksana Iliashenko[2], Dayana Alchinova[2] ✉

[1] Tashkent State Transport University, Tashkent, Uzbekistan;
[2] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ alchinovadayana@gmail.com

Abstract. This article discusses methods and schemes for using machine learning in automating document management business processes. The scenarios described in the article can be useful for companies involved in document workflow automation or related areas (for example, email services) or ECM systems in general, and represent a generalization of the experience of specialists in the use of machine learning methods in document management. Several machine learning models used in business process automation solutions in practice are also considered. The results of the study, based on the analysis of all the points considered in the article, identified the main possible areas of development that arise when using machine learning models in electronic document management systems, which will be useful for data scientists developing such areas of AI, as machine learning.

Keywords: business processes, documentation support for management, machine learning, electronic document management systems, data analytics

Citation: Abdukhalilova L., Ilyashenko O., Alchinova D. Applying machine learning methods in electronic document management systems. Technoeconomics. 2023. 2. 4 (7). 61−71. DOI: https://doi.org/10.57809/2023.2.4.7.6

Научная статья
УДК 330.47
DOI: https://doi.org/10.57809/2023.2.4.7.6

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В СИСТЕМАХ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

Лайло Абдухалилова[1], Оксана Ильяшенко[2], Даяна Альчинова[2] ✉

[1] Ташкентский государственный транспортный университет, Ташкент, Узбекинстан;
[2] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉ alchinovadayana@gmail.com

Аннотация. В данной статье рассматриваются методы и сценарии применения машинного обучения в автоматизации бизнес-процессов документационного обеспечения управления. Описанные в статье сценарии могут быть полезны компаниям, занимающимся автоматизацией электронного документооборота или смежными областями (например, почтовые сервисы) или в целом ECM-системами, и являются обобщением опыта авторов в области использования методов машинного обучения в системах электронного документооборота. Также будут рассмотрены несколько моделей машинного обучения, используемые в решениях автоматизации бизнес-процессов в практике. В результатах исследования на основе анализа всех рассмотренных пунктов статьи выведены основные возможные зоны развития, возникающих при использовании моделей машинного обучения в системах электронного документооборота, что будет полезно дата-сайентистам, развивающим такого направления искусственного интеллекта, как машинное обучение.

Ключевые слова: бизнес-процессы, документационное обеспечение управления, машинное обучение, системы электронного документооборота, аналитика данных

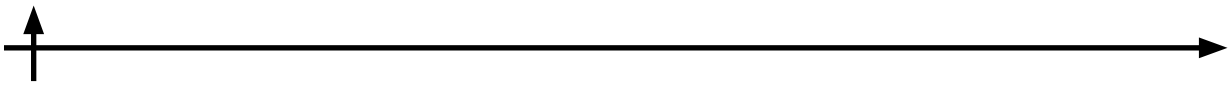
Для цитирования: Абдухалилова Л., Ильяшенко О., Альчинова Д. Применение методов машинного обучения в системах электронного документооборота // Техноэкономика. 2023. Т. 2, № 4 (7). С. 61–71. DOI: https://doi.org/10.57809/2023.2.4.7.6

Introduction

Modern business, as one of the priority tasks of strategic planning, announces process automation to maintain the company's sustainability in an increasingly competitive environment. One of the breakthrough technologies in solving automation problems is artificial intelligence (AI) technologies, including machine learning methods. Machine learning is a branch of computer science that allows systems to learn and improve their performance based on previous experience or history.

Machine learning is based on an algorithmic approach that allows systems to learn from data definitions and the resulting knowledge to solve problems at the first moment in time. Any AI algorithm is based on a mathematical model of varying degrees of complexity. A mathematical model is a representation of a system, process or phenomenon using mathematical explanations and results.

The purpose of this study is to study existing methods of using machine learning models in automating business processes related to document management support. The article uses an analytical approach based on an analysis of relevant literature and sources to consider the prospects for using intelligent microservices in a system for increasing document flow (EDMS). Scientific articles, books and electronic resources related to the use of AI and microservices to obtain documentation support for management have been studied. In addition, a practical



analysis of the results of research and implementation of intelligent microservices in the EDMS was carried out. The advantages and disadvantages of such a situation were explored, as well as issues of consideration, situation and depending on the quality of the data. The novelty of scientific research lies in the qualitatively new use of machine learning technologies to automate the processes of working with documents in the company's ecosystem by increasing the volume of routine operations performed by enterprise employees. The results of this study can help contracting companies implement EDMS, as well as improve the efficiency and effectiveness of management processes.

Materials and Methods

Tasks of an electronic document management system that can be modernized using machine learning:

1. Data storage

The task of storing data in electronic document management systems includes classification and selection of an acceptable storage method depending on the type of data. Machine learning models created for automatic classification will allow you to achieve fast and efficient document search by building connections between system objects. The model will detect classes of documents that have high value and require timely consideration first (Gogoulou, 2019). By combining the model with the notification service, the system will send such documents to users so they can start working quickly. This scheme allows the system to adapt the business process to changing conditions and makes it possible to achieve greater flexibility and stability.

2. Data processing

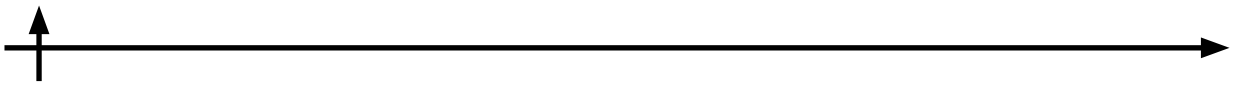
In an electronic document management system, objects have their own attributes to determine how data is stored, for example, data type, field type, section, etc. In cases where new objects appear in the system, a machine learning model that has been trained to recognize new object attributes and building a connection between these attributes and processes that use objects with similar attributes will enable the system to increase the speed of processing large arrays of different types and types of data (Iliashenko, 2020). The presence of such a goal in the company's strategy as system scalability from one process of one company to an entire consortium will contribute to the improvement of such models.

3. Data analysis

Analyzing data and making decisions based on this analysis is the main task of companies with a Data-driven approach. For such companies, the issue of using machine learning in data analysis will be relevant, since the effectiveness of using models is based on data on decisions made in existing successful cases, which will be embedded in the model, and which will serve as a model (teacher) for it (Devlin, 2019). The presence of such data in a company in an acceptable volume will allow the model to predict the consequences of decisions made with a high percentage of forecast quality. To understand the accuracy of the model, data scientists can increase the number of iterations of running these models. As with all decisions made, the consequences will not be immediately visible. Users of the system should remember this and not forget that ultimately, it is not the machine learning model that is responsible for the result, but the users themselves (Mogilko, 2020).

4. Generation of system objects

Neural networks, which are deep learning models, allow companies to minimize or completely eliminate the routine tasks of creating and working with objects. In relation to electronic document management systems, machine learning, namely deep learning, will be able to optimize the process of developing content and documents by generating memorized objects and their reproduction based on these objects of a system of the same type (Sambetbayeva, 2022).



Users only need to configure the generation of individual sections of documents according to specified rules.

5. Data security

Similar to the point on data processing, the system includes document flow, using machine learning, you can effectively manage documents with confidential information and personal data of the company's employees and its counterparties. Machine learning models built to automatically identify such information will classify documents containing confidential information and automatically determine access rights to them. To understand how machine learning models work with text and can be applied in various company cases, let's consider the stages of computer processing of texts in natural language:

1. Definition of linguistic resources – collection of text data (tagged and/or untagged) for processing and, if necessary, various kinds of dictionaries, for example, thesauruses in a form suitable and convenient for machine processing.

2. Cleaning text data – bringing texts to a general form without losing elements that affect the result of processing.

3. Graphematic analysis, segmentation or tokenization – breaking each text into units that are subject to further processing (tokens).

4. Morphological analysis:

– determination of the part of speech (part-of-speech, POS) and morphological features of the word, such as gender, number, case, declension;

– lemmatization is the process of transforming word tokens into the initial form, which in the case of processing Russian-language text means the representation of nouns in the nominative case and singular, adjectives, like nouns, but in the masculine gender, verbs, gerunds and participles in the indefinite form of the verb;

– stemming is the process of bringing word tokens to a base by cutting off suffixes and endings.

5. Exclusion of stop words, or "noise words" – removal of tokens representing frequently used but meaningful words (prepositions, conjunctions, parts, interjections, introductory words, some verbs, pronouns) (Popova, 2019).

6. Vectorization – consideration of each token in the text of a small numerical value.

The result of the steps is a representation of the source texts, suitable for further processing in order to solve a higher-level problem, for example, a text classification problem. Further processing of the prepared texts is carried out on the basis of two approaches: engineering and based on machine learning.

Results and Discussion

The engineering approach is to use rules compiled by linguists in the form of dictionaries, templates and other linguistic sources. Machine learning approaches use natural language texts as resources to build a mathematical model that can exhibit features in earlier samples of the data presented. Based on the identified features, the algorithm determines feature values for new data. Let's consider additional machine learning models:

Natural language processing (NLP) is an interdisciplinary field of computer science and linguistics. It's primarily about giving computers the ability to support and manipulate human language. It involves processing natural language datasets such as text or speech corpora using either rule-based or probabilistic machine learning approaches. The goal is a way to "understand" computer requirements, including the contextual subtleties of the language within them. The technology can then pinpoint the information and ideas contained in the documents, as well as categorize and organize the documents themselves (Cherkasov, 2018; Anufrieva, 2023).

To solve NLP problems, an engineering approach is being used less and less and increasingly based on machine learning. But the first one should not be neglected, since it is effectively "applicable to narrow subject areas with clear rules for naming significant objects and a small variety of required language constructs" (Ponomarev, 2023).

BERT model

The transfer learning approach represented by Bidirectional Encoder Representations from Transformers is a recent language representation model. BERT is designed to pre-train deep bidirectional representations to extract context-sensitive features from input text. These representations fall under the category of "embeddings," which is an important concept in the field of NLP (Ilyin, 2023).

The term embedding refers to fixed-length vector representations of text that are capable of encoding syntactic and semantic information. BERT embeddings can be successfully used for any NLP tasks such as language inference and name and object recognition.

What makes the BERT model powerful is that high accuracy is achieved without any specific network architecture at the fine-tuning level. The model consists of a sequence of embedding layers, transformers, and a prediction layer (Figure 1). Given the incredibly good performance of BERT embedding in capturing input context, it is expected that the size of the text completing the BERT model, with a conversation history encoding mechanism, can lead to a high-performance conversational machine model understanding (Obukhov, 2020).

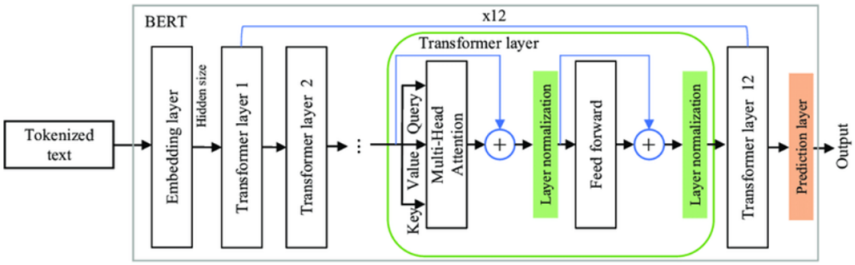


Fig. 1. BERT model architecture and its components

Model ELMo

The ELMo (Embeddings from Language Models) model takes context into account before embedding a word. This is how contextualized word-embeddings appeared (Davletov, 2023). Contextualized embeddings give words different vectors based on their semantics in the context of a sentence. Instead of using fixed word embeddings, ELMo looks at the entire sentence before assigning each word its embedding. It uses a bi-directional long short-term memory model (bi-directional LSTM), trained specifically for the task of creating such embeddings (Castelblanco, 2020).

As we train the model on a large data set, we begin to learn language constructs. ELMo creates contextualized embedding by grouping latent effects (other embedding) in this way (concatenation follows weighted addition) (Figure 2).

Let's look at the main scenarios for using machine learning.

Scenarios for the use of machine learning models largely depend on the implementation processes of management information support and EDMS components (Bolshakova, 2017).

The electronic document management system consists of the following elements:

– a scan of the document included in the system on paper;

– document registration card (DC), displaying information about the document, namely information about the sender, counterparty, subject of the document, date of sending, etc.;

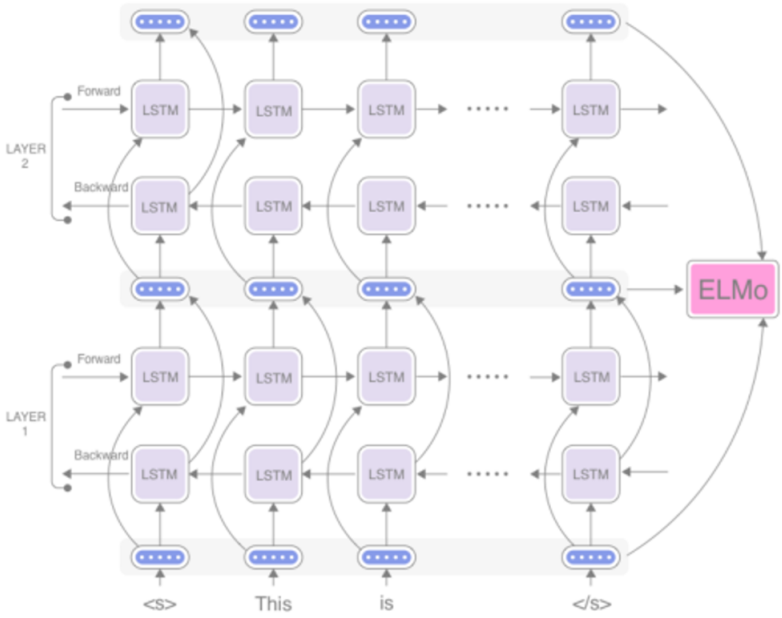


Fig. 2. ELMo model architecture

– information about the system user who works in the document's RC (Registrar) (Chernov, 2016).

Let's look at how machine learning models are used in the medium term and involve document flow.

1. Automatic routing of documents based on topic tagging can be represented in the form of the following script steps (Fig. 3):

1) maintaining a directory of thematic organization, where the responsible department of the employee is responsible for each topic;

2) highlighting semantic tags in the document text;

3) correlation of semantic tags with thematic ones (calculation of the probability of assignment, ordering by probability);

4) identifying the artist by tag.

The results of automatic routing of documents based on thematic tagging are the acceleration of delivery of the document to the department/performer, eliminating the processing of routine correspondence by the manager (Ivanovsky, 2021).

Examples of extracted entities are "Counterparty", "Related Documents and Instructions".

Results:

1. Related information on selected entities. For example, searching for correspondence with a selected counterparty.

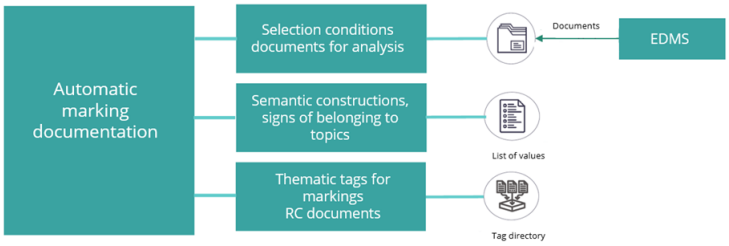


Fig. 3. Diagram of components and connections between them used in automatic routing

2. Automatic formation of links between documents.

In 15% of documents, on average, connections are not made at the registration stage.

Scenario stages:

1) registration of an incoming document, attaching files;
2) recognition of file scans;
3) analysis of the text of files, search for data from related documents;
4) search for related documents in the EDMS;
5) establishing connections between documents.

Outcome: the ability to move to other system objects with a logical connection.

3. Automatic creation of an object in the system based on a document included in the system (Fig 4).

Scenario stages:

1) scanning an incoming document - transferring information into the system;
2) creation of the RC document;
3) text processing in the document;
4) automatic filling of document RC attributes.

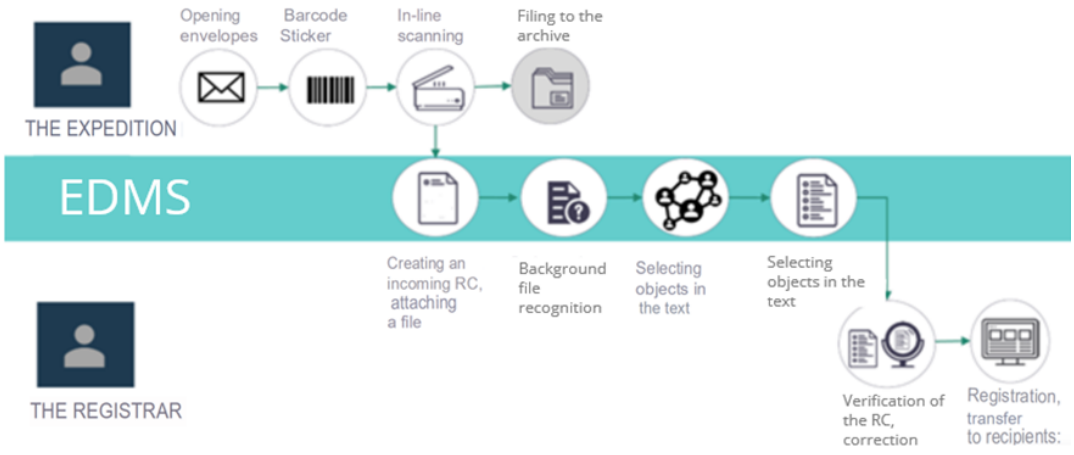


Fig. 4. Scheme of the process of automatically entering an incoming document into the system

4. PDF OCR

Optical Character Recognition is the process of processing and recognizing images or videos containing text content, as well as extracting text and layout of the information contained in them. OCR is most used to convert a non-editable PDF document file containing an image of the document into an editable Word document (Iliashenko, 2022).

Modern OCR systems use intelligent character recognition (ICR) technology to read text just like a human would (Figure 5). They use advanced machine learning techniques for human reading skills. A machine learning system called a neural network analyzes text at many levels by repeatedly processing the image. It looks for various image attributes (curves, lines, intersections and loops) and combines the results of different levels of analysis to produce a final result.

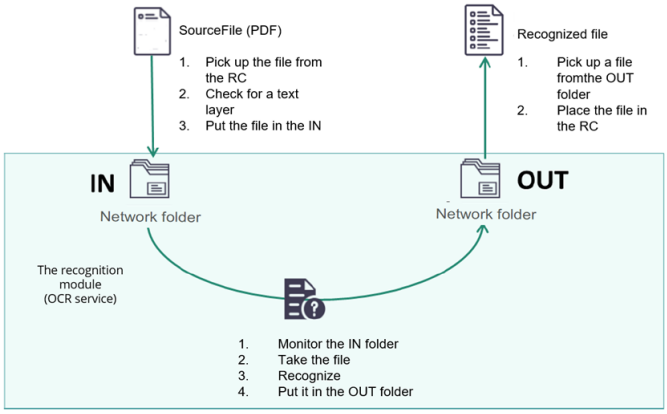


Fig. 5. Architecture of the OCR module in the document management system

Conclusion

The areas of development that arise when using machine learning models in electronic document management systems are the following:

1. Enriching the base of customers who have a need to solve business cases. Improvement of supervised machine learning models occurs at the expense of existing data, based on which the model can be calibrated, since there are real "prediction" values.

2. Expanding the pool of datasets for conducting experiments. Machine learning models must run through several epochs and iterations. To do this, you initially need a large array of data, on the basis of which new and varied data can be generated.

3. Development of the IT infrastructure of companies, namely IT capacities. During the time spent training the model, new data appears in living systems that can make changes to the model. For such maneuvers, companies must have the necessary capacity.

Integration of machine learning technologies into electronic document management systems will give companies the opportunity to solve problems of effective work with documents, from reorganization in the file structure to provide direct access to information to automated decision-making on this document. Electronic document management systems that integrate artificial intelligence technologies will have a positive impact on the company's financial policy in the future. Every organization in the future will save money by replacing routine operations performed by staff with an ensemble of digital digital services using AI technologies. To effectively integrate machine learning into an EDMS, a company must define the goals it wants to achieve through this merger. In addition, the company must have qualified specialists to select machine learning models for its specific task, because training a model requires resources.

The use of machine learning methods in electronic document management systems can significantly improve the efficiency and accuracy of document processing and analysis processes, which helps improve the productivity and competitiveness of enterprises.

REFERENCES

Anufrieva V. 2023. Improvement of the system of working with suppliers at enterprises of the power engineering sector. Technoeconomics, 2, 3 (6), 69–78. DOI: https://doi.org/10.57809/2023.2.3.6.6

Bolshakova E. I., Vorontsov K. V., Efremova N. E., Klyshinsky E. S., Lukashevich N. V., Sapin A. S. 2017. Automatic processing of texts in natural language and data analysis: textbook, 269.

Castelblanco A. 2020. Machine learning techniques for identity document verification in uncontrolled environments: A case study. Lecture Notes in Computer Science, 12088, 271-281.

doi:10.1007/978-3-030-49076-8_26

Cherkasov D.Yu., Ivanov V.V. 2018. Machine learning. Science, technology and education, 5 (46).

Chernov T.S., Ilyin D.A. 2016. Investigation of methods of image segmentation of text blocks of documents using structural analysis and machine learning algorithms. Bulletin of the Russian Foundation for Basic Research, 4(92), 55-71. doi:10.22204/2410-4639-2016-092-04-55-71

Davletov A. R. 2023. Modern methods of machine learning and OCR technology for automation of document processing. Bulletin of Science, 5 (67), 676-698.

Devlin J., Chang M.-W., Lee K., Toutanova K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

Gogoulou E. 2019. Using Bidirectional Encoder Representations from Transformers for Conversational Machine Comprehension. KTH Royal Institute Of Technology, 78.

Iliashenko O.Yu., Iliashenko V.M., Bikkulova Z.U. 2020. Artificial intelligence as a tool with an observable process of providing medical care. In the collection: System analysis in design and management. collection of scientific papers of the XXIV International Scientific and Educational-Practical Conference, 272-284.

Iliashenko V.M. 2022. Machine learning technologies in the oil and gas industry. In the collection: Increasing the managerial, economic, social and innovative-technical potential of enterprises, industries and national economic complexes. Collection of articles of the XIII International Scientific and Practical Conference, 116-119.

Ilyin I.V., Ilyashenko O.Yu., Devezas T.S., Shchenikov E.M. 2023. Forecasting electricity consumption using machine learning methods. Ecology, environmental protection, carbon neutrality and development. Collection of proceedings of the Chinese-Russian ASRTU Symposium, 36-38.

Ivanovsky N. I. 2021. Application of machine learning algorithms for classification of documentation. Vestnik VNIIDAD, 2, 35-39.

Mogilko D.Y., Ilin I.V., Iliashenko V.M., Svetunkov S.G. 2020. BI capabilities in a digital enterprise business process management system. Lecture Notes in Networks and Systems, 95, 701-708.

Obukhov A. 2020. Algorithm of adaptation of electronic document management system based on machine learning technology. Progress in Artificial Intelligence, 9 (4), 287-303. doi:10.1007/s13748-020-00214-2

Ponomarev O.V., Svetunkov S.G. 2023. Economics as a science: an explanatory dictionary of general scientific terms and concepts: textbook, 214.

Popova E. S. 2019. Using artificial neural networks to solve text classification problems. Proceedings of the 29th international conference on computer graphics and vision “Graphicon”, 270–273.

Sambetbayeva M., Kuspanova I., Yerimbetova A., Serikbayeva S., Bauyrzhanova S. 2022. Development of intelligent electronic document management system model based on machine learning methods. Eastern-European Journal of Ente rprise Technologies, 1 (2 (115), 68–76.

How transfer learning came to NLP. URL: https://habr.com/ru/articles/487358/ (accessed: 12.07.23).

Optical character recognition. URL: https://en.wikipedia.org/wiki/Optical_character_recognition (accessed: 12.07.23).

What is Optical Character Recognition (OCR): Overview and Use Cases. URL: https://www.labellerr.com/blog/optical-character-recognition-overview/ (accessed: 13.07.23).

What are the advantages of Optical Character Recognition? URL: https://makemeanalyst.com/what-are-the-advantages-of-optical-character-recognition/ (accessed: 14.07.23).

СПИСОК ИСТОЧНИКОВ

Anufrieva V. 2023. Improvement of the system of working with suppliers at enterprises of the power engineering sector. Technoeconomics, 2, 3 (6), 69–78. DOI: https://doi.org/10.57809/2023.2.3.6.6

Bolshakova E. I., Vorontsov K. V., Efremova N. E., Klyshinsky E. S., Lukashevich N. V., Sapin A. S. 2017. Automatic processing of texts in natural language and data analysis: textbook, 269.

Castelblanco A. 2020. Machine learning techniques for identity document verification in uncontrolled environments: A case study. Lecture Notes in Computer Science, 12088, 271-281. doi:10.1007/978-3-030-49076-8_26

Cherkasov D.Yu., Ivanov V.V. 2018. Machine learning. Science, technology and education, 5 (46).

Чернов Т.С., Ильин Д.А. 2016. Исследование методов сегментации изображений текстовых блоков документов с помощью алгоритмов структурного анализа и машинного обучения. Вестник Российского фонда фундаментальных исследований, 4(92), 55-71. doi:10.22204/2410-4639-2016-092-04-55-71

Давлетов А. Р. 2023. Современные методы машинного обучения и технология OCR для автоматизации обработки документов. Вестник науки, 5 (67), 676-698.

Devlin J., Chang M.-W., Lee K., Toutanova K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

Gogoulou E. 2019. Using Bidirectional Encoder Representations from Transformers for Conversational Machine Comprehension. KTH Royal Institute Of Technology, 78.

Iliashenko O.Yu., Iliashenko V.M., Bikkulova Z.U. 2020. Artificial intelligence as a tool with an observable process of providing medical care. In the collection: System analysis in design and management. collection of scientific papers of the XXIV International Scientific and Educational-Practical Conference, 272-284.

Iliashenko V.M. 2022. Machine learning technologies in the oil and gas industry. In the collection: Increasing the managerial, economic, social and innovative-technical potential of enterprises, industries and national economic complexes. Collection of articles of the XIII International Scientific and Practical Conference, 116-119.

Ilyin I.V., Ilyashenko O.Yu., Devezas T.S., Shchenikov E.M. 2023. Forecasting electricity consumption using machine learning methods. Ecology, environmental protection, carbon neutrality and development. Collection of proceedings of the Chinese-Russian ASRTU Symposium, 36-38.

Ивановский Н. И. 2021. Применение алгоритмов машинного обучения для классификации документации. Вестник ВНИИДАД, 2, 35-39.

Mogilko D.Y., Ilin I.V., Iliashenko V.M., Svetunkov S.G. 2020. BI capabilities in a digital enterprise business process management system. Lecture Notes in Networks and Systems, 95, 701-708.

Obukhov A. 2020. Algorithm of adaptation of electronic document management system based on machine learning technology. Progress in Artificial Intelligence, 9 (4), 287-303. doi:10.1007/s13748-020-00214-2

Ponomarev O.V., Svetunkov S.G. 2023. Economics as a science: an explanatory dictionary of general scientific terms and concepts: textbook, 214.

Popova E. S. 2019. Using artificial neural networks to solve text classification problems. Proceedings of the 29th international conference on computer graphics and vision "Graphicon", 270–273.

Sambetbayeva M., Kuspanova I., Yerimbetova A., Serikbayeva S., Bauyrzhanova S. 2022. Development of intelligent electronic document management system model based on machine learning methods. Eastern-European Journal of Ente rprise Technologies, 1 (2 (115), 68–76.

How transfer learning came to NLP. URL: https://habr.com/ru/articles/487358/ (accessed: 12.07.23).

Optical character recognition. URL: https://en.wikipedia.org/wiki/Optical_character_recognition (accessed: 12.07.23).

What is Optical Character Recognition (OCR): Overview and Use Cases. URL: https://www.labellerr.com/blog/optical-character-recognition-overview/ (accessed: 13.07.23).

What are the advantages of Optical Character Recognition? URL: https://makemeanalyst.com/what-are-the-advantages-of-optical-character-recognition/ (accessed: 14.07.23).

INFORMATION ABOUT AUTHORS / ИНФОРМАЦИЯ ОБ АВТОРАХ

ABDUKHALILOVA Laylo T. – professor, Candidate of Economic Sciences.
E-mail: laylo.66@mail.ru
АБДУХАЛИЛОВА Лайло Тухтасиновна – профессор, к.э.н.
E-mail: laylo.66@mail.ru

ILIASHENKO Oksana Yu. – Associate Professor, Candidate of Pedagogic Sciences.
E-mail: ioy12@yandex.ru
ИЛЬЯШЕНКО Оксана Юрьевна – доцент, кандидат педагогических наук.
E-mail: ioy12@yandex.ru
ORCID: https://orcid.org/0000-0001-9292-7759

ALCHINOVA Dayana Yu.– business analyst.
E-mail: alchinovadayana@gmail.com
АЛЬЧИНОВА Даяна Юрьевна– бизнес-аналитик.
E-mail: alchinovadayana@gmail.com
ORCID: https://orcid.org/0009-0005-0963-4506

Статья поступила в редакцию 15.11.2023; одобрена после рецензирования 26.11.2023; принята к публикации 27.11.2023.

The article was submitted 15.11.2023; approved after reviewing 26.11.2023; accepted for publication 27.11.2023.

Научное издание

Technoeconomics

Том 2, № 4, 2023

Учредитель, издатель – Федеральное государственное автономное образовательное учреждение высшего образования
«Санкт-Петербургский политехнический университет Петра Великого»

Телефон редакции 8 (812) 550-36-52

E-mail: technoeconomics@spbstu.ru

Компьютерная верстка *Д.М. Гугутишвили*
Редактирование английского языка *И.В. Ильина*
Ответственный секретарь *О.В. Воронова*
Выпускающий редактор *А.И. Лёвина*